Understanding variations in the outcome of a smoking cessation programme in Tuberculosis patients in Pakistan

- the role of fidelity?

OMARA FATIMA DOGAR

DOCTOR OF PHILOSOPHY

UNIVERSITY OF YORK
HEALTH SCIENCES

SEPTEMBER 2016

ABSTRACT

I conducted a review to show that smoking cessation in TB patients can have positive TB outcomes. Behavioural interventions (BI) for smoking cessation are known to be effective and cost effective but there are important variations in quit rates that are hard to explain. My aim was to develop and validate a fidelity index for BIs in smoking cessation, and use it to explore the extent to which variations in smoking quit outcomes reflect the degree to which providers implement the various components and the quality of delivery of a smoking cessation programme in TB treatment settings.

I developed and tested a theoretically mapped fidelity index comprising two sub-indices: 'Adherence' for measuring compliance (37 items) and 'Quality' for measuring provider competence (8 items). Items were rated as fully, partially, or not, implemented against a behaviourally anchored response scale. Fidelity was measured in a prospective study of 18 providers in TB clinics in Pakistan (154 patients) whose sessions were audio-recorded and then coded. These providers had participated in delivering the same BIs as part of the ASSIST trial four years earlier.

Reliability was assessed in three ways. There was good inter-coder reliability using Kripendorff's alpha, Principal Components Analysis showed the items of the index were coherent in measuring fidelity and Generalisability theory showed that the index reliably differentiated between providers by capturing the variation in their BI delivery practice.

Provider Adherence and Quality of provision were positively correlated. I tested the assumption that relative provider practice was consistent between the ASSIST trial in 2010 and the fidelity study in 2014. Using Kendall's W coefficient of concordance on data from a self-recorded checklist used in both studies, I found moderate to strong concordance. I then used binomial regression analysis to estimate the relationship between fidelity and ASSIST trial provider-level quit rates. This showed that the provider-level quit rate was positively associated with Quality of interaction (odds ratio: 2.15; 95% CI, 1.43 to 3.24) but negatively associated with Adherence to BI content (odds ratio: 0.55; 95% CI, 0.40 to 0.77). A negative interaction was found between Adherence and Quality and quit rates.

This research makes several contributions to the field. We have a better understanding of the impact of smoking on TB. I developed and validated a new, theoretically-informed, fidelity index which can be used to quantify and score delivery of BI ingredients in a standardised way and used to better understand how BIs influence outcomes. I report that the quality of delivery of BIs maybe as important as content in influencing smoking cessation.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF APPENDICES

ACKNOWLEDGEMENTS

I would like to express my gratitude to my supervisors, Professor Trevor A. Sheldon and Dr Kamran Siddiqi, for their continuous support throughout my research. Their patience, enthusiasm and immense knowledge kept me motivated in these few years. I could not have asked for a better mentoring for my PhD study and career path.

I would also like to thank members of my thesis advisory panel: Professor Ian Watt for his support, encouragement and insight; Dr Mona Kanaan for advising me on advanced statistical techniques; and Dr Cath Jackson for her encouragement, constructive critique, and brilliant suggestions.

I would also like to thank Dr Jan R. Böhnke for the countless afternoons spent teaching me psychometrics and Professor Martin Bland for his helpful insights on sample size considerations and statistical analyses.

This thesis was completed with the support of a studentship by the Department of Health Sciences- University of York, for which I am truly grateful.

My sincere thanks go to Dr Riaz Ahmad and Iqbal Mahmood for facilitating me in data collection for my research in Sargodha and Jhang districts of Punjab, Pakistan. In addition, I would like to thank the DOTS facilitators that participated in my research. I would also like to thank Dr Amina Khan and Qadeer Ashraf for the long hours spent coding audiotapes for my study.

I thank my friend (Dr Sarwat Shah) for listening to my ideas, for her kind words, support and encouragement towards finishing my thesis. I am grateful to Dr Najma Siddiqi for guiding me in writing up my thesis. I thank Dr Anne Readshaw for proofreading my thesis.

Last but not the least; I am ever so grateful to my mother (Khalida Adeeb Khanum) for being the inspiration in my life.

AUTHOR'S DECLARATION

I hereby declare that I am the sole author of this thesis and that this thesis is original and has not been previously published or submitted for a degree in this or any other institution. To the best of my knowledge this thesis does not contain material previously published or written by another person except where due reference or acknowledgement is made in the text.

I declare that the author solely undertook all the data collection and analyses. I further confirm that data collection that involved patients and DOTS facilitators has been conducted with the ethical approval of all relevant bodies and such approvals are acknowledged throughout this thesis.

I confirm that there are no known conflicts of interest associated with this thesis and there has been no financial support for this work that could have influenced its outcomes.

PART I- BACKGROUND

This section provides the background and context for the thesis. This discusses the significance of smoking cessation in high tuberculosis (TB) burden countries, implementing cessation interventions within existing TB care settings, variation in patient quit rates across implementation sites and the sources that potentially contribute to this variation; measuring fidelity to intervention delivery and its utility in explaining variation in outcomes.

I present 4 chapters in this section:

Chapter 1 provides an introduction to the thesis and outlines the focus of my research. This includes an introduction to important terms, such as fidelity, and describes complex interventions used in the thesis and their conceptualisation in terms of smoking cessation; and a brief overview of fidelity as a method for exploring variation and how behaviour change science contributes to it. I end the chapter by presenting an outline of the thesis and overview of the studies planned.

Chapter 2 provides literature on the association between smoking and TB outcomes. Furthermore, it introduces behavioural intervention as a cost-effective alternative to pharmacotherapy for cessation services in low-and-middle income countries (LMIC). It states the relevance of integrating such interventions within national TB programmes and supplies findings from a large cluster Randomised Controlled Trial (RCT) highlighting the implications of variation in smoking quit rates on the wider implementation of complex BIs.

Chapter 3 describes the sources of variation in the outcomes of the ASSIST study, using a framework to characterise them by: i) factors intrinsic to the intervention, such as its substantive features, ii) methods of evaluation, such as cluster RCT and iii) factors extrinsic to the intervention such as research context. I end the chapter by elaborating further on sources of variation due to intrinsic factors affecting the intervention.

Chapter 4 provides an insight into the elements of fidelity that need consideration for explaining the mechanism of impact of complex BIs on their outcomes. The literature reviewed pertains mainly to measurement of fidelity and its conceptualisation in terms of complex BIs, particularly focusing on identification of the active features of the intervention.

Chapter 1. Introduction

Smoking behaviour is common in countries with high TB burden and adversely impacts on the health of people living in these Low-and-Middle Income Countries (LMICs) (Siddiqi, 2014, Dogar et al., 2013). Implementation of smoking cessation interventions within national TB programmes is of utmost relevance to these countries (Dogar et al., 2013). Behaviour change interventions (or BIs) are a low-cost alternative to pharmacotherapies (alone) for smoking cessation in low resource countries. However, evaluations of these interventions are often limited in their description of what works and how it works.

RCTs evaluating such 'complex interventions' may provide evidence of their overall effectiveness in achieving smoking cessation, but the interventions being tested are frequently described as a 'black box', with little information about individual features (Craig et al., 2013). They do not indicate how the intervention works- what the 'active ingredients' are, and how they exert their effect (Lorencatto et al., 2013c). Neither do they attempt to explain the variation in effect within or between implementation sites. In order to identify and quantify the active ingredients of an intervention, first a fuller description of individual features of the intervention is needed (Lorencatto et al., 2013b). We then need a 'measure' of the delivery or implementation of these features, which would not only be reliable in measuring this construct (fidelity of delivery) under study but would also have the potential to explain variation in quit rates. This measure could help optimise intervention delivery for wider implementation, by facilitating identification of the active intervention ingredients and their respective effect on quit rates.

My aim in this thesis is to design a method for quantitative measurement of fidelity and then to use these quantitative scores to explore variation in smoking quit rates. This requires developing a measure that is reliable and takes into account the complexity of BIs. Developing a measure for an intervention with multiple facets requires wide-ranging knowledge, not only on scale construction but also on ways to capture intervention ingredients, aspects of patient-provider interaction and linkage with behaviour change in the patient, in a quantifiable manner. Accurately recording behaviour change and its underpinning causal mechanisms is challenging and requires a range of strategies. Process evaluation can give valuable insights into the causal mechanisms underlying the success of complex BIs and how they can be optimised (Craig et al., 2013).

I chose, therefore, to use the Medical Research Council (MRC) framework for process evaluation of complex interventions (Moore et al., 2015) to guide the methodology for examining variability in patient outcomes, based on provider fidelity of a BI. There is no single approach for carrying out this piece of work described in the MRC guidance. However, the guidance recognises fidelity as one of the key dimensions of implementation i.e. the quality and quantity of what is actually delivered as part of the intervention. It provides a framework linking fidelity to the broader functions of process evaluation, such as mechanism of impact and contextual factors. Although

this framework could not be applied in its entirety to quantify fidelity, some of its aspects could help design such a measure. For example, it describes approaches to delineate complex interventions in terms of active ingredients and how these act on behavioural determinants, introducing the use of logic models. It also introduces frameworks and methods for assessing fidelity that could be applied to BIs. I will describe some of these approaches in further detail here, as it helps explain the structure and process used in my thesis.

1.1 COMPLEX BEHAVIOURAL INTERVENTIONS

An intervention may be 'complex' in terms of the number of features, the nature of interactions between its features, challenges in its implementation, and how it interacts with its contexts (Craig et al., 2013). Complex interventions are more than the sum of their parts (Hawe et al., 2004). An intervention is more likely to be complex when it is difficult to precisely define its 'active ingredients' and how they relate to each other (Medical Research Council, 2000).

Complexity is a scientific concept, which asserts that some interventions comprise behavioural phenomena that are more complex than just the constituent parts of that intervention. Interventions aimed at changing behaviour add to their complexity in implementation, by the number and difficulty of behaviours (and skills) required on the part of the provider and the recipient of the intervention (Craig et al., 2008).

The complexity of BIs is determined, in part, by the number of features involved. Features include, the Behaviour Change Techniques (BCTs) or 'active ingredients', defined later in section 1.3.2, to facilitate behaviour modification, as well as procedures for delivery of the BCTs. Procedures for delivery include who delivers the intervention, to whom, how often, for how long, in what format, and in what context. The competences required to deliver the BCTs also contribute to the complexity of such interventions. To undertake specification of an intervention into its constituent ingredients, it is important to first understand all of its dimensions and its theoretical basis.

1.2 DELINEATING BEHAVIOURAL INTERVENTIONS

According to the MRC framework (Moore et al., 2015), a clear description of the causal assumptions of an intervention is a vital step in framing the pathways linking intervention to outcomes. Michie and colleagues (2009) argue that adoption of a uniform description of complex BIs may help evaluators in achieving this. In subsequent sections, I will present the behaviour change concepts, terminologies and their descriptions, including a logic model of behaviour change. This work is broadly based on behaviour change for smoking cessation.

1.2.1 Behaviour Change Theories

Theories that explain behaviour help to identify the determinants that can influence it to change (Michie et al., 2008) e.g. the theory of planned behaviour (Ajzen, 1991), health belief model (Green and Murphy, 2014), and the theory of reasoned action (Fishbein, 1979). These theories support a range of constructs which affect behaviour including, intention, self-efficacy, anticipated outcomes/attitude, norms (Fishbein et al., 2001), knowledge, skills, belief about consequences and capabilities, motivation and action planning (Michie and Abraham, 2004). Other theories help explain the process of change and the importance of tailoring interactions with patients e.g. the trans-theoretical (stages of change) model (Prochaska and DiClemente, 1984) and the social cognitive theory (Bandura, 1997).

Prochaska and DiClemente (1984) recommend assessing patient readiness to act on a new healthier behaviour, and providing strategies to guide them through the stages of action and maintenance of their changed behaviour. In this model, individuals can be characterised as belonging to one of the five 'stages', namely 'pre-contemplation', 'contemplation', 'preparation', 'action' and 'maintenance'. Although widely used, more recently this 'stages of change' model has received criticism for not being founded on evidence (West, 2005); mainly because the concept of 'stage' draws arbitrary dividing lines between stages which are not real, and represent traits of individuals rather than their 'states' of mind. Secondly, this model focuses on conscious decision-making, neglecting the fact that unhealthy habit patterns become entrenched and semi-automated through repeated reward and punishment - processes that fall outside conscious awareness and do not follow decision-making rules (West, 2005). Finally, the model gives no consideration to the concept of addiction, which is clearly an important determinant of behaviour change when it comes to behaviours like smoking (West, 2005).

All these constructs covered by different theories envelop the determinants of behaviour change. Some are described by overarching concepts, others by completely separate concepts and still others by similar concepts but different terminologies (Michie et al., 2008). This disparity in the classification of behavioural determinants limits the conceptualisation of activities within a BI in terms of causal assumptions underlying change in a particular behaviour.

1.2.2 Recent Advances in Behaviour Change Science

Recent progress in the science of behaviour change (Michie et al., 2009b, Abraham and Michie, 2008) has helped identify key behavioural determinants of healthy life-style changes including smoking cessation and the techniques that could influence these. Michie and Abraham (2008) identified determinants that could be targeted for modifying behaviour, e.g. knowledge, skills, attitudes, belief about capabilities, social influences etc. Similarly, the BCTs, defined in section 1.3.2, that have been established to address these determinants, show how to target particular behaviours, e.g. by providing information on consequences, setting goals, coping with triggers

and withdrawal symptoms, reward etc. (Michie and Johnston, 2012, Michie et al., 2011c). Both behavioural determinants and techniques that act on changing them can prove useful in disentangling ingredients within a complex BI and describing its activities.

1.3 THE 'LOGIC MODEL' OF BEHAVIOUR CHANGE

Modelling theory (commonly known as causal or logic modelling) involves hypothesising the concepts of what is targeted (for example, the behavioural determinants of smoking) and how these are targeted (Michie et al., 2008). A generic model linking behavioural determinants, causally through behaviour to physiological or biochemical measures, and eventually to health outcomes has been proposed (Hardeman et al., 2005). Michie et al. (2008) adapted this model (Figure 1.1) to explain behaviour change by identifying the behavioural determinants (Cane et al., 2012, Michie et al., 2008), the BCTs that are likely to be effective (Michie et al., 2009a, Michie et al., 2011b, Michie et al., 2011c, West et al., 2010) and link between the two (Michie et al., 2014–2017). In the illustration (Figure 1.1), each blue arrow represents a causal process and interventions are targeted at changing these causal processes (Hardeman et al., 2005). I added the blue circles to show the overarching work on behaviour change in the context of smoking cessation that has emerged in recent years.

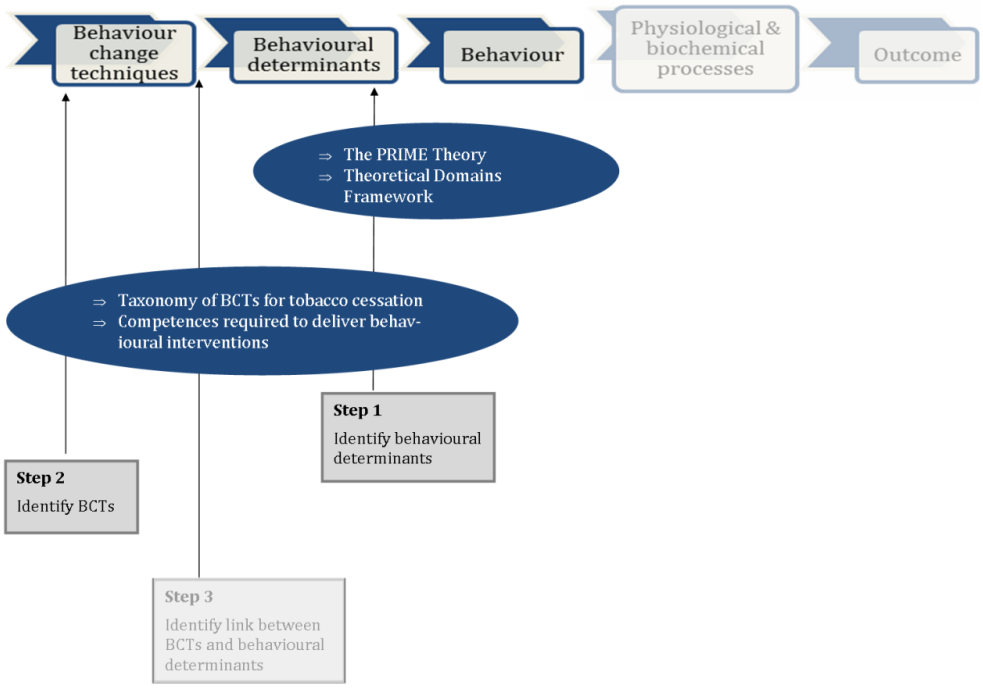


Figure 1.1: The logic model of behaviour change adapted for smoking cessation

*Adapted from: Hardeman, et al (2005), and Michie et al. (2008)

Subsequent sections provide an insight into the behavioural determinants of smoking cessation (Michie et al., 2011c), using the PRIME theory of motivation, (that provides a model to

encapsulate mechanisms in connection to tobacco dependence behaviour (West, 2009)) and the theoretical domains framework (TDF), (that provides a classification and categorisation method for behavioural determinants (Cane et al., 2012, Michie et al., 2005)). In these sections, I also list the relevant BCTs likely to influence these determinants (Michie et al., 2011b, Michie et al., 2011c, West et al., 2010).

1.3.1 Behavioural Determinants

Determinants identified by the TDF are generally applicable to healthy life-style behaviours (Cane et al., 2012). Determinants specific to smoking cessation that are used for developing the taxonomy of BCTs (Michie et al., 2011c) are conceptually derived from the PRIME theory (West, 2009). Both the TDF and the PRIME theory of motivation are briefly discussed, to shed light on their key parameters, as this is important for later descriptions of the BI in developing a fidelity measure.

Theoretical Domains Framework

Simplification of theoretical constructs based on the fact that theory can predict behaviour (Fishbein et al., 2001) shows good agreement on important behavioural determinants (Michie et al., 2008). However, overlapping theoretical constructs between theories hamper the identification of techniques specifically targeting individual determinants. The TDF was developed to categorise similar theoretical constructs into behavioural determinants and label them across all relevant theories of behaviour change.

Initially, the TDF comprised 12 behavioural determinants extracted from repeated evaluations of several theories (Michie et al., 2005). It was further validated for the number of extracted determinants, the constructs represented by each determinant, and its label (Cane et al., 2012). The revised framework consists of 14 determinants and 84 constructs including, but not limited to; knowledge, skills, identity, beliefs about capability, beliefs about consequences, intentions, goals and behavioural regulation.

The PRIME Theory of Motivation

Determinants critical to behaviour change for smoking cessation derive from the TDF, based on the PRIME theory of motivation (Michie et al., 2011c). Behavioural determinants of smoking and cessation are classified broadly into psychological (attitudes and beliefs), physiological (nicotine dependence) and social (normative behaviour). Nicotine dependence has multiple facets and requires understanding of the person's intention and motivation to change their behaviour (West, 2009). Dependence or addiction is a result of powerful motivation to carry out certain activities

repeatedly, to an extent that can be harmful and is often accompanied by impaired self-control (West, 2009).

Motivation means a reason for acting in a particular way (Oxford Dictionaries, 2015) i.e. deciding to do or not to do a certain activity, based on the known consequences. However, behaviour is often driven by habit or impulse and results in responding without thinking about the consequences. Actions of a particular behaviour might be merely driven by desire, ruling out the sense to make appropriate judgment about the best options (Mook, 1996). These actions often rely on instinctive thinking based on a previous idea or set rules, so that the brain does not have to think about the consequences each time (Mook, 1996). Enhancing a person's capability to achieve self-control based on self-imposed rules can lead to targeted behaviour change (West, 2009).

The *PRIME Theory of motivation* (West, 2009) provides a model to counter 'dependence behaviour' through stimulus-impulse associations, past experiences of pleasure, feelings of desire, beliefs about what is good or bad, self-conscious plans, rules or intentions and the ability for self-control. A key feature of this theory is that behaviour at each moment arises from the strongest of competing impulses at that moment and feelings of anticipated pleasure, satisfaction, or relief. These competing impulses and feelings are driven by past beliefs about what is good or bad, which in turn results from self-conscious intentions (West, 2009, West et al., 2010).

Behaviour change requires self-control, which requires generating sufficiently strong needs and wants from self-conscious intentions to overcome competing impulses (West, 2009). According to PRIME Theory, this involves a sense of self that creates 'rules,' which can generate motives from past experiences and beliefs about what is good or bad. For example, the belief that smoking is harmful does not usually in itself stop someone from smoking, but occasionally it causes the smoker to want self-consciously to create a rule not to smoke, which then generates the need not to smoke. While exercising self-control, the stronger the impulse against which it has to compete, the greater will be the effort required. The desire or habit to 'want' and motive or dependence to 'need' to smoke both determine the process of intended behaviour change and act as competing impulses against quitting.

In terms of dependence, nicotine from tobacco generates the impulse to smoke and undermines self-control by interacting with all of the levels of motivation (West, 2009). It leads to cue-driven urges, impairs inhibitory control and gives enjoyment, resulting in '*wanting*' to smoke. It causes nicotine craving, withdrawal symptoms and beliefs about benefits of smoking (e.g. stress relief), all of which can result in a '*need*' to smoke. Wanting to smoke is a major deterrent to 'making quit attempts,' but does not influence success. On the other hand, 'need' to smoke affects quit success, due to the cue-driven impulses to smoke, nicotine craving, adverse mood and beliefs about the benefits of smoking.

1.3.2 Behaviour Change Techniques

Behaviour Change Techniques (BCTs) form an integral part of BIs (Michie et al., 2011a, Albrecht et al., 2013), considered as coordinated sets of activities designed to change targeted behavioural patterns, based on an individual's capability, opportunity and motivation (Michie et al., 2011d). For BIs aimed at smoking cessation these can be defined as the ingredients that alter and redirect causal processes regulating behaviour connected to smoking e.g. feedback, self-monitoring and reinforcement (Michie et al., 2011a).

The Taxonomy of Behaviour Change Techniques for Smoking Cessation

Michie et al. developed a generic list of BCTs based on two published reviews (Abraham and Michie, 2008, Hardeman et al., 2000) and relevant text books (Michie et al., 2008). These BCTs aim at several health interventions including weight management, physical activity and healthy eating and are not specific to smoking cessation.

The goal of behaviour change in smoking cessation is to alter the balance of impulses and motivations by reducing impulses to smoke and increasing motivation and capacity to resist them (Michie et al., 2011c). This involves minimising motivation to smoke, maximising motivation not to smoke, maximising skills and capacity for self-control, and optimising use of smoking cessation medications. This is achieved in different ways, e.g. changing beliefs about what is good or bad, changing drivers of want or need to engage in the smoking behaviour, and changing exposure to stimuli that trigger the impulse to engage in smoking (Michie et al., 2011c).

The taxonomy for classifying BCTs of smoking cessation by their causal determinants was established (Michie et al., 2011c) from two source documents (McEwen et al., 2006, McEwen, 2008). Forty-four BCTs were classified (Michie et al., 2011c, West et al., 2010) according to four broad 'theoretical determinants' that can be mapped to the fundamental concepts of PRIME theory of motivation:

i. *Directly addressing/boosting motivation* e.g. by providing rewards dependent on abstinence,
ii. *Maximising self-regulatory capacity* e.g. facilitating barrier identification and problem solving,
iii. *Promoting adjuvant activities* e.g. advising on stop-smoking medication, and
iv. *Supporting other BCTs or general aspects of the interaction* e.g. building general rapport.

This taxonomy of BCTs for smoking cessation was further used to establish competences required by the BI providers in a manner similar to that achieved with cognitive behavioural therapy (Roth and Pilling, 2008), described below.

Competences identified to effectively deliver Behavioural Interventions

A list of competences for effective delivery of BIs was identified from relevant national and international guidance documents and Stop Smoking Service treatment manuals, centred on individual- and group-based sessions (Michie et al., 2011b). The main objective was to categorise the competences in terms of focus on 'skill versus knowledge' and their function in supporting smoking cessation. The competency for each BCT was derived simply to affirm that the provider should 'be able to' undertake this activity. For example, if the BCT was 'Measuring expired Carbon Monoxide (CO) level', then the competency would be the 'Ability to measure CO level'.

The list was classified in terms of the key behavioural determinants of competences in promoting smoking cessation (Michie et al., 2011b):

i. *BCTs that specifically target the behaviour* e.g. intervention content that directly promotes abstinence
    a) Address motivation: maximise motivation to abstain or minimise motivation to smoke
    b) Maximise self-regulatory capacity and skills: promote mental and physical activities that either reduce exposure to smoking cues or help with resisting motivation to smoke
ii. *Adjuvant activities* e.g. intervention content that promotes activities that indirectly facilitate abstinence
iii. *General aspects of the interaction* e.g. competences necessary for effective delivery of specific BCTs and adjuvant activities
    a) Delivery of the intervention: adapt the intervention according to the client and context.
    b) Information gathering: acquire relevant information.
    c) General communication: give relevant information and verbal and non-verbal behaviour underpinning effective delivery of specific BCTs and adjuvant activities.
    d) Professionalism: general aspects of conduct as a health professional working in the field.

According to the proponents of motivational interviewing (MI) (Hardcastle et al., 2016), another behaviour change approach (Heckman et al., 2010, Lai et al., 2010), the behaviour change taxonomies (Michie et al., 2011b, Michie et al., 2011c) have exclusively focused on describing the intervention content. They believe that the techniques classified in the above taxonomies do not include the interpersonal aspects of the intervention, that is, the manner or 'way' in which intervention content is delivered or expressed to the smoker (Hardcastle et al., 2016). They classify BCTs according to their function as 'content-based' or 'relational' for those related to the quality of interaction (Hardcastle et al., 2016). In Chapter 4, I will present an in-depth study of these ideologies, incorporating the content-based and relational techniques for describing features of a BI, the distinction between 'relational' techniques and 'generic competences', and whether these generic competences can afford specification of the interpersonal aspects of the intervention.

1.4 MEASUREMENT OF FIDELITY FOR BEHAVIOURAL INTERVENTIONS

The term 'fidelity' (also referred to as 'implementation', 'adherence' and 'integrity') describes the actual delivery of an intervention, versus the intended delivery (Moore et al., 2015). It is defined as the degree to which an intervention is implemented as it was designed in the original protocol (Peters et al., 2013). While it may be worthwhile to apply a common framework for evaluating fidelity (e.g. RE-AIM, Realist Evaluation or others (Pawson and Tilley, 1997, Glasgow et al., 1999, Carroll et al., 2007, Weiss, 1997)), it would be difficult to explain the variability in smoking quit rates using these approaches, for three reasons:

1) Linking behaviour change theory with fidelity: the lack of a methodology to capture the active features of a BI including provider expertise, in a measurable way;

2) Measuring fidelity: the lack of a measure to quantify fidelity objectively; and,

3) The extent to which fidelity predicts outcomes: lack of studies linking fidelity scores of BIs with quit rates.

Despite the potential importance of fidelity measurement in evaluations of BIs for smoking cessation, research focusing on developing fidelity measures to identify potential causal mechanisms of impact on effectiveness is rare. There is, however, a range of theories in behavioural science (Lorencatto et al., 2013b, Michie et al., 2011c), as described in section 1.3, that provides a consistent common taxonomy for describing the content of BIs for smoking cessation and the generic competences to deliver these in an effective manner (Michie et al., 2011b). The taxonomy of BCTs has so far been used to assess fidelity by coding and comparing the presence/absence of BCTs in English Stop Smoking services against the practice manuals (Lorencatto et al., 2013a), but this is not structured in a quantifiable way that can be used for rating and scoring for fidelity. Similarly, the competences required to deliver BIs have not been applied to assess the interpersonal aspects of delivery in a structured rating format (Michie et al., 2011b).

Fidelity indices can be used to document deviations from an intended practice and differences among the variations in practice (Bond et al., 1997b). Indices (also known as 'scales' or 'criteria') have been used conventionally for quantifying fidelity to compare an intervention, as implemented, to the empirically tested theory on which it is based (Drake et al., 2001, Mowbray et al., 2003). Some researchers have described common elements of fidelity to an intervention in the context of health behaviour change (Bellg et al., 2004, Borrelli et al., 2005, Borrelli, 2011). Others have provided models for developing measures of fidelity (Mowbray et al., 2003, Carroll et al., 2007). However, approaches combining the elements of fidelity specific to BIs and a measure predictive of quit rates do not exist. Further literature on fidelity and its application for BIs is provided in Chapter 4.

With guidance from the MRC framework for process evaluation, existing fidelity literature and the advent of BI specification taxonomies, I set out to develop a method for capturing active ingredients of a BI for smoking cessation in a format that would allow scoring based on fidelity.

1.5 FIDELITY TO BEHAVIOURAL INTERVENTION AS A METHOD OF INFORMING VARIATION IN OUTCOMES

The second part of my research concerns understanding of the mechanisms of impact of intervention features on the quit rates of the patients. To clarify, understanding the mechanisms of impact means establishing which of the techniques (content-based or interpersonal) that were actively delivered to the patients triggered change in the targeted behaviour.

1.5.1 Fidelity – as an intermediate implementation variable

The characteristics or different aspects of implementation response often studied in process evaluations are referred to as the implementation variables, e.g. acceptability, feasibility, fidelity etc. (Peters et al., 2013). In effectiveness evaluations, these variables are considered as intermediate factors that explain other important outcomes like smoking cessation, which are the prime target of the intervention (Brownson et al., 2012, Proctor et al., 2011). However, not all these variables are of equal importance in the delivery of an intervention (Proctor et al., 2011). In novel interventions, the focus remains mainly on factors relating to acceptability and feasibility, while in existing interventions, the degree to which the intervention is implemented as it was originally designed, or is faithful to the original (measured by the fidelity variable), is considered vital (Peters et al., 2013). For this thesis, the implementation variable of importance is fidelity, as I am interested in exploring variation in outcomes of an existing BI for smoking cessation. The requisite elements of fidelity in relation to BIs will be discussed in detail in Chapter 4.

1.5.2 Approaches to examining fidelity

Interventions in health services should be supported by evidence, with robust evaluation before widespread implementation. However, developing this evidence, where the interventions are complex is by no means straightforward. Evaluation purports to measure 'effectiveness', 'worth' and 'value'. Yet not only are these concepts value-laden, and subject to evaluation design, but also concern the nature of complexity in behavioural interventions and how these shape what works and how it works in reality. The necessary 'active ingredient' of the intervention may not be easily identifiable, and may influence outcomes through interactions with implementation in complex ways. In designing my evaluative research, I needed, therefore, to clarify my underlying assumptions in view of the aim of the proposed work.

I aimed to adopt a pragmatic paradigm, which is not committed to any one system of philosophy or reality. This paradigm places 'the research problem' as central and applies all approaches to

understanding the problem (Morgan, 2007). Data collection and analysis methods are chosen as those most likely to provide insights into the question, with no philosophical loyalty to any particular paradigm. The key point in evaluating complex interventions is whether they will work in real life scenarios, therefore making it important to understand the whole range of effects and how they vary among the recipients or across the sites (Craig et al., 2008).

Different approaches to examining the implementation process of an intervention exist, and the most commonly used term to describe such an exploration is 'process evaluation'. In the following paragraphs, I will briefly introduce the approach for linking fidelity with quit rates in relation to my thesis, which is a similar but distinct concept to process evaluation.

The theoretical models (Carroll et al., 2007, Linnan and Steckler, 2002, Grant et al., 2013), presented in Chapter 4, for process evaluation of complex interventions mostly apply formative or developmental research approaches (Moore et al., 2015). These approaches to process evaluation are undertaken during the design and pre-testing stages of interventions, targeted at improving the intervention theory and development (Rossi et al., 2003). Interpretive evaluation, on the contrary, employs some of the same concepts and methods as process evaluation, but its role is to illuminate the mechanism of action of the delivered intervention and enhance understanding of its impact or worth (Stetler et al., 2006).

Interpretive evaluation is a form of process evaluation that attempts to understand the complexity of an intervention (Neutens and Rubinson, 2001, Stetler et al., 2006). It attempts to understand the impact of an intervention, as opposed to experimental evaluations, which are used to validate simplified effects of the intervention through controlled comparisons (Neutens and Rubinson, 2001). It provides alternative explanations for results, helps to clarify the mechanisms of impact of the "black box" of an intervention, often, including associational relationships with outcomes (Stetler et al., 2006). Such interpretation occurs through the end point triangulation of qualitative and quantitative data collected for intermediary factors (i.e. fidelity) and the effect outcomes e.g. smoking cessation, including associational links between the two to furnish understanding of intervention impact.

MRC recommends validating association of fidelity with effectiveness outcomes of interventions where quantitative data is available (Moore et al., 2015). This could inform whether fidelity varied substantially between implementation sites or providers and if better delivery produced better outcomes. However, issues concerning timing of analysis play a significant role in interpretive approaches to process evaluation and the use of fidelity- as intermediate factor- with outcomes (Moore et al., 2015). Some perspectives advocate analysing fidelity data independent of trial outcomes, to avoid biasing these analyses (Oakley et al., 2006b). Others highlight the value of post-trial analysis of causal pathways and intervention fidelity in allowing emerging issues to be explored (Kinmonth et al., 2008).

1.5.3 Potential methods for examining fidelity

Process evaluation is a rapidly evolving science, which does not have a narrow set of specific research methods but rather it draws on a wide variety of qualitative, quantitative, and mixed-methods approaches (Peters et al., 2013). Mixed-methods approaches are highly recommended for studying fidelity, to maximise the interpretation of findings using different methods to complement each other (Leech and Onwuegbuzie, 2010). While these different approaches provide a basic toolkit for examining fidelity, it is the research question that is paramount in determining the type of research methods to be used (Peters et al., 2013).

Most process evaluation approaches use a combination of qualitative and quantitative methods (Moore et al., 2015). Qualitative methods and participatory action research approaches are suitable for research questions concerned with 'exploring', 'describing', or 'adequacy' of implementation of an intervention (Peters, 2009, Peters et al., 2013, Habicht et al., 1999). Realist review, on the contrary, provides an excellent approach to synthesising theoretical understanding and empirical evidence from documentary review, and is best suited to informing policy interventions (Onwuegbuzie and Collins, 2007). When studying questions concerned with 'plausibility', or 'probability' that the outcome is due to the implemented intervention, then trials are considered as the gold-standard (Curran et al., 2012). Eventually, research questions requiring 'explanation' of the mechanism of action of an intervention on its outcome, can be addressed through convergence of data and analysis using both quantitative and qualitative methods (Peters, 2009, Peters et al., 2013). Depending on the research focus, the study methods can weigh more towards qualitative or quantitative inquiry (Peters et al., 2013).

Fidelity can be assessed using quantitative monitoring and qualitative exploration during intervention implementation, while contextual factors can be explored qualitatively before and after the intervention is implemented (Grant et al., 2013). Grant et al. (2013) further highlight the use of quantitative data collected post implementation for exploring the potential effect of theory-linked causal mechanisms of an intervention on its outcomes. Common quantitative methods used for measuring fidelity to complex interventions include structured observations, self-record questionnaires, and secondary analysis of routine data (Moore et al., 2015). The qualitative methods combined with these often include one-to-one interviews, focus groups and non-participant observation. Subsequently, inferential statistics can be used to test significant associations using data collected from different sources by these methods (Stetler et al., 2006).

Most of the aforementioned methods provide insights into understanding implementation of an intervention and its mechanisms of impact on the outcome, which are defined as the key functions of process evaluation by MRC (Moore et al., 2015). I have adapted these key functions of process evaluation within an interpretive evaluation model (Figure 1.2). I will present the methods for designing a fidelity measure and using it to interpret variation in quit rates of smoking cessation, in detail, in respective chapters of the thesis.

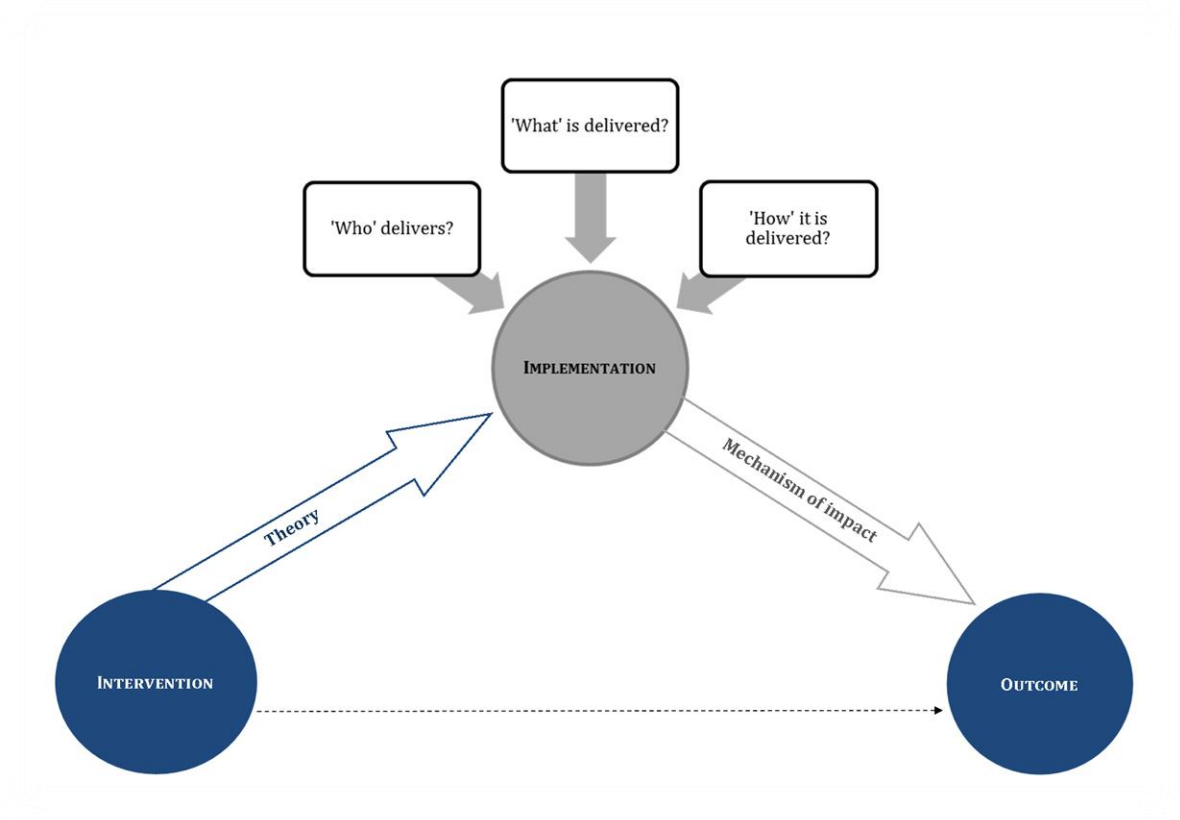


Figure 1.2: Interpretive Evaluation Model

1.6 THESIS OUTLINE AND OVERVIEW OF PLANNED STUDIES

In this thesis, I will first present four reviews of the relevant literature describing:

i. Smoking in TB patients- evidence synthesis,

ii. Smoking cessation in high TB burden countries,

iii. Sources of variation in smoking cessation-ASSIST study, and

iv. Description of intervention fidelity;

Reviewing the literature will help justify my research aims and build the foundation for designing a fidelity measure and evaluating fidelity as an intermediary factor in explaining quit rates. This work will lead into the three linked primary studies planned for the thesis:

Study A- describes the steps in the development of a fidelity measure for BIs in smoking cessation

Study B- describes reliability testing of this fidelity measure to see if it is fit for measuring fidelity to BIs as intended, and,

Study C- evaluates the intermediary role of fidelity in explaining quit rates of a BI.

All three studies use different methods, mainly quantitative, incorporating several linked investigations. Figure 1.3 gives an overview of the studies in the thesis.

A word of explanation is necessary before I present the case of variability in quit rates of a BI for smoking cessation. In this thesis I used data from a large cluster RCT- Action to Stop Smoking In Suspected TB (ASSIST), and linked it with fidelity scores obtained in a prospective observational study (see Figure 1.3). The decision to explore the wide variation in quit rates of a BI came out of findings from the ASSIST study. Therefore, I have presented an overview of the ASSIST study and the reported variability in quit rates first, in order to illustrate the process of arriving at this decision. In addition, given the paucity of research on fidelity measures for BIs, in developing a reliable and valid tool, I relied predominantly on extrapolating from the relatively richer literature on behaviour change science and applying it to conventional scale (or index) construction methods.

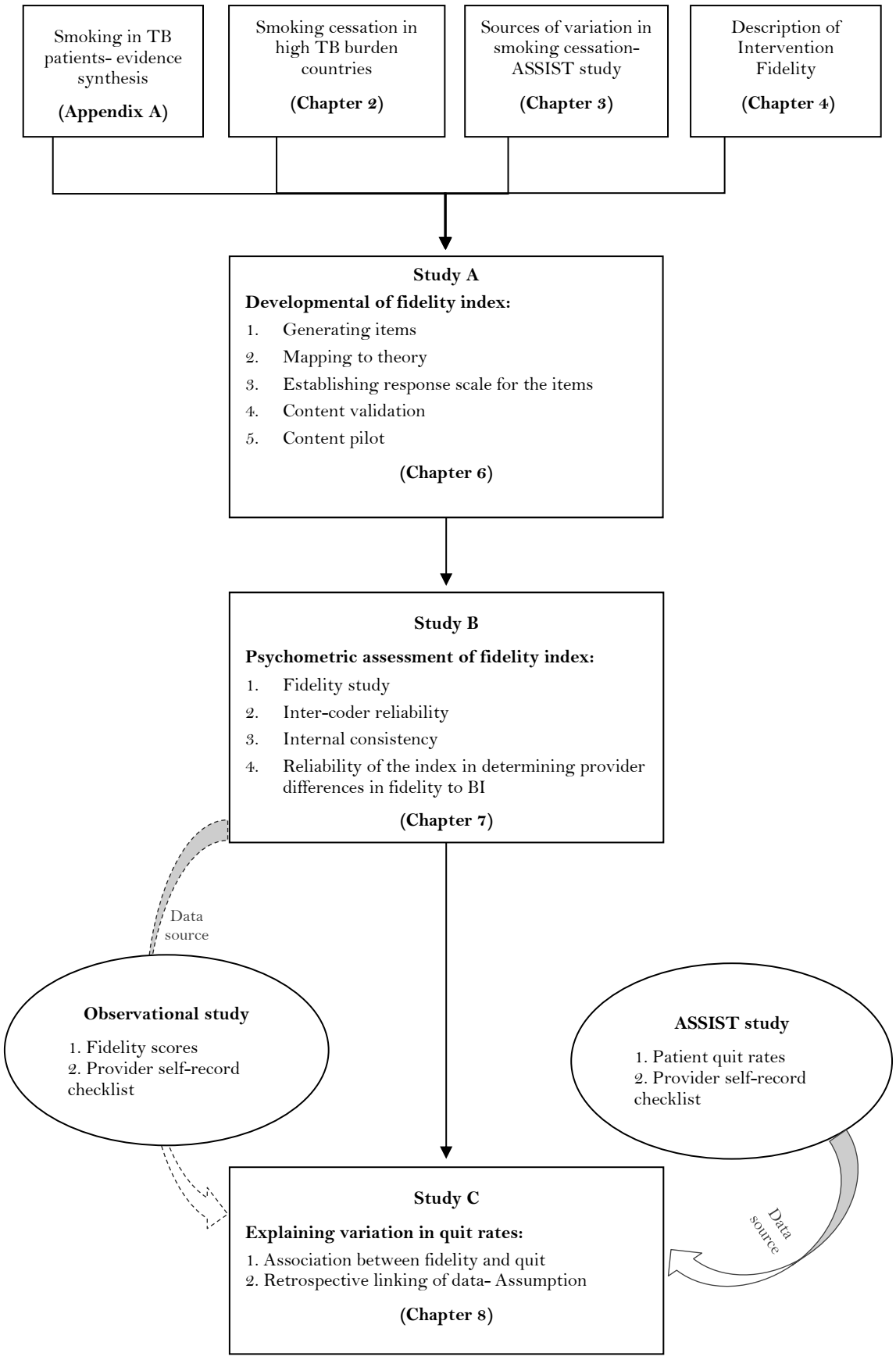


Figure 1.3: Overview of studies in thesis

Chapter 2. Smoking cessation in high TB burden countries

In this Chapter, I first give a brief overview of the association between smoking and TB, highlighting the uncertainty in this evidence. The review on smoking and TB is provided as an appendix. I then give a brief overview of behavioural interventions (BIs) as a cost-effective solution for cessation services in LMICs. I provide the significance of integrating BIs within routine TB care services (commonly called 'National TB programme'). I end the chapter by highlighting challenges of variation in cessation outcomes and the implications on wider implementation of smoking cessation, using a cluster RCT (ASSIST) as the case-study.

Tobacco consumption and TB both contribute sizably to the disease burden in LMICs (Dogar et al., 2013, Siddiqi, 2014). In the past, some systematic reviews have attempted to summarise the effect of smoking on TB outcomes (Bates et al., 2007, Lin et al., 2007, Slama et al., 2007). There remains an uncertainty in the conclusions drawn, in these reviews, due to the quality of observational studies included and the limitations in methods used to appraise them. Hence the usefulness of the evidence presented in these reviews in convincing policy makers and TB managers of the importance of smoking cessation in TB patients remains limited.

This is particularly true when drawing conclusions from evaluations of smoking cessation interventions in patients with conditions like TB. Smoking cessation trials often use quit rates as the primary endpoint, leading to a lack of direct evidence for the effects of smoking cessation on health outcomes like TB outcomes. Hence, it is important to synthesise all available evidence linking smoking with TB to inform policy and public health practice for integrating smoking cessation interventions in TB programmes of LMICs. I, therefore, carried out a review of the literature on the association between smoking and TB outcomes (Appendix A).

Regardless of the TB burden in any country, all those who use tobacco and who have TB may suffer excess mortality and morbidity (Lin et al., 2007, Bates et al., 2007). Understanding the mechanisms by which smoking cessation might alter TB disease in its natural course and identifying the most effective and cost-efficient means of influencing the two epidemics where these converge could hugely benefit research and policy directions in the area (Siddiqi, 2014). Effective smoking cessation interventions exist; however, their effectiveness and cost-implications for LMICs, particularly for TB patients, are under studied (Siddiqi, 2014). In contrast, well-established national TB programmes in these countries offer a unique opportunity for optimising smoking cessation within TB care and evaluating the effects of such integration.

One such smoking cessation BI (Action to Stop Smoking In Suspected TB [ASSIST] - a cluster RCT) was designed for a high TB burden LMIC and was found to be highly effective (Siddiqi et al., 2013). However, there was a wide variation in patient smoking quit rate between TB clinics (clusters) that could not be explained.

2.1 SMOKING AND TUBERCULOSIS

Smoking and TB both contribute significantly to the global burden of disease, not only individually but also by adversely influencing each other. Historically, both smoking and TB have been prevalent for centuries, but it was not until 1881 that tobacco smoking started to become widespread, after the invention of cigarette rolling machine. Coincidently, *Mycobacterium tuberculosis* was first isolated by Robert Koch in 1882 as the causative agent for TB. The potential association between the two (tobacco smoking and TB) was first hypothesised in 1918 (Webb, 1918).

2.1.1 Tobacco smoking

Tobacco smoking alone causes more deaths than any other preventable risk factor in the world. It kills one in two users, leading to approximately 5.7 million deaths in 2010. When second-hand smoke was included, the global burden of disease increased to 6.3 million deaths in that year (Institute for Health Metrics and Evaluation, 2010, Lim et al., 2013). It is a risk factor for six out of the eight leading causes of death worldwide, i.e. ischaemic heart disease, cerebrovascular disease, lower respiratory infections, chronic obstructive pulmonary disease, TB, and cancers of trachea, bronchus and lung (Mathers and Loncar, 2006).

Globally, 1.1 billion people in the world currently smoke tobacco, and 70% of these reside in the LMICs (Ng et al., 2014). Between 1980 and 2012, a large reduction (41% to 31% in men and 11% to 6% in women) in the prevalence of daily smoking was estimated, globally. However, the population growth contributed to an increase in the number of smokers worldwide from 4.96 trillion to 6.25 trillion during this time period (Ng et al., 2014). Moreover, as smoking prevalence declined in the high-income countries (HICs) overtime, an increase has been observed in the LMICs (Ng et al., 2014).

2.1.2 Tuberculosis

TB is a widespread, and in many cases fatal, infectious disease typically affecting the lungs, but can also attack other parts of the body. Every year 8.8 million new cases of TB are diagnosed worldwide, and 81% of these are from LMICs (World Health Organisation, 2014b). Approximately 3.9 million of these newly diagnosed cases are confirmed pulmonary TB on sputum smear microscopy, the most infectious and transmissible form of the disease (World Health Organisation, 2014b, Dye, 2006). Globally, TB is the leading cause of death from a curable infectious disease, causing 1.7 million deaths per year (Dye, 2006). The resurgence of TB in Eastern Europe and the advent of multi-drug resistant TB are some of the emerging challenges for the global TB control (Dye, 2006). Tuberculosis remains prominent on the international agenda to combat preventable diseases, as it kills people in economically productive age-groups.

2.1.3 Impact of smoking on TB burden

An estimated 15% of the pulmonary TB cases in the world each year could be attributed to smoking alone (Pai et al., 2007). This has been estimated from prevalence of exposure to tobacco smoke (30%) and the relative risk of developing TB (1.5), giving a global Population Attributable Risk of 15%. If the current rate of smoking persists, it is projected to cause 18 million additional cases of TB and 40 million excess TB related-deaths between 2010 and 2050, globally (Basu et al., 2011). Projections from these mathematical models of current smoking trends suggest increases in TB incidence ranging from 5%, in the western Pacific region, to 42%, in the Eastern Mediterranean region by 2050. Similarly the increases in TB related deaths are likely to range from 64% in Europe to 135% in Eastern Mediterranean region by mid-century (Basu et al., 2011).

2.1.4 The biological link between smoking and TB

Tobacco smoke influences the systemic and lung defences of the human body. Some biologically plausible explanations of how smoking affects these defences are given below.

Mechanical defences

Exposure to tobacco smoke impairs the function of cilia (tiny hairs) lining the inside surface of airways. This alters the normal clearance of secretions, leading to mucus hyper-secretion. Impaired bacterial clearance leads to increased peri-bronchial inflammation and permeability caused by epithelial damage to the airways (Di Stefano et al., 2012, van Zyl-Smit et al., 2010).

Peripheral immune defences

Tobacco smoke alters the function of innate and adaptive immune cells that collectively defend the body from invading bacteria (Di Stefano et al., 2012, Kumari and Meena, 2014). An important function of CD4+T cells is that they produce interferon gamma (INF-gamma) which signals phagocytic cells (like macrophages) and activates them to engulf the bacteria (Di Stefano et al., 2012). Exposure to tobacco smoke decreases the CD4+ T cell counts (van Zyl-Smit et al., 2010, Stead and Lancaster, 2012), predisposing individuals to respiratory infections. Nicotine and acrolein are constituents of tobacco smoke that increase the number of viable Mycobacteria inside host cells (Kumari and Meena, 2014). Nicotine also impairs antigen (bacteria related) mediated signalling in lymphocytes and T cells, leading to reduced responsiveness and antibody formation (Di Stefano et al., 2012).

Local alveolar immune defences

Tar is a by-product of tobacco smoke that inhibits the production of important markers of inflammation (IL-1b, IL-2), INF-gamma and Tumour Necrosis Factor alpha, which are

responsible for activating the body's defence mechanisms (Di Stefano et al., 2012, Kumari and Meena, 2014). Smoke from tobacco impairs the function of lung alveolar macrophages, which form an integral early defence mechanism to prevent the spread of bacteria by walling them off (Di Stefano et al., 2012). In addition, nicotine (Di Stefano et al., 2012) and acrolein (Kumari and Meena, 2014) inhibit the production of pro-inflammatory cytokines, leading to increased replication of bacteria.

Macrophages in lungs of smokers contain elevated amounts of iron (5-fold to 7-fold) in comparison to non-smokers (Boelaert et al., 2003). Macrophage iron overload impairs defence against bacteria through reduced production of inflammatory markers (Boelaert et al., 2003). This iron overload could be related to the high iron content of tobacco, leading to inhalation of 1·12 g of iron per pack of cigarettes smoked (Boelaert et al., 2003, Mateos et al., 1998).

2.1.5 The behavioural link between smoking and TB

In addition to the biological mechanisms that support the causal chain of adverse effects of smoking on TB, smoking and TB also interact and enhance the behavioural risk for each other.

The prevalence of smoking is found to be higher (OR: 1.9, 95%CI: 1.5 to 2.5) in TB patients than the general population (Wang and Shen, 2009). A high number of TB patients quit smoking when diagnosed with TB, however, 18% (OR: 3.5, 95% CI: 1.3 to 9.5) of these go back to smoking within the next 15 months.

TB patients' smoking pattern and smoking cessation behaviour can alter their clinical disease outcomes (Wang and Shen, 2009). Poorer compliance to TB treatment and thus default among TB patients has been associated (OR 1.8; 95%CI 1.0 to 3.3) with smoking (Lavigne et al., 2006). Poorer compliance to TB treatment in turn prolongs infectiousness and increases the chances of drug resistance, relapse and death (Zignol et al., 2006, Schneider and Novotny, 2007).

2.2 EFFECTIVE SMOKING CESSATION INTERVENTIONS

There is a substantial body of evidence on the efficacy of a range of pharmacological interventions and BIs offered by healthcare professionals for smoking cessation (Lancaster and Stead, 2005, Stead and Lancaster, 2012). However, the evidence of effectiveness and cost-efficiency for offering these interventions to TB patients is limited (Siddiqi, 2014, Piné-Abata et al., 2013). In high TB burden countries, well-developed TB programmes harness the primary health care (PHC) systems, providing TB diagnosis and treatment care nationwide. In countries with concomitantly high tobacco use, such systems offer a unique opportunity to integrate smoking cessation into routine TB diagnosis and treatment care.

2.2.1 Offering help to quit smoking- MPOWER

Globally, there are six highly promoted policies designed to prevent young people from initiating smoking, helping current smokers quit, protecting non-smokers from exposure to second-hand smoke, warning people about the harms of smoking, raising taxes on tobacco and enforcing bans on tobacco promotion (World Health Organization, 2008). One of these policies is to offer help to quit smoking.

The majority of the smokers who wish to stop, find it extremely difficult to give up on their own without proper professional support, leading them into several but futile quit attempts (Fiore, 2000). Those who seek professional support to stop smoking are four times more likely to give up their habit successfully than those attempting to stop on their own (NICE, 2008).

2.2.2 Treatments for cessation

Behavioural support (based on BIs) for smoking cessation refers to any form of advice, discussion, encouragement and activity designed to increase the probability of success in quit attempts (Lancaster and Stead, 2005). Behavioural support for smoking cessation is found to be effective either alone, (relative risk (RR) of stopping being 1.39 (95%CI: 1.24 to 1.57) (Lancaster and Stead, 2005)) or in combination with pharmacotherapies including Nicotine Replacement Therapies-NRTs, Bupropion and Varenicline, (RR: 1.82 (95%CI: 1.66 to 2.00), respectively) (Dogar and Siddiqi, 2013, Stead and Lancaster, 2012).

2.3 CONTEXTUALISING SMOKING CESSATION IN LMICS

Epidemiologic transition in the past has been described by the changing patterns of health and disease over time as populations tend to age (Omran, 1971). However, this concept of demographical transition entailing population growth and longevity of life has not consistently explained the epidemiological transition from communicable diseases to Non-Communicable Diseases in developing countries. Changing life-style, food market globalisation, increasing urbanisation and economic development are shown to predict the current epidemiological situation in developing countries (Stuckler, 2008). The sharp demarcation between communicable diseases and Non-Communicable Diseases seen in the past now appears to be merging, as both types of disease affect the same individuals and populations (Bygbjerg, 2012). Many LMICs are worst impacted by this convergence, attempting to contain the communicable diseases, while at the same time being faced with the rise in Non-Communicable Diseases (Stuckler et al., 2010).

2.3.1 Conceptualising the 'Syndemic'

'Syndemic' is a relatively new concept used to describe the synergistic interaction of two or more conditions at multiple levels of causation and linkage contributing to a greater burden of disease in a population. It includes not only the impact due to interaction between the two conditions but

also the forces that cluster these conditions in persons, places and/or times (Littleton and Park, 2009). Conceptualising tobacco and TB as a syndemic allows understanding of the synergies at multiple levels between both epidemics, whereby their mutual presence has an amplified negative impact on population health (Littleton and Park, 2009). It acknowledges the burden of these two mutually enhancing epidemics as borne by those countries that are the most marginal, globally. Both tobacco and TB tend to interact with other aspects of socio-economic and environmental context in these populations, such as HIV-infection, Non-Communicable Diseases, malnutrition, poor housing and indoor air pollution (Lonnroth et al., 2009). Thus, the concept allows recognising that over time this burden might become still more concentrated, due to the additive effects of the intergenerational deprivation if not intervened effectively (Littleton and Park, 2009, Lonnroth et al., 2009). A number of possible pathways linking the two conditions (tobacco and TB) and the co-existence of multiple inter-linking factors/diseases point towards the syndemic context of the situation in LMICs (Littleton and Park, 2009).

2.3.2 Smoking trends in LMICs

The health related effects of smoking in any population pertain to the intensity of its consumption and its prevalence. In LMICs, the prevalence of smoking is rising at a fast pace compared to the HICs, where it is either stable or decreasing over time (Esson and Leeder, 2004). Between years 1970 and 2000, the smoking prevalence has more than doubled in the middle-income countries and is expected to rise further by 60% between 2000 and 2025 (Esson and Leeder, 2004). Similarly, prevalence in low-income countries has shown a steep increase since 1990, which is expected to increase by more than 100% by the year 2025, if this trend of smoking continues (Esson and Leeder, 2004).

This increasingly high prevalence of smoking in LMICs may impact TB disease incidence and treatment outcomes substantially (Patra et al., 2015, Bates et al., 2007, Lin et al., 2007). Tobacco smoking is associated with a higher risk of developing TB and poorer TB treatment outcomes, such as death, default (non-adherence to treatment), and treatment failure or relapse (Lin et al., 2007, Bates et al., 2007, Slama et al., 2007). Second-hand smoking is also associated with an increased risk of acquiring TB infection and developing TB disease (Patra et al., 2015, Dogar et al., 2015). These moderate increases in relative risks of TB due to tobacco smoking could translate into marked increases in the absolute risks at the population level, due to the high smoking prevalence.

Nevertheless, the intensity of smoking consumption in LMICs remains low (i.e. < 10 cigarettes per smoker per day) (Ng et al., 2014). This low consumption of tobacco smoking in most LMICs is remarkable, given its high prevalence contrary to the pattern of consumption and prevalence in HICs (Ng et al., 2014). Although, the less harm attributed to the lower intensity of consumption

is offset by the high rates of smoking prevalence in these LMICs, the low tobacco consumption could be used to advantage for promoting smoking cessation. A combination of cost-effective smoking cessation interventions and approaches to promote health behaviour change e.g. teachable moments targeted at this potentially susceptible population in TB clinics can have a multiplicative effect on smoking cessation impact.

2.3.3 Smoking cessation in TB patients

An established evidence-base exists for effective smoking cessation interventions in HICs (Lancaster and Stead, 2005, Stead and Lancaster, 2012), however, very little evidence is found for LMICs. Incorporating cost-effective smoking cessation interventions within TB programmes might have the potential to improve outcomes in confirmed TB cases. Additionally, it could be a means to reach and benefit the much larger group of those with suspected disease presenting to TB clinics in LMICs. Three distinct benefits of offering smoking cessation in TB patients are:

1. Greater health benefits than the general population
2. Opportunity of 'teachable moments'
3. Potential to integrate smoking cessation within TB programmes

Greater health benefits

Most of the immunological abnormalities in TB patients induced by smoking tobacco are shown to reverse within six weeks after stopping smoking (Arcavi and Benowitz, 2004). Findings from the 50 year cohort of British doctors showed that for men born in 1900–1930, stopping smoking at the age of 50 reduced their risk of premature smoking-related death by half, compared to persistent smoking; while stopping at the age of 30 avoided almost all associated risk (Peto et al., 1999, Doll et al., 2004). There are only a handful of empirical studies evaluating smoking cessation interventions aimed at smokers attending TB clinics in LMICs. All of these studies have a more pragmatic focus in terms of evaluating cessation interventions in public health programme settings. The evidence from these studies is summarised in the following paragraphs.

A quasi-experimental (non-randomised) multi-centre study conducted in Malaysia found that the impact of integrating smoking cessation services in conventional Directly Observed Treatment Short-course (DOTS) TB programmes may confer advantages for short-term TB treatment outcomes in these patients and possibly for their long-term lung health (Awaisu et al., 2011). High smoking cessation rates (77.5% vs. 8.7%) and improved TB treatment response (no cavitary lesion at treatment completion: 68% vs. 35%), treatment adherence and significant reductions in treatment failure rate were also observed among the TB patients assigned to the intervention condition (behavioural support and Nicotine Replacement Therapy-NRT) compared to the standard care.

Another feasibility study was conducted in PHCs in Brazil, to determine whether TB DOTS paramedics could be trained to deliver smoking cessation support effectively in a TB clinic (Sereno et al., 2012). This study found a dose-response effect between the duration of individual face-to-face patient-provider interaction session and successful cessation outcomes.

A cluster RCT (see 'Overview of ASSIST study'- Appendix B) in Pakistan found that an inexpensive behavioural intervention to stop smoking was effective in supporting patients who attend PHCs for diagnosis and treatment of TB, to achieve sustained smoking cessation (Siddiqi et al., 2013).

Teachable moments

'Teachable moments' is a strategy advocated for promoting health behaviour change in a variety of settings (Lawson and Flocke, 2009). It is often conceptualised as events or a set of circumstances which leads patients to alter their health behaviour positively; it can also be created by the provider in a face-to-face interaction with the patient as an opportunity to promote healthy behaviour (Lawson and Flocke, 2009). Utilising the 'teachable moments' to promote smoking cessation in TB patients can markedly benefit them from cessation advice and success in quitting compared to general smokers (McBride et al., 2003). General smokers do not face the same level of anxiety or fear of death and the possibility of infecting others, as those with TB. This often-missed opportunity can be utilised by health professionals to tailor their smoking cessation advice for TB patients and reinforce the benefits of stopping tobacco smoking in those with TB.

Integrating smoking cessation in TB programmes

Very few LMICs have the basic infrastructure and systems in place to offer cessation support to tobacco users (Siddiqi, 2014). In a recent survey of smoking cessation services in 121 countries, only a quarter of the low income countries were found to have an official national smoking cessation strategy; one-fifth had an official responsible for it; one-tenth had treatment guidelines; none had an identified budget for cessation; 5% had quit-lines; and few had any specialist behavioural support (Piné-Abata et al., 2013).

While setting up specialist cessation services could be costly and require an infrastructure that does not currently exist, integrating smoking cessation within existing public health programmes like a TB programme is feasible and could be highly cost-effective in these resource-limited countries. Moreover, the TB programmes often run laterally with the health care systems in these countries for successful identification and treatment of TB and can be utilised for targeted smoking cessation. These National programmes deliver TB clinics as part of the health infrastructure of the country at all health care levels (tertiary, secondary and primary). The programme appoints a '*TB DOTS facilitator'* at each government health centre. These paramedics

are responsible for: (a) assisting in identifying, recording and reporting new TB patients; (b) educating them about TB; and (c) ensuring direct observation during the intensive phase of their TB treatment. Health care providers, being closely involved in the vast majority of clinical contacts with their patients, have an opportunity to influence the patients' smoking behaviour and reduce related harm (Whyte and Kearney, 2003, Rice and Stead, 2008).

In the ASSIST study, it was found that an inexpensive behavioural support intervention embedded in TB care was very effective in supporting patients attending PHCs to achieve sustained smoking cessation (Siddiqi et al., 2013). Such integration is not only likely to be cost-efficient, because it brings together two different services utilising the same infrastructure, but also implies the utilisation of the 'teachable moment' concept in addressing smoking cessation within TB programmes (Siddiqi, 2014).

2.4 VARIATION IN SMOKING CESSATION EFFECT

ASSIST was a pragmatic cluster RCT designed for testing behavioural support (with and without bupropion) in a high TB burden LMIC (see 'Overview of ASSIST study'- Appendix B). The overall effect estimate for the intervention was significantly large; however, there was a wide heterogeneity in patient quit rates when the effect sizes were disaggregated by TB clinics (clusters).

The intra-cluster correlation co-efficient (ICC) (described in Chapter 3) for the effect estimate was 0.28, denoting the presence of a large clustering effect and indicating a strong influence of clinics in determining the success of the intervention. Despite overall highly effective cessation rates, they were seen to vary substantially across individual clinics, especially in the intervention groups (Table 2.1). Therefore, some clinics delivering the intervention had an average quit rate as low as 7%, while others had cessation rates as high as 72% at the intervention sites.

Table 2.1: Proportion abstinent by TB clinics (intervention groups)- ASSIST study (2010- 2011) Pakistan

Clinic ID	Proportion abstinent† n/N (%)
26	41/57 (71.9)
28	37/55 (67.3)
25	34/55 (61.8)
01	35/61 (57.4)
11	34/60 (56.7)
03	33/60 (55.0)
21	33/60 (55.0)
19	34/62 (54.8)
23	31/57 (54.4)
05	31/60 (51.7)
09	27/59 (45.8)
04	26/64 (40.6)
22	22/55 (40.0)
27	19/54 (35.2)
02	19/60 (31.7)
10	17/60 (28.3)
24	16/61 (26.3)
07	12/60 (20.0)
08	11/59 (18.6)
06	7/60 (11.7)
12	6/60 (10.0)
20	4/60 (6.7)

† Number abstaining among the total participants in the TB clinic.

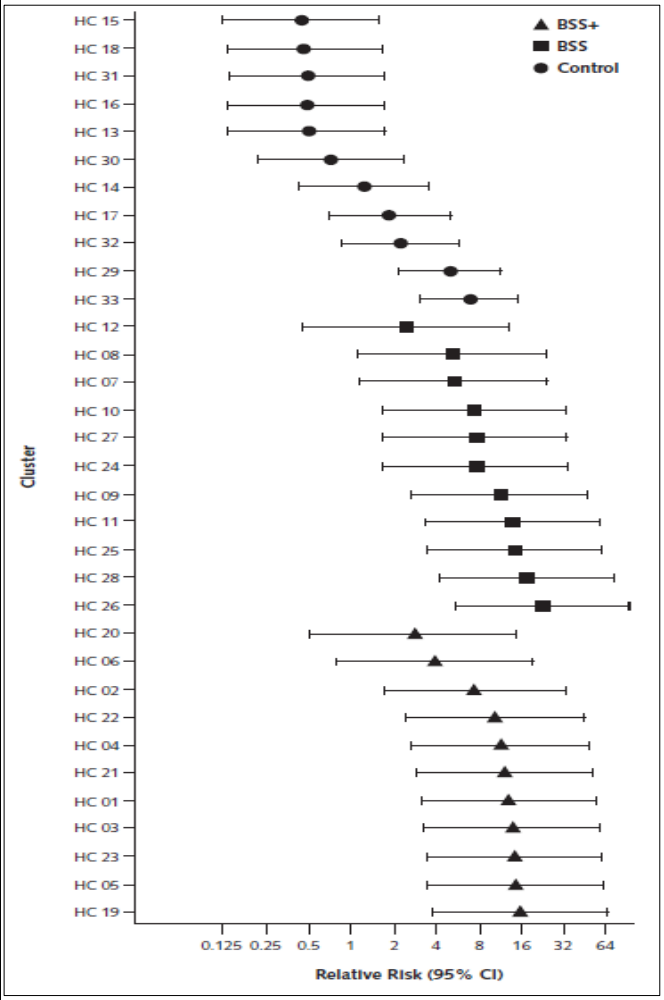


Figure 2.1: Variation in quit rates (RR of abstinence) across all TB clinics- ASSIST study (2010- 2011) Pakistan

Although there was some variation between the clinics in the control group, in general, these clinics had lower-than-average cessation rates (refer to Relative Risks of abstinence by clinics, Figure 2.1; Siddiqi et al., 2013).

The ICC value of 0.28 means that 28% of the total variance in cessation rates was due to the differences between TB clinics rather than the true effect of the intervention itself. Risk estimates for smoking cessation were characterised by large variation across the clinics in the three trial groups. In general, clinics assigned to the control had lower-than-average cessation rates. However, some control clinics (HC 29 and HC 33) reported cessation rates similar to those of the best performing clinics in intervention groups.

Achieving a significant effect of an intervention in the ASSIST study means that the intervention works. However, it is also very important to understand the context and conditions in which that intervention worked best and to explore the conditions that are necessary for delivering intervention in the best possible way to achieve the optimal effect (Petticrew et al., 2013).

2.5 SUMMARY

Inexpensive BIs are effective in supporting patients to quit smoking during their visit to PHCs for diagnosis and treatment of TB, as indicated in the ASSIST study. Incorporating such cost-effective smoking cessation interventions within TB programmes has the potential to improve outcomes in patients attending TB clinics in the affected countries. However, there is a knowledge gap about the factors that are likely to cause variation in outcomes of complex BIs and research to address this gap is clearly needed. More importantly, variation in an intervention's effect might have programmatic and policy implications for a wider implementation of smoking cessation within an existing health service such as a TB programme. Research into the sources of variation in the outcomes of smoking cessation will help focus on factors to improve intervention delivery and selection of the active ingredients for wider scale-up, which could potentially lead to conservation of resources.

Chapter 3. Sources of Variation in smoking cessation- ASSIST study

In this Chapter, I first describe the phenomenon of clustering of individual level outcomes in evaluations using cluster RCT design. I explore the potential sources of variation in smoking cessation using relevant literature and data from the ASSIST study. I provide the significance of factors intrinsic to the intervention in shaping outcomes of complex BIs. In the end, I summarise the sources likely to contribute to the differential effect of BI on smoking quit in the ASSIST study that need further investigation.

An insight into the between clusters differences in outcomes could further aid the applicability and validity of findings from RCTs of complex interventions (Bellg et al., 2004). However, the sources of variation in outcomes of complex BIs for smoking cessation have rarely been explored. Interpretation of findings from well-conducted RCTs of complex interventions can be misinforming, given that the reasons for wide variations in outcomes are not explored and reported. The findings from these RCTs are not only hard to interpret, but also challenging to generalise, unless a better understanding of the active ingredients of the intervention is gained (Craig et al., 2008). The term 'cluster', for the purposes of description in this thesis refers to a TB clinic, unless otherwise specified.

Therefore, evaluating the effectiveness of a complex intervention entails more than just assessing whether or not it works. It involves the characterisation and examination of the potential sources of variation (Pigott and Shepperd, 2013). Exploration of these sources requires taking into account the various dimensions of complexity of BIs, in order to understand the causal mechanisms through which the intervention exerts its effect. Disaggregating sources of variation can also help understand the differences in outcomes between multiple sites of implementation and help optimise intervention delivery.

Achieving a significant effect of an intervention is the first step in establishing that it works, prior to further exploration about how it actually works (Petticrew et al., 2013). Complex interventions (as outlined in Chapter 1) are made up of several interacting features and the range of their effects may vary by the number and difficulty of behaviours required by those delivering or receiving the intervention, the number of groups or levels targeted by the intervention, the number and variability of outcomes and the degree of adaptation of the intervention allowed (Craig et al., 2008). Such interventions are consequently difficult to describe, standardise and administer consistently to all patients (Boutron et al., 2008).

Behaviour change interventions often contain several features, and do not merely comprise a sum of these features but much more complex interactions between them (Hawe et al., 2004). There can be many potential sources of complexity in the relationship between a BI and its outcomes. Two key factors that might contribute to this variation are the interacting nature of its features and the potential variation in the way these are delivered (Craig et al., 2008).

3.1 CLUSTERING OF INDIVIDUAL OUTCOMES BY UNIT OF INTERVENTION ALLOCATION

The cluster RCT design is increasingly being used for evaluating effectiveness of complex interventions where health-care provider, health centre or community is the unit of allocation, rather than the individual participants (Eccles et al., 2003). Variation in outcomes is a different and more difficult type of problem in cluster trials compared with simple RCTs. When a cluster is the allocation unit for the intervention, standardising the way that the intervention is delivered across all clusters can be challenging; potentially leading to clustering of individual outcomes (Hawe et al., 2004). Conversely, cluster design offers the utility for further exploration of such variations in individual outcomes between clusters, by allowing identification and examination of factors influencing at cluster-level (Grant et al., 2013).

3.1.1 Understanding clustering effect

The Intra-Cluster correlation Coefficient (ICC) is calculated from the variance in the outcome. This parameter can be understood in the analysis of variance (ANOVA) framework as the proportion of the total variation in outcome that can be attributed to the difference between the clusters, i.e. $\mathrm{ICC} = {\sigma_B^2}/{(\sigma_B^2 + \sigma_e^2)}$ where σ_B^2 represents variance between clusters, and σ_e^2 the variance within clusters. Strictly speaking it measures the correlation between units (e.g. individuals) within a higher level of unit (e.g. TB clinics), ranging between 0 to1. The lower (approaching 0) the value of ICC, the more independently the individuals within clusters act, and higher (approaching 1) values of ICC demonstrate that the individuals within clusters act as a pact i.e. they act similarly to each other, thus introducing a higher clustering effect in the study.

Furthermore, the higher the ICC is (towards 1) the more variability between the TB clinics it represents and thus it contains more information and there is less chance of this variation being due to measurement error.

3.1.2 Factors affecting the magnitude of clustering

The variability or the magnitude of an ICC might increase, due to a variety of factors including; the type of the primary study endpoint (e.g. intermediate factors vs. outcome), the effect of setting (community vs. clinical), the prevalence of the endpoint (e.g. smoking quit rate, as variability is low at both extremes of prevalence), the size of the clusters, and the characteristics of individuals and clusters (Adams et al., 2004, Campbell et al., 2005).

Effectiveness studies of complex interventions usually focus on measures of behaviour change dependent on process variables, such as adherence to protocol or compliance with recommended best practice amongst the intervention providers, rather than on the ultimate goal of behaviour change i.e. change in patient outcomes such as smoking quit (Mason et al., 1999). However, the

ASSIST study outcomes were focused on the quit rates of the individual participants who received BI from the TB care providers. Observations depending on measures of behaviour change or process outcomes tend to be more correlated and thus give ICCs of higher magnitude (Marion et al., 2001).

3.2 EXPLORING SOURCES OF VARIATION- ASSIST AS A CASE STUDY

A useful framework for characterising sources of variation in complex interventions consists of three components: firstly, substantive features of the complex intervention and how it is delivered in a given study; secondly, the procedures and methods used to conduct its evaluation such as the cluster RCT design (as discussed above); and thirdly, factors extrinsic to the intervention, such as research context (Pigott and Shepperd, 2013). I used this framework for conceptualising and describing the sources of variation in quit rates in the ASSIST study (Figure 3.1), in subsequent sections.

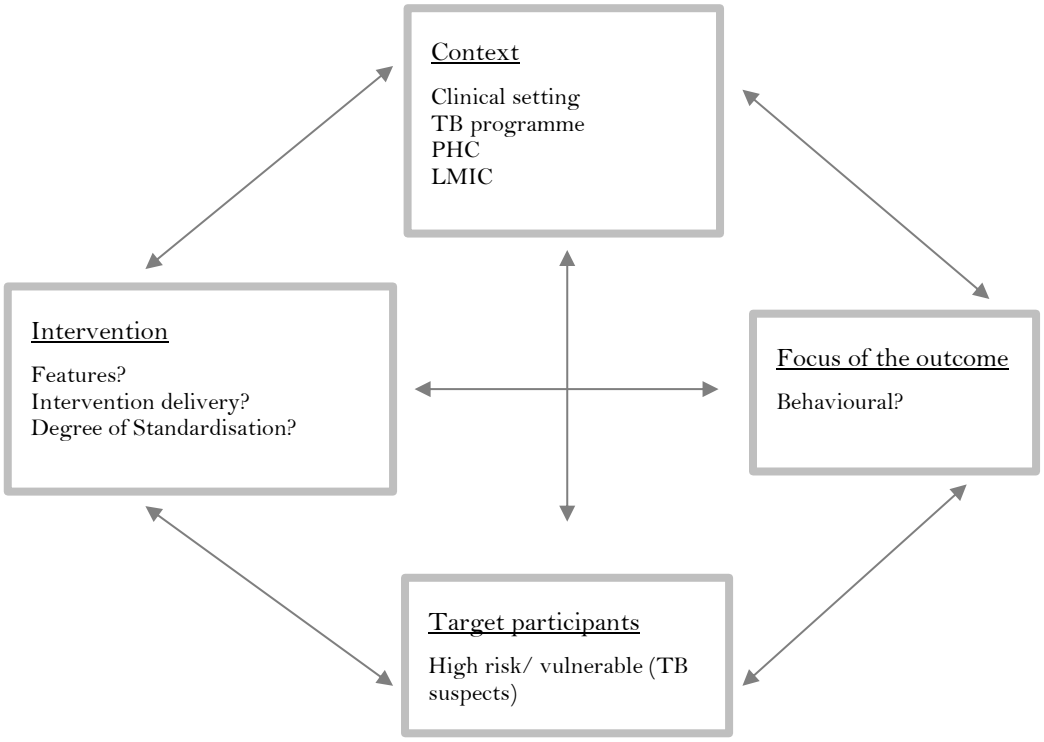


Figure 3.1: Factors that can lead to variation in quit rates of ASSIST study

3.2.1 Target participants

Clustering of individuals' smoking quit rates by TB clinics as observed in the ASSIST study (ICC of 0.28) is considerably higher than the usual ICC of 0.036 seen in other studies conducted in similar settings (Parker et al., 2005). Differences in individuals' characteristics (i.e. their demographics and personal attributes, socio-economic status, education level, motivation level

etc.) may contribute to clustering of similar outcomes, thus it is important to take account of these factors while exploring sources of variation in outcomes of complex interventions (Craig et al., 2008).

More commonly, individuals clustered by sites or providers tend to interact frequently and as a result may respond in a similar manner (Donner and Klar, 2000). This is more likely to be an issue in group therapies. The patients recruited in the ASSIST study were not related to each other and they received individual face-to-face BI sessions, making it less likely for them to interact with each other during or after the sessions. On average, one patient was delivered behavioural support on any given day.

To explore the variation in quit rates and control for its effect, the target participant's characteristics (age, gender, duration of smoking, income and quantity smoked) in the ASSIST data were accounted for in a statistical model. However, adjusting for the patient characteristics only increased the variation further to 32% (ICC= 0.32) from the crude association, showing 28% clustering effect (refer to Table 3.1; Siddiqi et al., 2013).

Table 3.1: Quit rates at 6 months adjusted for individual level characteristics- ASSIST study (2010- 2011) Pakistan

Trial Group	**Relative Risk [Ъ] (95% CI)**	***p*-value**	**ICC**
Behavioural support + bupropion (BSS+)	9.3 (4.0-21.6)	<0·0001	0.32
Behavioural support alone (BSS)	8.5 (3.7-19.6)	<0·0001	

[Ъ] Adjusted for age, gender, income, smoking duration & quantity smoked per day

Hence, the variation in quit rates by TB clinics in the ASSIST study could not be explained or reduced by controlling for the individual characteristics of the TB suspects.

3.2.2 Context

Context in terms of complex interventions means any factor extrinsic to the intervention itself that tends to hinder or bolster its effects (Moore et al., 2015). An intervention might have different results if delivered in different settings, as causal pathways underlying the issues targeted by the intervention might differ from one context to another (Bonell et al., 2006).

A provider's ability to change is predetermined by factors like skills, training, resources and attitudes. Similarly, pre-existing factors determine to an extent the response of a target group to a certain intervention. It is this mutual ability to transform and adapt between an intervention and its immediate context that partly determines the effect estimates (Jansen et al., 2010). However, interventions that appear outwardly simple may still be considered highly complex when they interact with their context. According to Donner and Klar (2000), the variation in outcomes seen in cluster RCTs could be due to individuals tending to select the clusters to which

they belong more often (e.g. preference in being seen by a particular provider) but more importantly it could also be due to factors at the cluster level that affect all enrolled individuals similarly, e.g. provider practice (Donner and Klar, 2000).

Context at the micro-level for an intervention, including training and technical support to the providers, is likely to have a bearing on the mechanism of impact of its key features on the cessation outcomes. At a macro-level, an implementer's engagement in the preplanning and planning stages of the intervention development, integration with the goals and missions of the implementer's organisation, and the organisational environment all can impact outcomes (Dane and Schneider, 1998)

A final aspect of implementation regards the motivations for adopting an intervention that often dictate its success or failure in being delivered as intended (Dane and Schneider, 1998). Interventions initiated within organisations that are deemed relevant after empirically documenting needs, and are in line with the stated goals and missions of the implementing organisation, are more likely to be delivered, with better quality, than those initiated by forces external to it (Dane and Schneider, 1998, Durlak and DuPre, 2008). Moreover, positive organisational climate, stability, shared decision making and staff support all influence the mechanism of impact of an intervention (Durlak and DuPre, 2008). Some of these macro- or organisational level contextual factors important in the integration of complex behavioural interventions in LMICs settings were explored using ASSIST as a case-study (Dogar et al., 2016). These factors will not be discussed in detail here as they deviate from the focus of the current research.

Potential cluster level factors like its size (big/ small) and locality (urban vs. rural) might also contribute to the clustering of outcomes. Studies with larger cluster sizes (the number of individuals within each cluster) tend to have lower clustering effects than those with smaller cluster sizes, as one would expect greater similarity of response in smaller clusters (e.g. households), compared to larger clusters (e.g. diverse communities) (Siddiqui et al., 1996). The size of clusters in the ASSIST study ranged between 55 and 74 patients per clinic, which is much less likely to produce greater similarity of responses between individuals.

PHC settings act largely autonomously compared to secondary or more specialised care facilities that have better standardised and consistent management practices across hospitals, thus leading to higher clustering within these settings (Campbell et al., 2005). Seven out of the 33 TB clinics in the ASSIST study were secondary care clinics, while the rest were primary care clinics, classified as urban/rural for reference. However, when explored in a cluster-level statistical model for urban vs. rural this did not significantly predict the quit rates (refer to Table 3.2; (Siddiqi et al., 2013)). To explore the cluster-level variation further, all known study characteristics (mean age, proportion of men, mean duration of smoking, median quantity smoked and median household income) summarised by clusters, were statistically modelled to adjust for

the unexplained variation and none were found to be significantly associated with the quit rates (refer to Table 3.2; (Siddiqi et al., 2013)).

Table 3.2: Cluster-level characteristics and Relative Risks with quit rates at 6 months- ASSIST study (2010-2011) Pakistan

Site ID	**Location**	**Mean (SD)** age in years	**Number (%)** males	**Mean (SD)** duration of smoking in years	**Median (IQR)** quantity smoked per day	**Median (IQR)** *household income
1	Rural	32.4 (10.6)	61/61 (100)	15.5 (10.1)	20 (10)	116 (81)
2	Rural	44.1 (11.9)	54/60 (90)	18.9 (8.8)	20 (13)	116 (105)
3	Rural	36.9 (12.1)	58/60 (97)	18.8 (12.3)	20 (18)	96 (67)
4	Rural	38.1 (13.2)	62/64 (97)	19.4 (10.6)	20 (14)	70 (47)
5	Urban	42.6 (13.5)	54/60 (90)	21.1 (10.7)	17 (12)	116 (116)
6	Rural	35.9 (9.5)	59/60 (98)	15.7 (8.8)	27 (23)	105 (47)
7	Rural	39.4 (10.2)	56/60 (93)	20.7 (10.2)	20 (13)	58 (58)
8	Rural	39.1 (11.2)	59/59 (100)	19.5 (9.6)	20 (30)	116 (81)
9	Rural	40.4 (11.5)	57/59 (97)	21.3 (10.7)	20 (25)	128 (116)
10	Rural	43.6 (16.1)	58/60 (97)	21.1 (16.5)	17 (14)	105 (105)
11	Urban	44.2 (11.9)	59/60 (98)	22.5 (14.1)	20 (23)	105 (110)
12	Rural	40.2 (6.2)	56/60 (93)	16.5 (6.2)	10 (6)	174 (70)
13	Rural	41.0 (13.3)	59/61 (97)	21.5 (12.2)	20 (10)	64 (47)
14	Rural	40.2 (12.2)	59/60 (93)	19.5 (11.6)	21 (11)	47 (41)
15	Rural	38.4 (14.9)	60/60 (100)	18.7 (13.5)	10 (7)	93 (38)
16	Rural	37.7 (12.1)	59/59 (100)	22.0 (12.3)	20 (7)	70 (47)
17	Urban	45.3 (13.1)	73/74 (99)	23.6 (12.8)	20 (18)	58 (70)
18	Urban	42.6 (13.7)	60/60 (100)	22.5 (13.9)	18 (15)	81 (86)
19	Rural	40.7 (13.7)	57/61 (93)	24.3 (12.2)	20 (12)	70 (47)
20	Rural	37.9 (10.6)	60/60 (100)	20.5 (11.2)	20 (10)	116 (105)
21	Rural	32.9 (9.3)	57/57 (100)	12.1 (9.9)	17 (10)	93 (47)
22	Urban	43.7 (13.8)	47/55 (86)	17.6 (10.8)	20 (8)	93 (47)
23	Urban	36.8 (11.9)	50/51 (98)	15.9 (11.2)	20 (22)	81 (58)
24	Rural	44.4 (16.9)	54/61 (89)	25.9 (17.0)	20 (30)	58 (23)
25	Rural	42.2 (12.1)	53/55 (96)	22.6 (11.8)	20 (14)	58 (41)
26	Rural	50.0 (14.5)	47/53 (89)	27.5 (12.4)	39 (10)	116 (58)
27	Urban	44.6 (16.4)	52/53 (98)	24.4 (16.0)	20 (5)	70 (67)
28	Rural	43.4 (14.8)	47/55 (86)	23.8 (13.8)	24 (18)	64 (47)
29	Rural	44.6 (16.2)	55/55 (100)	24.5 (15.2)	12 (12)	58 (47)
30	Rural	41.2 (12.0)	59/60 (98)	20.5 (10.2)	25 (15)	93 (41)
31	Rural	38.4 (12.3)	52/55 (95)	15.9 (8.5)	20 (5)	47 (29)
32	Rural	44.8 (11.7)	48/55 (87)	20.7 (9.9)	20 (9)	140 (81)
33	Rural	43.7 (13.9)	48/56 (86)	20.6 (12.6)	12 (7)	81 (58)
ǂ RR (95% CI)	0.88 (0.44-1.73)	1.05 (0.91-1.21)	0.96 (0.91-1.02)	1.06 (0.92-1.22)	0.97 (0.92-1.01)	0.99 (0.98-1.00)

* Based on monthly household income in USD ($1= 86 Pakistani Rupees)
ǂ Multilevel modelling for level 2 (cluster) predictors
IQR is inter-quartile range
RR is Relative Risk

Hence, the variation in quit rates could not be explained by the individual or the cluster-level characteristics explored using data from the ASSIST study.

3.2.3 Intervention

Variation in quit rates across clusters has been linked with variations in the process of delivery and the content of complex BIs (Boutron et al., 2008, Craig et al., 2008). The variation in smoking quit rates between TB clinics in the ASSIST study might possibly be explained by these two factors, in addition to the individual participant and cluster level characteristics. However, process evaluation of the BI was not considered in the ASSIST study.

Some brief insights into the factors that can contribute to variability in the delivery of a BI, and hence to the variation in quit rates in ASSIST study, are considered in the following paragraphs. These factors are explored further in section 3.3.

Complexity of the intervention features

Complex interventions do not simply comprise checklists of things to read out to the individual, they are more focused on how the message in the checklist is delivered to the patient. (Moore et al., 2015). In the case of ASSIST study, the BI was not only complex as it involved behaviour change and was adapted from the developed world settings, but also because it targeted vulnerable population (TB suspects), was implemented in health care settings and was embedded in a public health programme. The competences of the TB paramedics and their skills to deliver behavioural support, their motivation and intention to provide cessation support, the efficiency of the systems and organisational structure, stretching practice time to include additional cessation provision and their own smoking status all supplemented the complexity of the devised intervention (Dogar et al., 2016).

Intervention delivery

Whether the intervention is delivered in the intended way may be affected by the provider's competences in the use of intervention techniques or their behavioural attributes (e.g. empathy or being a smoker themselves) (Bellg et al., 2004, Hardeman et al., 2008). Patient-provider interaction is the mutual behaviour of the person delivering the intervention and the one receiving it that develops in a face-to-face intervention session, where this behaviour then influences the effect of the intervention on the intended outcomes.

One of the factors that can influence this mutual behaviour is the health care provider's motivation to change behaviour that can alter ways in which intervention is delivered (Lennox et al., 1998). The degree of training required by the intervention provider to learn essential skills for application of complex intervention models can also influence patient-provider interaction (Lennox et al., 1998). In addition to the expertise of the providers, the volume of patients attending the clinics for care might also influence the interaction, thus leading to unintended changes in the estimate of the intervention effect (Boutron et al., 2008).The complexity of

intervention ingredients, the provider's skill in delivering complex ingredients to the patients and the interaction between these two can impact on the estimate of intervention effect.

Degree of standardisation

Standardisation is likely to be less stringent in pragmatic RCTs that attempt to evaluate whether an intervention works under the usual conditions in which it is applied. However, to allow adequate replication of intervention ingredients, it is essential to have some degree of consistency in its delivery (Boutron et al., 2008).

Practice manuals or protocols are often developed for a structured execution of an intervention; it outlines the rationale, goals and the recommended content (e.g. the BCTs) to be delivered (Lorencatto et al., 2013a). These manuals are widely used in BIs for smoking cessation. However, it is seen that different providers delivering a BI for smoking cessation using the same manual can have widely varying cessation outcomes for the respective patients (Brose et al., 2012). It is often believed that the outcomes of interest ought to be consistent across various sites of implementation in a RCT, for it to be considered standard practice, giving higher confidence in its effect estimates (Tones, 2000, Nutbeam, 1998). However, this does not mean that the intervention should look exactly the same or completely 'standardised' in terms of its content and delivery across implementation sites, as has been thought to be paramount in the past (Campbell et al., 2000). Recent thinking on the subject projects 'standardisation' as allowing the form to be adapted while keeping the process and function of delivering it consistent (Hawe et al., 2004).

In the ASSIST study, which was a pragmatic cluster RCT, the BI was not strictly standardised, although implementation of the research protocol was strictly monitored.

3.2.4 Focus of the outcome

Outcomes might also cluster due to a specific behaviour or behaviour change that is similar in a group of people. Therefore, it is necessary to consider this factor as a likely contributor to any clustering effect in trials of BIs. It is theorised that rare behaviours with lower prevalence are associated with smaller clustering effects than behaviours with higher prevalence (Taljaard et al., 2008). Theoretically, clustering for factors with prevalence near zero or near 100 percent (i.e. everyone behaving exactly alike or absolutely different from each other) tend to be smaller than those for factors with mid-range prevalence (Campbell et al., 2005). Therefore, in minimally prevailing behaviours (where the prevalence is near zero), cues are difficult to establish; leading to less variation in human behaviours and smaller clustering effects. On the contrary, behaviours that show more variety in performance tend to cluster together more. (Campbell et al., 2005). This might provide an insight into the behaviour of intervention providers, as the process or intermediate variable that can lead to clustering of individual outcomes of the patients attending their practice.

3.3 EXAMINING SOURCES OF VARIATION FOR COMPLEX BEHAVIOURAL MECHANISMS

The intermediate procedures through which intervention activities produce intended (or unintended) effects make up the mechanisms of impact of an intervention on outcomes of interest (Moore et al., 2015). From exploration of sources of variation in the ASSIST study, it seems plausible that substantial variation is contributed by factors intrinsic to the intervention i.e. the extents to which providers follow the protocol, and the quality of delivery. These factors and their contribution to variation in outcomes of complex BIs can be captured by assessing intervention fidelity based on content-related and interpersonal behaviour change techniques (described in Chapter 4). In this section I briefly look at factors within the intervention or closely connected to it (referred to as 'intrinsic' throughout this thesis) that can affect its mechanism of action on the outcome. The literature I present here pertains mainly to the mechanisms of action of complex BIs and the factors that might influence these processes. More detailed discussion of these factors within a fidelity measurement frame will be covered in the next chapter.

3.3.1 Delivery mechanics of Behavioural Interventions and the variation

In an ideal situation, a BI containing BCTs A, B and C would be delivered consistently to the participants; in the same way and the same order as pre-specified in the practice manual (Knittle, 2014). However, in real life scenarios: i) some participants might not receive C as the time allocated for the session might have run out; ii) others might have received A, B & C in a different order (e.g. BAC, CBA); and iii) others still might have received A, B & C, but also received D, E & F that were not pre-specified in the manual. This could be considered a 'drift' rom the practice manual or an 'innovation' as described by Bumbarger et al. (2008), depending on whether the additional BCTs were as effective or less so than A,B & C (described further in chapter 4). Alternatively, the order in which they were delivered might have been tailored by the provider according to participant's choice. Such variations in delivery of a BI can affect outcomes and intervention effectiveness overall. If not accounted for and explained, such variations could hinder the application of the findings from high quality studies (Knittle, 2014).

3.3.2 Controlling variation in intervention delivery practice

In addition to the content that is delivered, an important factor affecting variation in outcomes depends on how delivery of this content is achieved (Carroll et al., 2007, Montgomery et al., 2013). Facilitation or support strategies are sometimes used to enhance intervention delivery according to protocol, with an aim to reduce variation in practice (Carroll et al., 2007). The adequacy of these strategies to standardise practice depends on the degree of complexity of an intervention. These are considered even more important for complex interventions, which are often multi-faceted and, therefore, more vulnerable to variation in the delivery process (Medical Research Council, 2000).The crux of standardising practice across implementation sites is a clear

specification of the intervention features (both its content and interaction quality), as well-defined interventions are more likely to be delivered uniformly than less well-structured interventions (Dusenbury et al., 2003, Carroll et al., 2007).

Strategies that can facilitate uniform intervention delivery practice include practice manuals, guidelines, training, monitoring, cultural considerations and feedback (Bellg et al., 2004, Gearing et al., 2011). These strategies, if not used adequately, can themselves lead to more variation in delivery practice and hence, differential effects of the intervention on respective outcomes.

3.3.3 The impact of provider expertise on variation

Differences in delivery practice between providers can be related to the provider's expertise. Controlling for these can ensure standard quality across various implementation sites (Bellg et al., 2004). The characteristics and behavioural interactions of the intervention provider and their qualifications and expertise play a significant role in delivering the intervention as intended. Providers who are motivated for the change, and believe that it is specifically needed to bring about the desired benefits, feel more confident in their ability to deliver the intervention (Durlak and DuPre, 2008). Those who have the requisite skills are also more likely to deliver it with better quality (Dane and Schneider, 1998). There is evidence that, although the concepts of such models (complex BIs) themselves are easy to grasp, it is much more difficult to acquire a useful knowledge of the associated intervention delivery skills (Hall et al., 1997). Therefore, short training sessions (e.g. one day) might not be enough to gain and retain effective skills for delivering complex interventions, even for those intervention providers who are motivated to change behaviour (Sereno et al., 2012, Lennox et al., 1998).

BIs are complex, as they involve several interacting features targeting different behaviours, but implementing them to have the intended impact is even more complex, because it relies on a certain degree of skill of the provider. Moreover, it requires an in-depth understanding of the behaviour that needs changing e.g. a proper understanding of the concept of 'tipping points' in smoking cessation and utilising this knowledge to the advantage of the recipient. Tipping points refer to a sudden successful attempt as a result of a single event (or perfect motivational level achieved) by an individual who might have tried several times to quit smoking in past but failed (Resnicow and Vaughan, 2006). BIs may therefore be delivered at the 'right' moment to trigger change, whereas delivery to the same individual at a different time might not have the desired effect (West, 2013). Also, understanding or identifying this 'right' moment tailored to each individual's situation requires knowledge of the latest techniques that are effective for behaviour change, to build on relevant skills.

3.3.4 Aspects of patient-provider interaction and the variation

The concept of 'patient-centred focus' appropriate for negotiating behaviour in complex interventions is not new e.g. motivational interviewing (Miller and Rollnick, 2013), but it is very different from the predominantly practised 'professional-centred focus,' where the health professional is the expert giving advice. In PHC settings, the latter is the usual case, which could be an important factor in the consistent application of intervention features and might lead to variations in their delivery (Lennox et al., 1998).

Participants tend to interact in context with interventions, resulting in outcomes produced by these interactions, rather than passively receiving them (Pawson and Tilley, 1997). Therefore, the mechanism of impact of an intervention on the outcome cannot simply be controlled by setting standards of practice and adequate expertise of the provider, but also requires an in-depth understanding of and training on patient-provider interaction features and how these work in regard to BIs. In the context of individual BIs delivered in settings like a TB programme in a LMIC, in a PHC by personnel not primarily skilled in behaviour change counselling (e.g. the TB paramedics), an emphasis on building efficacy on relevant patient-provider interaction skills can be beneficial in reducing variation.

3.4 SUMMARY

In summary, description of data and analysis of the individual and cluster level characteristics from the ASSIST study, do not reveal any contribution of these factors to the clustering effect. The literature and empirical evidence on the subject of clustering, in outcomes of complex interventions, highlights factors more intrinsic to the intervention. As discussed in this chapter, these factors might relate to behaviour change focus of the intervention, the nature of the ingredients, the way these ingredients are delivered, and the level of their standardisation across clusters, as highlighted below in Figure 3.2.

Even when interventions are designed in a standard format, implementation occurs differently in different contexts, showing many different effects (Peters et al., 2013). Understanding the variation in cessation outcomes of complex interventions is critical to making these interventions work as well as possible. Cluster RCTs, in which the same intervention may be delivered and received in different ways, might present an opportunity to characterise and examine several potential sources of variation, including those explored in this chapter. However, before setting out to evaluate these effects to interpret variation, identifying the sources of variation and mapping aspects of complexity in the intervention onto the appropriate sources is a necessary step (Petticrew et al., 2013).

Such features, intrinsic to a complex intervention involving behaviour change, and the ways to capture or assess their role in explaining variation in patient outcomes, will be studied further in the next chapters of this thesis.

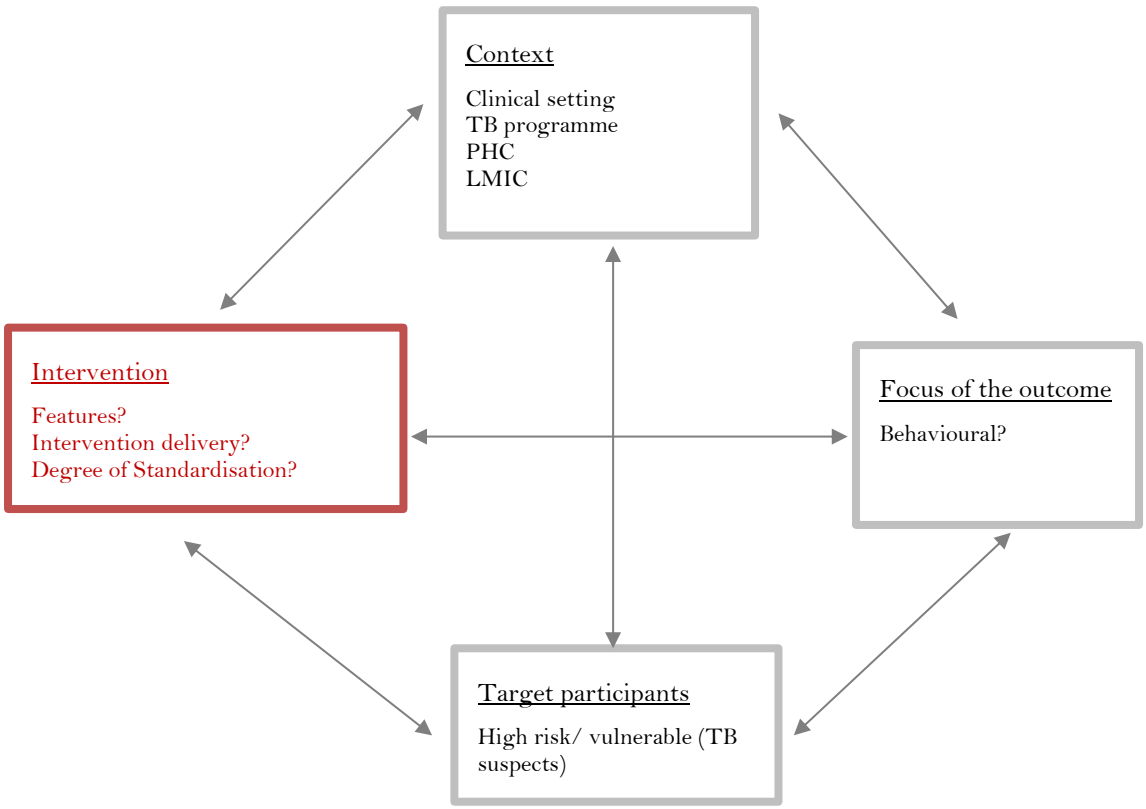


Figure 3.2: Sources of variation in quit rates as identified for further exploration

Chapter 4. Description of Intervention Fidelity

As the intervention delivery process can introduce variation in smoking quit rates, this chapter presents the evidence for assessment of fidelity to intervention delivery. The literature reviewed pertains mainly to measurement of fidelity and its conceptualisation in terms of complex BIs. The focus remains on the identification of active ingredients of the intervention, including the quality related aspects of its delivery.

The appraisal of evidence on the link between smoking and TB (Appendix A) demonstrates the need for integrating smoking cessation interventions in TB care and management practice. Scientifically tested and effective BIs to support smoking cessation exist, both in general smokers and in those with TB symptoms. However, as highlighted in the previous chapters (2 & 3), optimizing delivery of such interventions is hindered by the limited understanding of variation in their effect- partly because the active ingredients of the intervention are not known. To understand this variation in effect with complex BIs, potential contributing factors need to be examined, as theorised in Chapter 3.

Intervention 'fidelity', an implementation variable (Peters et al., 2013), encompasses the three main contributing factors related to intervention delivery: namely 'what' is delivered, 'who' delivers it and 'how' it is delivered (Carroll et al., 2007). As conceptualised and presented in Chapter 1 (see Figure 4.1), intervention fidelity forms the core of the interpretive evaluation model for studying intervention effect in all of its variation. The term 'fidelity' in this thesis will refer to two main concepts: providers' adherence to intervention content ('depending on what' is delivered) and the quality of interaction (depending on 'who' delivers and 'how' it is delivered).

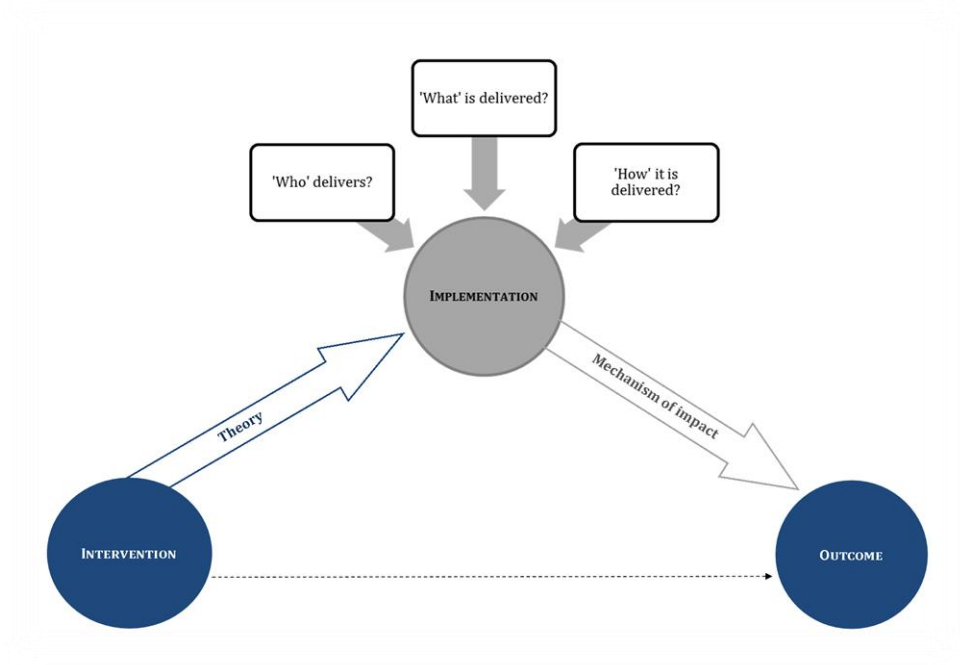


Figure 4.1: Interpretive Evaluation Model

4.1 THE CONCEPT OF FIDELITY MEASUREMENT

The key research question for my thesis, regarding explaining variation in smoking quit rates (presented in Chapter 5), emerged from the ASSIST study. Since, individual patients' responses (i.e. the quit rates) were clustered by implementation sites (single intervention provider per site); an inquiry into the reasons for this clustering of outcomes would require measurement of factors influencing intervention delivery. As hypothesised (in chapter 3), fidelity measurement might help describe the mechanism of impact of the BI on its outcomes (Moore et al., 2015). However, first I will describe the concept of measurement of fidelity and the things to consider when measuring fidelity for complex behavioural mechanisms.

4.1.1 Quantifying fidelity as opposed to subjective assessment

Despite the preference of using objective measures of fidelity for BIs (Knittle, 2014), progress in developing fidelity measures (including scales and indices) has been hampered by several factors in the past. Similar to the psychotherapy literature, complex BIs have lacked even basic descriptions of intervention models and details of their active ingredients (Lorencatto et al., 2013a). The description of BIs is further complicated by the complexity of interactions of behaviour change, such as the causal linking of BCTs with the behavioural determinants of the targeted individuals, as discussed in chapter 1 (under 'The logic model of behaviour change').

Nonetheless, psychotherapeutic research realised the potential of fidelity measures much earlier than behavioural or health sciences research. The actualisation of this concept of objectivity in fidelity measurement for BIs targeting tobacco use derives from this literature (Bond et al., 2000b). Additionally, concepts and methods from fidelity assessment in health and education will also be used (Mowbray et al., 2003, Streiner and Norman, 2008).

4.1.2 Origination of fidelity measures

Fidelity measures originated in psychotherapeutic research in the 1960s, when the need to measure elements of psychosocial interventions was first recognized, due to the difficulty in interpreting outcomes in studies (Moncher and Prinz, 1991). Psychotherapy approaches were fundamentally different from each other and poorly defined, while practice varied greatly among therapists (Bond et al., 2000b); which is similar to the observations from ASSIST study. This led to the development of the first 'fidelity measures' in a scale format: the process rating scales for client-centred therapy (Bond et al., 2000b).

Fidelity measurement advanced in two directions, the intervention integrity, 'extent to which intervention is implemented as intended ', and its differentiation, 'whether interventions differed from one another in the intended manner' (Moncher and Prinz, 1991). Concurrently, two aspects, the provider 'competence' (Waltz et al., 1993) and the 'dose delivered' (Scott and Sechrest, 1989), started being considered as elements essential to measuring fidelity. Soon the need to disseminate

active intervention ingredients was recognized, resulting in operationally defining intervention features and linking them to the theory (Bond et al., 2000b).

4.1.3 Where does the science of behaviour change stand? – Articulating the causal assumptions of complex behavioural interventions

A decade later, complex BIs for smoking cessation are at an identical stage. The potential of fidelity measures having valid links to theories of behaviour change is acknowledged (Hardeman et al., 2005, Cane et al., 2012, Lorencatto et al., 2013a, Michie et al., 2014-2017), and progress to operationalise these ideas using reliable methodologies is underway (Nelson et al., 2012, Moore et al., 2015, Bellg et al., 2004, Borrelli et al., 2005).

The taxonomies of BCTs (Michie et al., 2011c, Lorencatto et al., 2013b, Michie et al., 2013) have so far been used to assess fidelity (Lorencatto et al., 2013a) in terms of presence or absence of BCTs. These provide a coding framework for identifying and categorizing the BCTs within a BI and have been used reliably to code English Stop-Smoking Service treatment manuals (Michie et al., 2011b, West et al., 2010) and transcripts of audio-recorded BI sessions delivered by these services (Lorencatto et al., 2013b).

Behaviour change is complex and a description of the features of a BI by simply coding 'absent' vs. 'present' might not be adequate for identifying active ingredients and processes by which the intervention leads to behaviour change (Hardcastle et al., 2016). The above taxonomies, although used to describe BIs 'compositionally' have not been used to score fidelity of BIs in relation to the 'functionality' (explained in the next section) of BI features. This later concept surpasses the simplicity in coding 'present vs. absent' for fidelity assessment and could help identify the active ingredients of a BI, assuming that the devised fidelity measure is found to be reliable and psychometrically valid.

The MI approach for behaviour change proposes that a description of an intervention should entail both content and relational techniques (as briefly mentioned in Chapter 1). That is, techniques that relate to what is included in an intervention (its content) and how it is delivered (interaction quality) (Hardcastle et al., 2016, Hagger and Hardcastle, 2014). Coding systems (including 'motivational interviewing treatment integrity-MITI' and 'motivational interviewing skill code-MISC') designed to assess specific domains of patient-provider interaction within the MI approach have been developed to assess provider adherence to and competence in intervention delivery (Moyers et al., 2010, Moyers et al., 2005, Madson and Campbell, 2006). These, however, remain true to the 'spirit' of MI (Hagger and Hardcastle, 2014), which comprises four components: collaboration, evocation, autonomy and compassion (Miller and Rollnick, 2013), described using the 'relational' techniques (Hardcastle et al., 2016). Furthermore, these fidelity measures were designed with the intention of using them in a training and supervision

environment and not for research into the mechanisms of impact of an intervention (Madson and Campbell, 2006). Further insight into existing fidelity indices is provided in Chapter 6.

The importance of generic or inter-personal competences, focusing on the actions of the provider in delivering BI content to individuals (Roth and Pilling, 2008), such as those identified for smoking cessation (Michie et al., 2011b) is also acknowledged by behaviour change scientists. However, these competences have not been used to describe interaction quality, particularly in relation to its 'functionality'.

Of further note is that these taxonomies have not been used to code BIs outside of the UK context and it is not known if they can reliably code such interventions for the purpose of assessing fidelity.

4.1.4 The fidelity and adaptation debate

Most of the theoretical models and concepts recommended by UK MRC for evaluating complex interventions emphasise the need to understand 'what' was delivered and 'how' it was delivered (Moore et al., 2015). These models assume that the intervention must have some standardised features in different settings (Moore et al., 2015), that can form the basis of the evaluation. However, as discussed in Chapter 3, standardising the practice of delivering complex interventions is a subject of controversy among researchers and one to be carefully considered when measuring intervention fidelity.

Fidelity measurement is not straight forward for complex interventions. In evaluations where active ingredients need to be identified and selected, strict standardisation may be required across practices (Craig et al., 2008). However, some aspects of the intervention might require adaptation to the local context or individual patients and would not be possible to standardise completely (Craig et al., 2008). Some researchers argue that when little is known about the active ingredients of an intervention, allowing adaptations in the content to be delivered might inhibit effectiveness (Mihalic, 2004). Others advocate distinguishing between the fixed and variable aspects of an intervention, for example, the BCTs might be considered as fixed, while their structure and application process in different contexts might be considered as variable (Hawe et al., 2004). Variable aspects can then be adapted to the local context. Yet another school of thought suggests keeping the core intervention features as fixed and less critical features as variable (Durlak and DuPre, 2008, Firpo-Triplett and Fuller, 2012). Working towards a compromise, this last concept can help achieve the balance between fidelity and degree of standardisation permitted.

A slightly different concept from distinguishing between the fixed and variable ingredients or the core and less critical ingredients is looking at fidelity and adaptation under the same spectrum, while distinguishing between 'innovation' and 'drift' (Bumbarger and Perkins, 2008). 'Innovation' is described as the skilful attempts at tailoring intervention to better fit the target

population or local context needs, while 'drift' is considered as the unintentional shortcomings that can act as the barriers to full implementation (Bumbarger and Perkins, 2008). Hence, according to this concept, skilful providers are permitted to deviate from the protocol in response to feedback from the participants, while carefully remaining consistent with the underlying intervention theory. This concept is more easily applicable to the theories of BIs, as the delivery of intervention relies heavily on the skill of the provider.

In addition, integrity of an intervention is described as having the 'dose' delivered at an acceptable level and in the same way at each site (Hawe et al., 2004). Integrity in terms of a complex intervention needs to be defined 'functionally' rather than 'compositionally' (Hawe et al., 2004). Therefore, standardisation by means of 'delivering the same intervention consistently across sites' for a complex intervention is an oversimplification of the concept. In complex interventions, the 'functionality', which can be described as the process of delivery, can be standardised, while keeping the intervention content ('composition') adaptable (Hawe et al., 2004).

In my thesis the balance between standardisation and adaptability of intervention features, in establishing a quantitative measure of fidelity for BIs, is based on the concepts of functionality and innovation/drift theory by Hawe et al. (2004) and Bumbarger et al. (2008).

In order to identify and quantify the active ingredients of an intervention, first a fuller description of individual features of the intervention is needed (Lorencatto et al., 2013b). We then need a 'measure' of the delivery of these features, which would not only be reliable in measuring this construct (fidelity to BI) under study, but would also have the potential to explain variation in quit rates.

4.2 Derivation of a Fidelity measure for Behavioural Interventions

From discussions in previous chapters, it is clear that adherence to BI content and the quality of its interaction may influence smoking cessation outcomes amongst intervention recipients. The concepts exist for evaluating the delivery process of complex interventions (Hawe et al., 2004, Bumbarger and Perkins, 2008) and the methods for quantifying fidelity in general (Bond et al., 2000b, Mowbray et al., 2003, Streiner and Norman, 2008). But how do we judge that the BI is being delivered as intended? This will first require some degree of 'description' of the intended or planned features of the intervention.

In this thesis, fidelity is conceptualised as a combination of 'what' is delivered and 'how' it is delivered; two very important aspects of the process of intervention delivery. Answering these questions requires understanding the 'intended or planned' practice and the 'actual' practice.

4.2.1 Intervention delivery: 'Intended practice'

Intended practice is the specification and description of the intervention content and the processes involved in its delivery, often in the form of practice manuals and the training offered to enhance effective delivery skills. Practice manuals are structured, procedural booklets outlining the rationale and goals of an intervention, as well as the recommended content to be delivered while administering an intervention (Lorencatto et al., 2013a).

Going back to the concept of 'what' is delivered, 'who' delivers it and 'how' it is delivered, it is clear that although intervention features can be simple, the interaction between them and behaviours of those involved can be complex; leading to variation in these concepts. Nevertheless, the starting point for fidelity measurement can be the specification of the features of an intervention that are key to its execution. This entails classification and labelling of the content and the interaction quality of the BI.

Clear definitions and descriptions of an intervention are vital to gain an accurate appraisal of applicability and validity of the findings from its evaluations (Moore et al., 2015). Considerations for describing complex interventions include description of its key features and processes, as well as uniform labelling of the active ingredients. Describing complex interventions is a science in itself; frameworks to guide this scientific process are also summarised in this section.

Adequacy of intervention description

Inaccurate descriptions of an intervention can lead to its content being replicated with poor fidelity (Lorencatto et al., 2013a). Interventions with clear specification of the core effective ingredients are more likely to result in uptake in practices that require cessation services (Michie et al., 2009c). Published descriptions of BIs for smoking cessation are found to report less than half of the BCTs specified in the intervention manuals (Lorencatto et al., 2013c). Failure to adequately report intervention features could undermine attempts to replicate and optimize intervention delivery in practice (Lorencatto et al., 2013c). Furthermore, it could jeopardise the development of a fidelity measure, if it fails to capture all of the key intervention features. Hence, incomplete descriptions of an intervention would leave the reader confused about whether or not that intervention works, with very little indication of the active ingredients that worked and why (Lorencatto et al., 2013c).

Frameworks for describing interventions

Intervention description involves more than merely labelling and providing a list of the various ingredients (Hoffmann et al., 2014). Key characteristics, including duration, intensity, mode of delivery, essential processes and monitoring, could all influence the replicability and ideal implementation of an intervention (Hoffmann et al., 2014).

An extension of the CONSORT (Consolidated Standards of Reporting Trials) statement for Non-pharmacologic treatments (pertaining to BIs) indicates reporting on the following intervention features (Boutron et al., 2008):

- Precise details of the intervention and comparator, as intended and as implemented
- Description of the various ingredients of the interventions
- Description of the procedures for tailoring the intervention to individual participants
- Details of how the intervention was standardised
- Details of how adherence of providers with the protocol was assessed or enhanced
- Description of the providers (case-volume, qualification, expertise etc.)

Another useful guide, 'The Template for Intervention Description and Replication' (TIDieR) provides a checklist for improving the completeness of reporting for health care interventions (Hoffmann et al., 2014). The 12 item TIDieR checklist is an extension of the CONSORT 2010 statement (http://consort-statement.org) and the SPIRIT 2013 statement (http://spirit-statement.org), consisting mainly of items on: brief name, why (rationale, theory or goal of the key features), what (materials/tools used), what (procedures, activities and/or processes), who provided (provider expertise, background, training), how (mode of delivery), where (settings and location), when and how much (number of sessions, schedule, duration and intensity of dose), tailoring (if intervention was personalised, titrated or adapted), modifications (any changes made in the intervention), how well-planned (how fidelity was assessed and by whom), and how well-actual (the extent to which intervention was delivered as intended).

Of particular interest to BIs reporting and perhaps most challenging to describe are the items, namely, 'what procedures' and 'who provided' (Johnston, 2014). The TIDieR item 'what procedures' basically involves both the processes that reflect the content and the way this is delivered. The TIDieR item 'who provided' also relates to the inter-personal aspects of the interaction and hence the quality of delivery.

Therefore, in short, intervention features according to the above mentioned frameworks fall into two broad categories: one, mode and style of delivery that relate to the quality of the interaction, and two, the content that is delivered (Lorencatto et al., 2013c).

Uniformity in intervention description

A classification or labelling standard for specifying BIs is the ultimate goal towards standardising intervention delivery practices (as outlined in 'Controlling variation in intervention delivery practice' in 3.3.2). Additionally, it finds its use in reporting of findings and replication of the content of interventions, which is vital to designing process evaluations of complex interventions.

The taxonomies of BCTs (Michie et al., 2011b, Michie et al., 2011c) are theory-linked nomenclatures (Michie et al., 2011a) that provide a common language for uniform coding of intervention ingredients in manuals and transcripts of BIs, both in research and in practice (Lorencatto et al., 2013c). Forty-three BCTs have been identified (so far) and labelled for use in describing content of BIs for smoking cessation (Michie et al., 2011c). Each BCT has a specified criterion for its operationalization, is defined using consistent terminology, and has a clear label that can be used to categorize and consistently report intervention features.

The taxonomies for BCTs concerning smoking cessation and the competences to deliver these BCTs have been described in detail in Chapter 1 (section 1.3.2). Whilst the BCTs focusing on individual behaviour change 'content' reflect the information and knowledge provided to intervention recipients to promote behaviour change (e.g. exploration of pros and cons), the competences for specifying 'quality of delivery' reflect the inter-personal way in which the content-based BCTs are presented by the provider to increase their effectiveness (Hardcastle et al., 2016).

The interpersonal aspects, that is, 'who' delivers it and 'how' it is delivered can be described using the generic competences (Michie et al., 2011b) for intervention delivery (as summarised in Chapter 1) and content-based aspects, that is, 'what' is delivered can be described using the BCTs specific for smoking cessation (Michie et al., 2011c, Lorencatto et al., 2013b).

Clarifying assumptions about how the Behavioural Intervention works

I have previously (in Chapter 1) presented the 'advances in behaviour change science' by providing details of the logic model of behaviour change (Hardeman et al., 2005, Michie et al., 2008), the PRIME theory of motivation for smoking cessation (West, 2009), identification of the behavioural determinants using the TDF framework (Cane et al., 2012, Michie et al., 2005) and the development of taxonomies to label the BCTs (Michie et al., 2011b, Michie et al., 2011c). In this section I briefly describe the 'basis of behaviour change science' in an attempt to bring together the various concepts and frameworks described so far.

The core of the taxonomies (Michie et al., 2011b, Michie et al., 2011c) for behaviour change is formed from the COM-B model (Figure 4.2), which acts as a 'behaviour system' analysing behaviour in terms of the extent to which an individual's capability, opportunity and/or motivation need to change for a certain behaviour to change (Michie et al., 2011d). This COM-B system forms the hub of the Behaviour Change Wheel, which is a synthesis of 19 frameworks of behaviour change interventions.

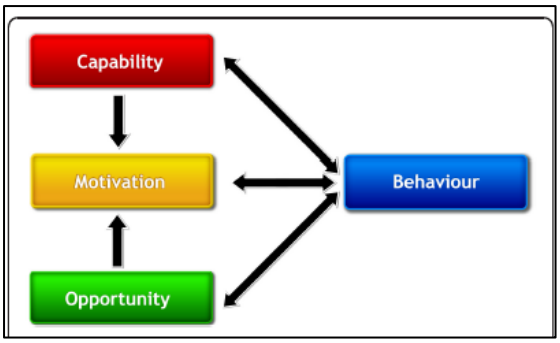


Figure 4.2: The COM-B system- a framework for understanding behaviour

(Adopted from: Michie S, et al. The behaviour change wheel: a new method for characterising and designing behaviour change interventions. *Implementation Science*. 2011;6:42.)

The Behaviour Change Wheel was developed to inform the process of selection of interventions according to an analysis of the nature of a certain behaviour, the mechanisms that need to change in order to change that behaviour, and identification of the points to intervene with those mechanisms (Michie et al., 2011d). The behavioural determinants identified using the TDF have also been mapped to the COM-B system (Figure 4.3) (Cane et al., 2012, Michie et al., 2011d).

COM-B component		TDF Domain
Capability	Psychological	Knowledge
		Skills
		Memory, Attention and Decision Processes
		Behavioural Regulation
	Physical	Skills
Opportunity	Social	Social Influences
	Physical	Environmental Context and Resources
Motivation	Reflective	Social/Professional Role & Identity
		Beliefs about Capabilities
		Optimism
		Beliefs about Consequences
		Intentions
		Goals
	Automatic	Social/Professional Role & Identity
		Optimism
		Reinforcement
		Emotion

Figure 4.3: Mapping of the behaviour change wheel's COM-B system to the TDF

(Adopted from: Cane J, et al. Validation of the theoretical domains framework for use in behaviour change and implementation research. Implement Sci. 2012;7(1):37.)

The taxonomies of BCTs and competences to deliver them are standardised nomenclatures developed to facilitate communication in literature and to label ingredients of intervention models (Michie et al., 2011c). These cannot be considered fidelity criteria. They can, however, serve as the basis for developing a standard measure for scoring fidelity of BIs for smoking cessation.

In summary, a reasonable account of fidelity for BIs will entail standardising the functionality of the intervention represented by the procedures (mode and style) of its delivery, while keeping its composition or intervention features adaptable. Working on the principles for describing

interventions, as discussed in this section, the taxonomies can be used to code BIs and disaggregate the intervention features into ingredient BCTs and competences, in a standard format that can be potentially scored for fidelity.

4.2.2 Intervention delivery: 'Actual practice'

Once intervention description is uniform and adequate and in line with the causal assumptions of the underlying theory about how the intervention works, the core structure is ready on which to base fidelity measurement. 'Actual practice' represents the extent to which BCTs and competences mapped to the BI (the core structure) are subsequently delivered. It refers to the extent to which intended practice is actualised adhering to specifications in the practice manual (Bellg et al., 2004, Borrelli, 2011) and competences acquired through training.

In essence, actual practice is the fidelity to the intended practice.

In this section, I will explore the elements of fidelity essential to capture the actual practice that can be applied to the core structure of intervention features for scoring. The elements, as used by others to assess fidelity, will be described and appraised for inclusion in my research, with a particular focus on BIs and the elements that are intrinsic to the intervention delivery process.

Theoretical perspectives for examining fidelity

Early evaluations of intervention delivery process aimed at monitoring whether interventions were delivered as intended, in order to determine the extent to which changes in the outcomes were a true reflection of the effect of intervention theory (McGraw et al., 1989, Pirie et al., 1994). Several theoretical perspectives have since emerged, as the demand for process evaluation increased, shifting the focus towards understanding how implementation is achieved and moving away from simply monitoring what is delivered (Moore et al., 2015). While it is beyond the scope of this chapter to provide an exhaustive review, this section briefly describes concepts that are commonly used in examining fidelity to complex interventions.

At the start of this century, 11 priority areas for understanding the process of intervention delivery were identified: recruitment, maintenance, context, resources, implementation, reach, barriers, exposure, initial use, continued use and contamination (Baranowski and Stables, 2000). Soon after, Steckler and Linnan (2002) narrowed down the focus to six priority functions for evaluation, namely: context, fidelity, dose delivered, dose received, reach and recruitment (described in next section). Later, a theoretical model which defines fidelity as a combination of content, frequency and duration of delivery, and coverage was proposed (Carroll et al., 2007). This is very similar to Steckler et al's. (2002) definition of implementation being a combination of 'fidelity, dose and reach'.

However, Carroll et al's. (2007) model goes further and describes the 'moderators' of fidelity, as essential elements to be considered in addition to adherence. It describes intervention complexity, facilitation strategies, quality of delivery and participant responsiveness as the main factors that may moderate the degree of fidelity with which the intervention is delivered. As described by the UK MRC framework for process evaluation, these present-day models typically focus on quantification of fidelity, receipt and context (Moore et al., 2015). However, when studied in depth these models only describe the key priorities for investigating intervention delivery process and not necessarily the methods suitable for carrying out such assessments.

A recent model introduced for cluster RCTs demonstrates methods suitable for investigating these key priorities and the consideration of different timings in the process being evaluated (Grant et al., 2013). Besides, it outlines the methods for evaluating the processes during trial execution that can be used to explore variations in effectiveness between implementation sites. Despite the fact that these focus primarily on design and reporting of process evaluation conducted parallel to the trial execution, some of the aspects (e.g. the mechanism of impact) have the potential to be adapted for a post-hoc (after completion of trial) exploration.

These models provide a comprehensive list of elements important for assessing intervention fidelity. The selection of elements for any evaluation, however, depends completely on the purpose of the fidelity assessment (Peters et al., 2013); i.e. monitoring delivery versus understanding the mechanism of impact of delivery. These latest models for understanding what was delivered in an intervention, how it was delivered, and its mechanism of impact can be promising for the design and conduct of an appropriate exploration of the intervention delivery process and might help understand the range of its effects on smoking cessation.

'What' is delivered: Intervention content

Intervention fidelity is noted as a potential moderator of the relationship between an intervention and its intended outcomes (Carroll et al., 2007). Fidelity has been defined as the 'adherence of the intervention delivery to the specifications in intervention manuals' (Bellg et al., 2004).

Multiple descriptions of the important aspects of fidelity have been in use in the literature. Some define fidelity as the quality of delivery, emphasising the need to capture the qualitative nature of what was delivered, not just the technical aspects of delivery (Linnan and Steckler, 2002). Others argue that fidelity specifically concerns whether core, prescribed intervention ingredients are delivered, rather than the separate yet associated question of how they are delivered, for example, in terms of quality and tailoring of delivery (Lorencatto et al., 2013a). In respect to capturing intervention content adequately, the most recent thinking on the elements of fidelity for consideration is discussed below.

Elaborating on the elements for understanding the process of intervention delivery put-forward by Linnan and Steckler (2002), I attempt to describe these elements one-by-one, in the next few

paragraphs, and provide justification for selecting those that are important for fidelity measurement of a BI, particularly keeping the focus on those elements that relate to the intervention delivery.

'Context' can include anything external to the intervention e.g. the organisational structure or environment, the provider's readiness or ability to change that can be influenced by pre-existing circumstances, skills, resources and attitudes (Lawton et al., 2015, Moore et al., 2015). For the purpose of accounting fidelity in the current research, context encompasses contributing factors at the cluster (organisational) level, which might influence fidelity or moderate its effect but is not considered an element of fidelity. This is based on the fact that cluster-level variables were not found to explain the variation in quit rates in the ASSIST study, as described in Chapter 3. The other connected but separate concept of providers' skill and attitude being considered a part of 'context' (Lawton et al., 2015), makes sense when looking at the micro and meso factors concerning intervention delivery. However, this concept of the context for an intervention in intrinsic terms better fits with the quality of delivery ('how it is delivered'), and will be discussed in the next section.

'Adherence' refers to the extent to which the intervention is delivered as planned according to the protocol (Dane and Schneider, 1998). Adherence is the most fundamental element of fidelity; if an implemented intervention adheres completely to the content, frequency and duration as designed, then the fidelity is likely to be high (Carroll et al., 2007). The content of an intervention may be seen as its 'active ingredients', while frequency and duration of the intervention (sometimes referred to as 'dose delivered)' are essentially sub-categories of adherence (Carroll et al., 2007). Dose delivered or intensity of intervention, also called 'provider commitment' (Lawton et al., 2015) is the time spent in delivering the intervention. It is sometimes quantified by the number of sessions of therapy given, in addition to the duration of each session. Adherence will form an integral aspect of the fidelity measurement in this thesis; it will help establish scoring criteria for the core structure (intended practice) described for the BI.

'Dose received' or exposure is the extent to which the recipients of intervention report receiving the intervention features (Dane and Schneider, 1998, Lawton et al., 2015). Intervention receipt could be assessed by monitoring session attendance and key features delivered in the session, along with comprehension of these features by the recipient (Gearing et al., 2011). In my research, devising a measure that can cover this aspect of fidelity is difficult, as it could not be directly observed and would require comprehension by the intervention recipients. However, considering the complexity of behavioural interactions between the provider and the recipient in BIs, this element can be affected by the quality of delivery. It can be considered an indirect consequence of patient-provider interaction, rather than a direct product of fidelity itself and, therefore, might be covered under quality of intervention delivery.

'Reach' is the extent to which the intervention programme is received by the targeted group and 'recruitment' relates to strategies used to attract participants for the intervention programme

(Linnan and Steckler, 2002). Despite its importance in implementation for scale-up at a wider level, the extent to which reach can be considered a necessary aspect of intervention fidelity depends on how strongly it is linked to the effectiveness of a specific intervention. To elaborate on this point, consider the example of a smoking cessation referral scheme which cannot be implemented if no one attends, while a national tobacco taxation policy will impact even those who do not know such a policy exists (Moore et al., 2015). For the purposes of the current research, and considering the nature of individual BIs, as compared to mass interventions for smoking cessation, (where coverage for an optimal impact is necessary), this element falls outside the scope of the proposed fidelity measurement.

Therefore, adherence to the intended content for delivery is one of the main elements of fidelity that will be considered for measuring 'what' is delivered in my research work.

'Who' delivers and 'How' it is delivered: Quality of interaction

Quality of interaction is another important aspect of fidelity with which the intervention is delivered. It is considered by some as an element of fidelity (Lawton et al., 2015, Mihalic, 2004, Dane and Schneider, 1998), while by others it is regarded as an associated factor that moderates fidelity to intervention delivery (Carroll et al., 2007, Lorencatto et al., 2013a).

Quality of interaction necessarily concerns whether the content of an intervention is delivered in a way appropriate to achieving what was intended and relies on factors like provider expertise (skill and competence) (Gearing et al., 2011), training, enthusiasm, attitude and ongoing supportive monitoring (Resnick et al., 2005, Dane and Schneider, 1998, Lawton et al., 2015).

If the content of an intervention is delivered badly, then this may affect the degree to which full fidelity can be realised (Carroll et al., 2007). The way in which intervention content is expressed to the recipient can mediate (explain the relationship between an intervention and outcome) (Hasson, 2010, Carroll et al., 2007) the impact of the intervention on cessation outcomes or moderate (influence the strength of relationship between the intervention and outcome) the effect of adherence to its content in explaining this mechanism (Pawson and Tilley, 1997). As taxonomies evolve, they need to identify and incorporate the inter-personal techniques or competences that fulfil the conditions to be satisfied if the intervention is to be effective (Hagger and Hardcastle, 2014). It is also important to determine whether these inter-personal competences moderate the effect of the content-based BCTs (Hardcastle et al., 2016).

All these factors as sources of variation in outcomes of complex interventions have been discussed previously in detail, in Chapter 3.

4.3 SUMMARY

Fidelity is delivery of the intervention according to the plan or protocol. In this chapter, I have described the concept of fidelity and its elements important for assessing the process of intervention delivery.

Two main concepts covering fidelity in connection to the intrinsic nature of an intervention are: 'what' is delivered, or the individual behaviour change content and 'how' it is delivered, or the quality of interaction. The extent to which adaptation of the intervention to local context represents poor fidelity or beneficial tailoring to client needs, is a consequential consideration when prioritising elements to assess fidelity (Moore et al., 2015). An acceptable account of fidelity for BIs might be achieved by standardising the functional features of the intervention, such as the interaction style, while keeping its compositional features (both content and quality of interaction) adaptable to the patient need.

Quantifying fidelity or developing a measure on which fidelity can be scored is more complicated than assessing fidelity subjectively. It is further complicated for BIs because of the interacting nature of the behaviours and intervention features involved. Therefore, to account for fidelity of a BI quantitatively it is essential to first describe the intervention features using a standard nomenclature. This gives the fidelity measure its core structure or the back bone, which can be referred to as the 'intended' practice.

The elements of fidelity deemed important for the current research include adherence to the content of the BI intended for delivery and the quality of the interaction with which the content is delivered. This can be referred to as the 'actual' practice and thus represents intervention fidelity.

The theoretical perspectives and concepts described in this chapter will help establish the methods for measuring fidelity to BIs (chapter 6 & 7), and whether fidelity is associated with smoking quit rates (chapter 8). Such a fidelity measure does not exist for BIs, particularly in connection to smoking cessation. Furthermore, it provides an opportunity to step forward, in the science of fidelity measures, and establish a method for accounting the quality of interaction as an aspect of fidelity within implementation and execution of a BI.

PART II- AIMS & OBJECTIVES

This section contains one chapter, in which I will outline the overall aim and describe the specific objectives of this thesis.

Chapter 5. Aims and Objectives

From the literature presented in previous chapters, it is evident that smoking is associated with significantly poor outcomes of TB disease, mortality and recurrence. Despite its proven significance, there is poor uptake and implementation of smoking cessation within TB care and management services in LMICs. This is partly due to our limited understanding of what works in these settings. Policy makers and programme managers are likely to implement shorter versions of interventions if they do not have enough staff or resources to implement the full BI package. Giving them options to select the most effective features of a BI delivered in the most efficient manner might improve the uptake of smoking cessation services in routine clinical care. A quantitative scale, with the capability to measure what is being delivered (fidelity) and which components are effective, is likely to help. However, fidelity to BIs for smoking cessation is rarely assessed and our understanding of the contribution made by the fidelity to patient quit rates is limited. Moreover, the research tools necessary to carry out a robust measurement of intervention fidelity are also lacking.

My overall research goal, therefore, is to investigate whether smoking quit rates are associated with the fidelity with which a BI for smoking cessation is delivered.

My aim, in this thesis, is to carry out the preliminary research required to answer the above question. This includes first designing a fidelity index, and then delineating features of a BI for smoking cessation in a format suitable for rating. This could determine those features that are effective, ready to be replicated and validated in future research studies. I also aim to explore any association between intervention fidelity and the smoking quit rates, in order to generate hypotheses and inform the methods for further explanatory work in a future trial.

Three aims of my thesis are:

- To develop an index to measure fidelity to a BI for smoking cessation
- To validate the fidelity index by assessing its psychometric properties
- To explore whether fidelity to a BI explains variation in quit rates

I will describe each of these aims and their specific objectives (summarised in Table 5.1) in further detail below.

5.1 Development of the fidelity index- Aim I

The first aim is to develop an index to measure fidelity to BIs for smoking cessation. I have already discussed the paucity of the published research, on which to base the fidelity index (Chapter 4). What evidence there is suggests that when used in research, as opposed to use in practice as conventional monitoring checklists, fidelity measures can be efficient and effective tools to measure intervention adherence in process evaluations, to facilitate communication of

intervention content in literature, to interpret effectiveness findings and to identify critical ingredients of the intervention (Bond et al., 2000b). In developing a fidelity index, it makes sense then, to focus on the research utilisation of the measurement tool.

I have also described (in Chapters 1, 2 and 3) how BIs for smoking cessation are complex, and the difficulties this poses for successful implementation. In developing an appropriate fidelity measure, I will define the functional and compositional features of the intervention. This is to take account of the factors affecting implementation that are intrinsic to the intervention, and to balance standardisation versus adaptation to local context. I have reviewed these concepts above, in Chapters 3 & 4, and the aim will be to integrate these concepts in conjunction with the fundamental elements of fidelity.

Further objectives in developing the fidelity index are:

- To ***describe the features of BI*** for smoking cessation, including its content and quality of interaction related aspects, to inform the 'intended' practice
- To ***map*** these intervention features to theories of behaviour change utilising the taxonomies of BCTs to understand their mechanism of action or potential effectiveness
- To ***formulate a rating scale*** for measuring fidelity to each intervention feature to inform the 'actual' practice

5.2 PSYCHOMETRIC PROPERTIES OF THE FIDELITY INDEX- AIM II

The second aim is to assess the psychometric properties of the fidelity index to assess its reliability when used in research. In determining the reliability of a measure, the consistency in its use by different people, coherence of independent intervention features within the fidelity construct, and its ability to capture variation in intervention delivery practice, are considered important.

Specific further objectives, then, are:

- To test the ***feasibility*** of use in primary research
- To determine its ***replicability*** for use by people with diverse backgrounds, to inform its stability as a research tool
- To determine the ***coherence*** of items in the index, to refine its structure and inform its reproducibility in future research
- To investigate the underlying ***dimensions of the fidelity construct*** within a behaviour change framework that the index can capture, to inform its utility in explaining variations in smoking cessation outcomes

5.3 EXPLAINING VARIATION IN QUIT RATES- AIM III

As described in Chapter 2, there are particular challenges in wider scale-up of effective BIs, due to the variability in their outcomes. These include identification of the active intervention ingredients and identification of the competences that in turn determine the training requirements of the providers of these complex interventions.

Specific objectives are, therefore:

- To describe ***intervention fidelity*** to the BI used in the ASSIST study
- To assess ***provider consistency in practice*** behaviour over time
- To explore any association between ***provider fidelity to the BI and the patient quit rates,*** to explain the observed variation
- To explore the relationship between ***Adherence to intervention content and Quality of interaction,*** to generate hypotheses for appropriate mediation or moderation pathways

The research goal is to use the outputs from this work to inform the detailed methodology of robust evaluations of BIs for smoking cessation, and to provide information on methods for identifying the active intervention features through hypothesised mediation or moderation pathways in future studies. In addition, further validation of the fidelity index in future trials of smoking cessation, based on the psychometric assessments in this thesis, will determine if fidelity predicts quit rates.

Table 5.1: Aims and Objectives of the thesis

Study Aims	Specific Objectives
I- To develop an index to measure fidelity to a BI for smoking cessation **(Study A, Chapter 6)**	**1**. To describe the features of BI for smoking cessation, including its content and quality of interaction related aspects to inform the 'intended' practice **2**. To map these intervention features to theories of behaviour change utilising the taxonomies of BCTs to understand their mechanism of action or potential effectiveness **3**. To formulate a rating scale for measuring fidelity to each intervention feature to inform the 'actual' practice
II- To validate the fidelity index by assessing its psychometric properties **(Study B, Chapter 7)**	**1**. To test the feasibility of use in primary research **2**. To determine its replicability for use by people with diverse backgrounds to inform its stability as a research tool **3**. To determine the coherence of items in the index to refine its structure and inform its reproducibility in future research **4**. To investigate the underlying dimensions of the fidelity construct within a behaviour change framework that the index can capture to inform its utility in explaining variations in smoking cessation outcomes
III- To explore whether fidelity to a BI explains variation in quit rates **(Study C, Chapter 8)**	**1**. To describe intervention fidelity to the BI used in the ASSIST study **2**. To assess provider consistency in practice behaviour over time **3**. To explore any association between provider fidelity to the BI and the patient quit rates, to explain the observed variation **4**. To explore the relationship between Adherence to intervention content and Quality of interaction, to generate hypotheses for appropriate mediation or moderation pathways

PART III- METHODS & RESULTS

This section describes and discusses the methods and the results of the three linked studies conducted to achieve the three research objectives set out in previous section, respectively.

There are three chapters in this section, each linked to one study:

Chapter 6 presents *Study A*. It provides an overview of the approaches for developing and using fidelity indices. Based on this, I then describe the methods used in the study. I end the chapter by presenting the results and discussing strengths and limitations of the study.

Chapter 7 presents *Study B*. It explains the design and methods of the primary investigation to test psychometric properties of the fidelity index and the literature used to inform these methods. I end the chapter by presenting the results and discussing strengths and limitations of the study.

Chapter 8 presents *Study C*. It explains the design and methods of the secondary investigation to explore any association between the intervention fidelity and quit rates of a BI and the literature used to inform these methods. I end the chapter by presenting the results and discussing strengths and limitations of the study.

Chapter 6. Study A: Development of the Fidelity Index

This chapter presents the first study in a series of three linked studies. It describes the first of the two studies concerning development and validation of the fidelity index. These studies were closely linked, as the assessment of psychometric properties of the fidelity index was a sequential process in the development of a reliable measure. However, I have presented the development process first (study A) and separately from the psychometric testing (study B), for clarity.

The rationale for developing and validating a quantitative measure of fidelity was to explore the variation in patient quit rates by testing its association with provider fidelity for BI. Study A builds on the research evidence established by literature presented in Chapter 1 on behaviour change science, and, on the possible sources of variation (Chapter 3) and intervention fidelity in Chapter 4.

Measurement instruments that are collections of items combined into a composite score, and intended to reveal levels of theoretical variables not readily observable by direct means, are often referred to as scales (DeVellis, 2003). When items in such measurement instruments share a common cause, they constitute a 'scale', but when they share a common consequence, they constitute an 'index' (DeVellis, 2003). As fidelity is considered an intermediary variable in an interpretive evaluation model (refer to Figure 1.2, page 29), the quantitative measure of fidelity for a BI is referred to as the 'fidelity index' throughout the thesis.

Study A (this chapter) and study B (Chapter 7) describe the methods used to develop the fidelity index and assess its reliability. Study C (Chapter 8) explores the relationship of quit rates with fidelity.

6.1 CONSIDERATIONS IN DEVELOPING A FIDELITY INDEX

Three key decisions influence the development of an index; defining the purpose of the index, developing a new index or adapting an existing one, and the level of theoretical description of intervention features.

6.1.1 Purpose of Index

The first step in developing a fidelity measure is to define its purpose. This would influence the approaches taken and methods used in its design. If the purpose is to develop an index for demonstrating intervention fidelity, then the approach will likely be more comprehensive (Borrelli, 2011), that is, identifying features of the intervention that are causally linked to theory and features that capture fidelity in a way that can be used to explain outcomes. In such cases, the researcher is more likely to consider multiple measures, to conduct detailed reliability studies, and to administer the fidelity index repeatedly (Bond et al., 2000b). Conversely, if the purpose of

the index is self-recording by the providers or for auditing services or for training in practice, where the goal is to ensure a certain basic level of adherence to an intervention programme, then a more pragmatic and less intense approach might be employed (Bond et al., 2000b). For this study the index will be developed for measuring fidelity that can be used as an intermediary variable to explain variation in smoking cessation, therefore a more comprehensive approach to its design will be adopted.

6.1.2 Existing Indices

Existing indices can be useful and inexpensive resources because these offer items validated through repeated studies (Streiner and Norman, 2008). However, a careful assessment of existing indices by searching literature on topics related to 'fidelity' and 'behavioural interventions' and speaking with experts in the field did not identify an index fitting my purpose. Some indices were developed to measure fidelity; however, they were inadequate for answering my research question (as discussed in section 4.1.3, chapter 4). These were designed either for training purposes or for assessing behaviour change strategies different from those used in the BI evaluated in my study. A brief account of the existing indices for intervention fidelity is given below.

Behavioural and psychological intervention research provides elements of fidelity in a rating format e.g. intervention content, provider behaviour and competence, protocol drift etc. (Gearing et al., 2011). It does not, however, provide the specific features within each of these elements that would allow their proper measurement. Those that provide specific features within the broader elements of fidelity are unique to psychosocial (Clarke, 1998) and psychotherapy (Weisman et al., 2002) interventions. Similarly, indices developed for educational and psychiatric practice are specific to the respective conditions under study (Dane and Schneider, 1998, Bond et al., 2000b). Other indices on behaviour change (as discussed in section 4.1.3, chapter 4) include competence rating scales i.e., The Cognitive Therapy Scale (Young and Beck, 1980) and the Motivational Interviewing Skill Code (Moyers et al., 2005), which are specific to provider competences concerning cognitive therapy for depression and motivational interviewing. The Behaviour Change Counselling Index (BECCI) (Lane et al., 2005) assesses change in provider behaviour before, during and after training, with a focus on knowledge acquisition. It is by far the closest index that matches with the behaviour change theories relating to provider competence for BIs of smoking cessation used in my study. Although the items can be adapted, the purpose of the BECCI, which is to assess knowledge acquisition of providers, differs altogether from evaluating provider fidelity for a BI.

6.1.3 Description of intervention theory

The degree to which intervention features are described theoretically determines the approach to item generation for an index, discussed below in 6.2 (Bond et al., 2000b).

Modified behaviour is an outcome of BIs (Michie and Johnston, 2012). Theory represents an integrated summary of the hypothesised causal processes involved in behaviour modification (Michie et al., 2008). The role of theory is, therefore, to identify the key concepts that are causally related to behaviour and are thus scientifically plausible targets for intervention (Michie et al., 2009c). Most BIs are based on formal theories but even those that are not, assume and hypothesise that a course of action is a potential solution to a problem (Cane et al., 2012). Hence, a careful consideration of theories and their underlying causal assumptions is important to understand the likely processes of change targeted by an intervention when developing a new index to measure fidelity for linking with the outcome.

Theories that explain behaviour and those that focus on the process of change, as detailed in Chapter 1, informed the design of BI in the ASSIST study. However, the study did not benefit from the taxonomies of BCTs, as these were not published at the time. Hence, the BI (used in ASSIST study) was theory-based, but its various features were not formally mapped to the theoretical constructs of behaviour change for smoking. Furthermore, the intervention features were not specified in a way that could be readily unpicked and rated separately.

The BCTs could help label the intervention ingredients and explain the focus of behaviour change specific to behavioural determinants of smoking; hence providing theory-linked 'compositional features' of intervention delivery.

6.2 APPROACHES TO ITEM GENERATION

Once the key conceptual decisions are made, the first step in index development is to generate the items. An item has two elements: the item stem, which is an ingredient of the intervention and the response scale, which is a set of response options that each ingredient is scored on (Bond et al., 2000b). In this thesis, the term 'item stem' is referred to as the '*item*' which represents the compositional features of the fidelity index and the '*response scale'* is called as such, representing the functional features of the fidelity index.

A fidelity index is unique from other scales because the items can be built from intervention features (Gearing et al., 2011). One of the two approaches, described below (inductive vs. confirmatory), can aid construction of items from intervention depending on its level of specification and clarity in terms of theory (Bond et al., 2000b).

6.2.1 Inductive approach

If an intervention is not well-defined theoretically and its key features cannot be disentangled based on underlying causal mechanisms, then an inductive approach is considered best for generating items (Bond et al., 2000b). In an inductive approach, the key intervention features are discovered, using interviews and brainstorming with people, as well as observations, and content analyses of documents. This approach requires developing a coherent understanding of the

theoretical underpinning of the intervention for identification and definition of its important features and the generation of new items for the index. As the inductive approach is used when interventions are not well-based in theory, larger sample sizes, often more than 20 individuals might be required (Bond et al., 2000b).

The Delphi technique, concept mapping (through focus groups), ethnography (through structured interviews), critical incidents technique, and content analysis are the different types of methods used to aid the inductive approach for generating items. Many of these methods are also used in the confirmatory approach, as described below.

6.2.2 Confirmatory approach

Confirmatory approach is used to specify intervention features that are already deeply rooted in theory (Bond et al., 2000b). This approach relies on well-established intervention theory and intervention resources such as the practice manuals that provide details on key intervention features. Confirmatory approach comprises of two types of techniques, the critical components technique and the expert consensus, described here:

The critical components technique

The critical components technique is used to generate items *a priori* (Bond et al., 2000b). This technique focuses on the activities of an intervention, rather than the broader underlying theoretical concepts. It is quite challenging as it requires sifting through the intervention features and identifying those that are measureable and representative of the construct (e.g. behaviour change) under scrutiny. Furthermore, it involves research to ensure that the selected intervention features have been shown empirically to be linked to theory (Streiner and Norman, 2008).

Expert consensus

Often used in conjunction with the above technique is the process of inviting expert insight into the generated list of items and establishing the expert consensus on the content (Bond et al., 2000b). The experts usually comprise intervention users and other individuals who are knowledgeable about the subject matter. This technique is discussed further in section 6.5, page 82, under '*content validation*'.

The selection of best approach (inductive or confirmatory) for constructing items of the fidelity index relies on the BI used in the ASSIST study and its specification using the theories of behaviour change.

6.3 DESCRIBING BEHAVIOURAL INTERVENTION USED IN THE ASSIST STUDY

Here I outline the BI used in the ASSIST study and how its key features can be prepared for use in the fidelity index.

The BI (briefly described in Appendix B) was developed in 2008. It used the World Health Organisation's (WHO) '5As to quit' model for structuring the counselling sessions (World Health Organisation, 2014c). Activities within 5As were designed, focusing on behaviour change theories, as described above in section 1.2.1 (Green and Murphy, 2014, Fishbein, 1979, Ajzen, 1991, Prochaska and DiClemente, 1984) to assist smokers with their quit attempt. These activities can be easily divided into 'compositional' and 'functional' features as discussed in Chapter 4, to describe the *content* and *delivery* of the intervention (Figure 6.1).

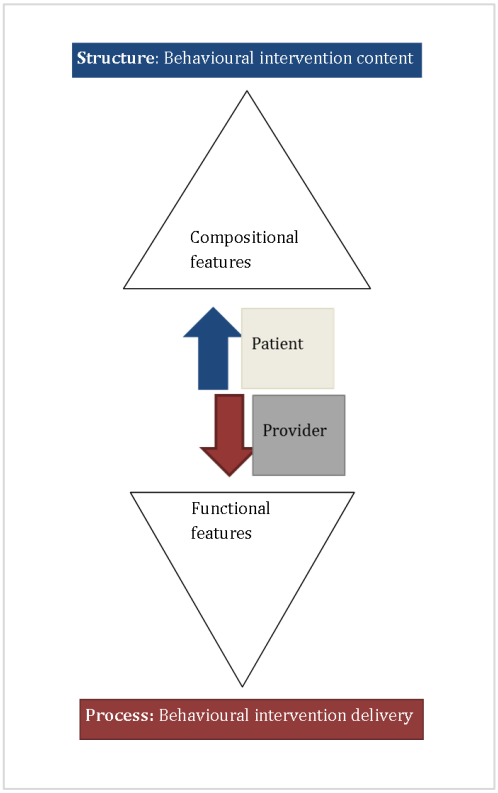


Figure 6.1: Behavioural intervention content and delivery

In the wider behaviour change literature 'quality of intervention delivery' equates with the notion of intervention provider *competence* (Lorencatto et al., 2016). Using this concept the term 'quality' has often been used to 'qualify' intervention ingredients on 'how well' these are delivered in practice. However, in my work, keeping true to the original idea of the 'quality of intervention delivery' as distinct from the 'process of delivery' (discussed below), I used the evidence-based competences (Michie et al., 2011b) required to deliver behavioural support to describe the 'inter-personal aspects' of the BI and called it 'Quality of interaction'. For the 'content-based aspects' of the BI, I used the BCTs developed for individual behaviour change for smoking cessation (Michie et al., 2011c) and called it 'Adherence to BI content'. Both of these terms in my study represent the 'compositional features' of the index or the 'intended practice'. To 'qualify' each ingredient of the BI on 'how well' it is delivered, which I perceived as the 'process of delivery', a scoring criterion or scale is needed for each individual ingredient. This qualification of the ingredients is the essence of a quantitative measure for a complex intervention and forms the 'functional features' of the index and captures the 'actual practice'.

6.3.1 Behavioural Intervention 'Content'

The 5As quit model (Table 6.1) summarises the activities that a health care provider could do to help a smoker quit tobacco use (World Health Organisation, 2014c). This model guides the

provider through the process of talking to patients who are ready to quit smoking and delivering advice. The key features of the BI (Table 6.1) within the 5As quit model, in the ASSIST study, were built to facilitate patient-provider interaction in clinical settings, using an assessment questionnaire and pictorial messages (in the form of a flip-chart).

Table 6.1: Behavioural Intervention used in ASSIST study

5As quit Model	**Key features of Behavioural Intervention**
Ask (about tobacco use behaviour and history)	1. Ascertain about tobacco use
Advise (about consequences of tobacco use and quitting)	2. Information about harms of smoking and benefits of quitting
Assess (willingness to quit)	3. Assessments of dependency and motivation to stop 4. Preparation for quit attempt
Assist (in quitting)	5. Identifying the social/psychological or environmental cues that trigger a smoker's desire to smoke and advising on ways to monitor and avoid these 6. Assessing the withdrawal symptoms (i.e. the strong urges/cravings to smoke) and advising on ways to address and overcome these
Arrange (follow-up)	7. Offering the BI leaflet as a reminder

6.3.2 Behavioural Intervention 'Delivery'

The BI flip-chart and the integrated questionnaire were designed for use in a specific way. Each slide in the flip-chart had a sketched scenario facing the patient and key messages on the back facing the provider to help them deliver the messages effectively. The *interaction style* of the provider described the way that the intervention is intended to be delivered, that is, delivery using three essential steps:

Step1- asking the patient to describe the slide

Step2- facilitating them with understanding the message

Step3- clarifying concept and re-emphasizing the key message

The interaction style was more complex than these 3 steps of delivery because the BI also involved a structured questionnaire interview, which would require defining the process of delivery of the assessment questions, in addition to the flip-chart slides.

The processes involved in delivering BI ingredients form the 'functional features' of the index which would capture intervention fidelity. The integrity of the index in measuring fidelity relies heavily on defining these functional features based on provider interaction-style. This would give the fidelity index its quantitative ability to measure 'actual practice' assessed on the intended practice.

6.4 APPROACHES TO SCORING OF ITEMS IN THE FIDELITY INDEX

The goal for the fidelity index is to develop items and their corresponding response scales that will accurately capture the provider fidelity to intervention delivery. Each ingredient of the BI

identified and described using the taxonomies of behaviour change will qualify as an '*item*' of the index and thus represent the 'intended practice' (as described in Chapter 4).

After creating these items, the next step is to establish the '*response scale*' for rating and scoring the items to capture 'actual practice'. Developing the response scale would require careful consideration for this fidelity index, as it encompasses two levels of behaviour (Figure 6.2). The first level is the individual smoker behaviour (compositional features) that is targeted by the intervention to aid quitting. The second level is the provider behaviour in delivering the intervention or the interaction style (functional features). Whilst the first level (intended behaviour change of the target individual) can be captured within the structural formation of the index represented by items, the second level (provider practice or interaction style) needs anchoring within the scale points, by creating response options that accurately capture the 'actual practice' (as described in Chapter 4).

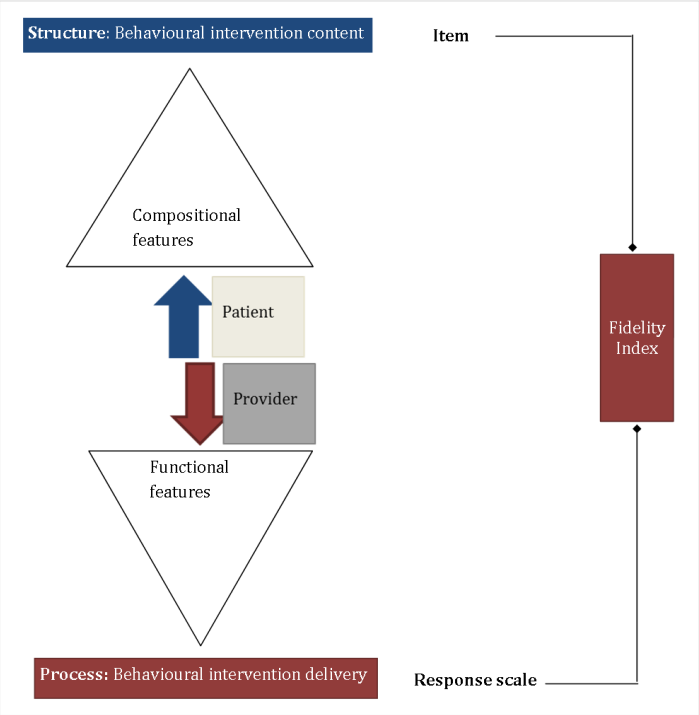


Figure 6.2: Developing items and their corresponding response scales

6.4.1 Establishing response scale points

This step, crucial in the index development, relates to the precision of measurement of fidelity (or 'actual practice') for a BI.

Linking fidelity to outcomes requires indices to have the ability to detect variation in delivery of the intervention (Nelson et al., 2012). Indices involving rating of items on a binary scale (e.g. implemented or not implemented) cannot detect sufficient variance in implementation to link these to outcomes (Nelson et al., 2012). Therefore, three, five or seven point ordinal scales are often used, to ensure that all relevant variability in intervention delivery practices for each item could be captured.

The following characteristics are important in devising a comprehensive yet precise response scale for use in research (Bond et al., 2000b):

- A standard number of scale points for every item
- Ordinal scale points approximating equal intervals between each option
- Response scale points for each item. These can be numerical representing gradients or behaviourally anchored linking to specific behaviours or a set of behaviours. If the behavioural anchors do not include all of the possible response options, or if only part of the range of options captures all of the possible responses, then the scale is not sensitive. The response scale is more reliable for use by different coders if all response options have the potential of being chosen (Mowbray et al., 2003).
- No gaps in the response alternatives (or options), and
- No overlap in the response options

Several methods have been introduced for coding intervention fidelity data (Borrelli, 2011). The simplest method is to rate the occurrence or non-occurrence of intervention ingredients. A coder simply checks off the prescribed and proscribed ingredients that occur, while listening to an audio/video-tape or observing an intervention session. In most of the fidelity indices for psychological and behavioural interventions, each item was rated on a three point ordinal scale. For example, Gearing et al. (2011), in the comprehensive intervention fidelity guide, used: absent/minimal (score= 0), moderate (score= 1), and extensive (score= 2); and Clarke et al. (1998), in the fidelity scale for psychosocial therapy of depression, used: no adherence (score= 0), partial adherence (score= 1) and complete adherence (score= 2).

Whilst the high and low ends of scale response options are often the easiest to develop, the middle options, that need to be independent yet meaningful, pose a challenge to define. (Bond et al., 2000b). Because few interventions are fully implemented as intended, or rarely, not adhering at all, it is often the middle alternatives on the response scale that might be crucial for capturing the variability in actual practice and thus the selectivity of the items (Bond et al., 2000b). Coders need to be able to differentiate between the response alternatives in the scale for a respective item stem, because, if the alternatives are too finely differentiated, then the coders may have a difficult time choosing one of them. On the other hand, if the response alternatives do not discriminate finely enough, coders might end up choosing the same option, in spite of their responses differing widely on the dimension (e.g. behaviour change) being measured (Bond et al., 2000b).

An accepted way of objectively measuring intervention delivery processes in health services research and psychometrics is to utilise linear additive scales (Nunnally and Bernstein, 1994). These consist of scoring of the item stems based on an ordered response (ordinal categories) scale. The scores for each item are then summed up to give a total or 'composite' score for the entire index.

6.5 CHECKING ADEQUACY OF GENERATED ITEMS

Once the fidelity index is formed, the next few steps relate to its field measurement, validity and reliability of use; all of which delineate its validation as a sound measure. The terms 'face validity' and 'content validity' are the technical descriptions of the judgement that the index looks reasonable and measures the intended concepts (Streiner and Norman, 2008).

6.5.1 Content validation

Once the items are listed, this step ensures that the index adequately covers the construct being measured (Streiner and Norman, 2008). The technical term for this is *'content validation'* (described in section 6.5), although many theorists believe it should be replaced by more accurate descriptors such as, 'content relevance' or 'content coverage' (Messick, 1980).

Content rating by experts on relevance and importance to context

As noted in 6.2.2 above, 'expert consensus' has been used for a wide variety of purposes in health research, such as the development of practice guidelines for disease prevention and management (Bond et al., 2000b). While there are important differences between practice guidelines and fidelity indices, they are both developed for a common purpose: to inform provider practice (Bond et al., 2000b). Where guidelines are developed to standardise practice, fidelity indices are used to measure practice attributes.

Building consensus depends on several factors: the selection of participants, selection and presentation of scientific information, the way in which the interaction is structured, and the method of synthesising individual judgements (Black et al., 1999). The principle qualitative research methods (Peters et al., 2013) for content validation include: the Delphi technique, nominal group technique and consensus development conference (Black et al., 1999).

The 'consensus development conference' is an informal method that involves open meetings amongst experts, followed by private group discussions to reach agreement (Black et al., 1999). Sometimes recommendations are produced in a single meeting (Bond et al., 2000b). The 'Delphi technique' and the 'nominal-group technique' are comparatively formal approaches to consensus development, involving face-to-face meetings or mailings to a carefully chosen group of experts (Black et al., 1999).

The advantage of using the Delphi technique over other consensus methods is that if experts are chosen carefully, they can be assumed to represent the most recent thinking in the area, giving the index developer access to the accumulated knowledge and experience. However, the method can prove to be disadvantageous if it does not reflect a range of opinions covering all aspects of the subject under study (Streiner and Norman, 2008). The Delphi technique can be administered in a group i.e. face-to-face, or during an individual consultation (aka *modified Delphi technique*)

(Black et al., 1999). Two to three rounds are likely to result in some convergence of individual judgements, while more than three rounds show little impact on agreement and might adversely affect the response rate (Black et al., 1999).

Criteria for item selection based on content rating

Expert consensus is usually obtained on the 'importance' i.e. the ingredients deemed essential for the index and the 'intervention specification' i.e. the clarity in describing the key intervention features as items (Peters et al., 2013). For indices where well-established specification methods (such as the taxonomy of BCTs) are used, expert consensus might only be required on 'importance' of the items to be retained.

To identify the items to retain in the index, it is also necessary to establish criteria for item inclusion based on the ratings given by the experts on relevance and importance of the items. There are no established rules for setting this criterion. Researchers have arbitrarily used the criterion that at least 50% of the experts rank an item as "very important" or "very relevant" (McGrew and Bond, 1995).

6.5.2 Content pilot

Content piloting is often undertaken to determine the *'face validity'* of a tool. It focuses on respondent perceptions or other stake holders beliefs about the content of the index, rather than the empirical evidence by experts (Streiner and Norman, 2008). This ensures that the items on the surface tend to measure what they were designed to measure, that is, the items and their response scale retain their meaning (as intended by the researcher) when they are administered.

Piloting the content of a new measure is critical in identifying problems with the sequence of the items, ambiguity in items, confusion in response alternatives, and other issues in its administration. For instance, if data are collected through an interview, a pre-test pilot gives an opportunity to the interviewer to practice asking the questions, as well as coding the responses. The goal of a content pilot is vast and often covers aspects from content clarity to unfolding field test errors; these include (Bond et al., 2000b):

- Determining feasibility of data collection methods,
- Giving observers/coders practice,
- Identifying problems with the pace or placement of the items,
- Identifying terminology or jargon problems,
- Identifying whether the response scale options are appropriate, and
- Assessing whether the provider has additional information that would be vital to the intervention fidelity, and, is not currently being covered by the index.

All of the above inform decisions in refining the content of the index and measurement procedures via identification of the problematic content (e.g. items with multiple meanings, double barrelled questions or misunderstood meaning of the items) and process issues (e.g. logistical or administrative challenges in the field etc.).

6.6 AIMS OF STUDY

> "To develop an index to measure fidelity to a BI for smoking cessation" (described above in 5.1).

The specific objectives of study A are to:

- ***Describe the features of BI*** for smoking cessation, including its content and quality of interaction related aspects, to inform the 'intended' practice
- ***Map*** these intervention features to theories of behaviour change utilising the taxonomies of BCTs to understand their mechanism of action or potential effectiveness
- ***Formulate a rating scale*** for measuring fidelity to each intervention feature to inform the 'actual' practice

6.7 METHODS

The development of the fidelity index consisted of several sequential steps as depicted in Figure 6.3, which illustrates the 'structure' (in blue) and 'processes' (in red) of intervention delivery.

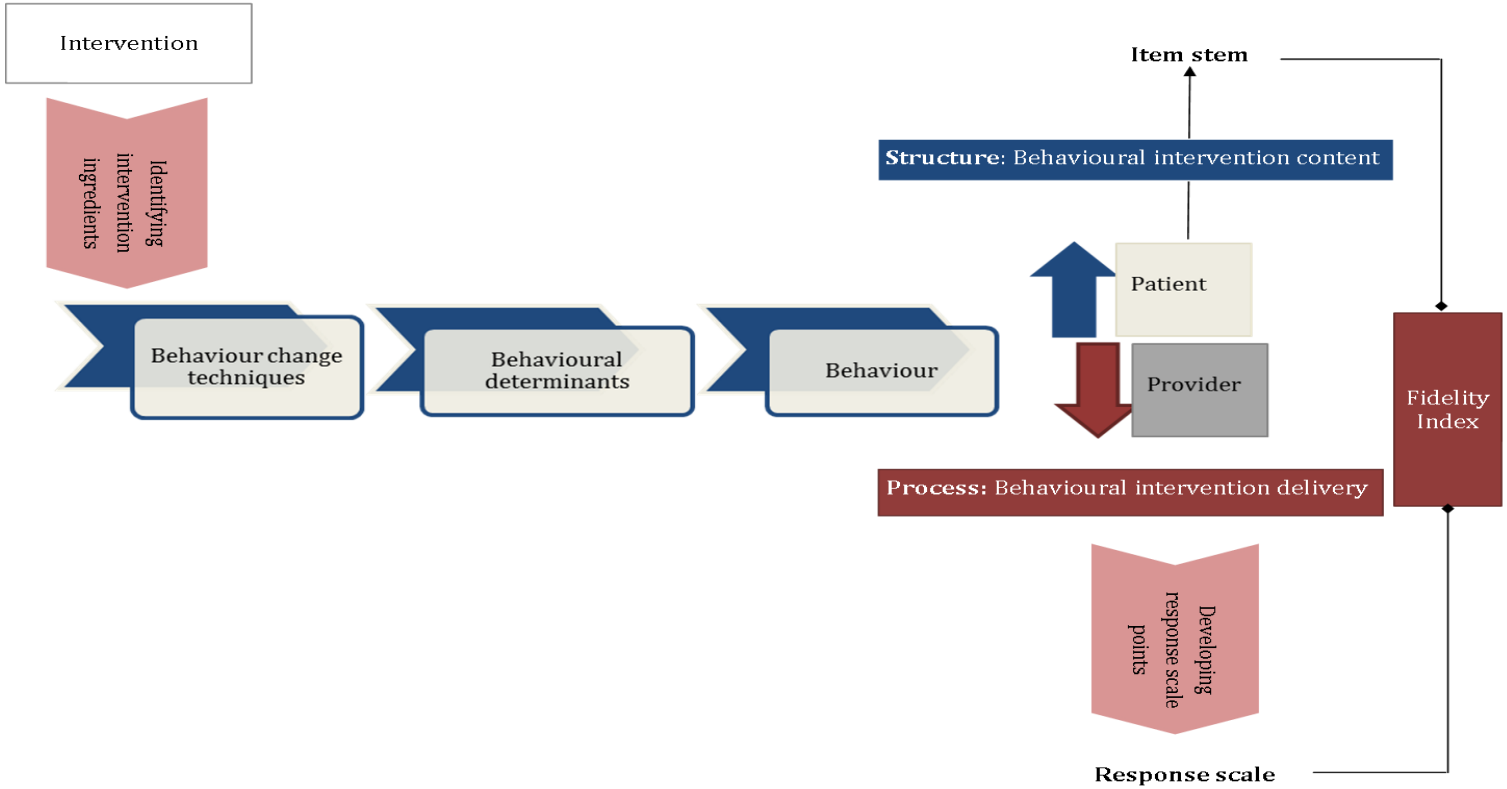


Figure 6.3: Design of the fidelity index

6.7.1 Generating items for the fidelity index

The fidelity index was based on the BI used in the ASSIST study- well-defined in theory; therefore a confirmatory approach was considered appropriate for item generation. Within this approach, the critical components technique was used to sift through the BI and pick intervention features. The 5As to quit model (World Health Organization, 2008) for smoking cessation provided the skeleton to the index, within which the key features of the BI (Table 6.1) were unpicked and specified using the taxonomies of BCTs (see next heading).

6.7.2 Mapping items to theory

As noted in Figure 6.3 (above), the 'logic model' of behaviour change (Figure 1.1, page 20) described in chapter 1, was used to map the BI ingredients in terms of the behavioural determinants and the BCTs likely to be influential.

The taxonomies of behaviour change (Michie et al., 2011c) and competences (Michie et al., 2011b) to deliver BIs for smoking cessation were used to specify BI ingredients. Features that were core to individual behaviour change (e.g. boosting patient motivation, maximising their self-regulatory capacity etc.) were labelled to represent BI content for *Adherence* (Michie et al., 2011c). Features of provider competence (e.g. focus on emphasising patient choice within the bounds of evidence-based practice, focus on providing reassurance regarding treatment outcomes and tailoring interaction according to the patient needs) were labelled to delineate the *Quality* of interaction (Michie et al., 2011b).

Hence, a list of items was assembled for '*Adherence*' to BI content and '*Quality'* of interaction forming the 'compositional' features of the fidelity index (Figure 6.4). These mapped features of the BI known as 'items' represented 'intended practice' and gave the index its necessary structure.

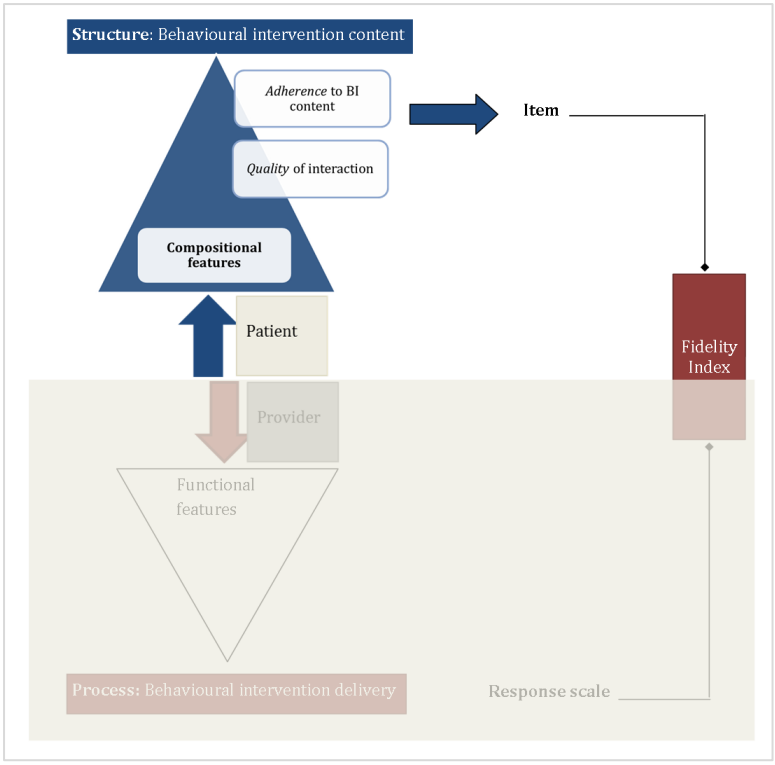


Figure 6.4: Compositional features of the index forming items

6.7.3 Establishing a response scale for the items

Similar to existing fidelity indices (Gearing et al., 2011, Clarke, 1998), a three point ordinal scale was chosen for each item; the scores against each item would then be summed up for all items in the index, to give a 'composite index score' (Nunnally and Bernstein, 1994), which makes it a linear additive scale. The psychometric properties of the linear additive index, both for *Adherence* to BI content and *Quality* of interaction, are presented in next chapter.

The most important aspect, in establishing response scale points for the items was defining the anchors that help distinguish between the three point scale alternatives.

Defining selection anchors

As fidelity represented provider practice or process of delivery of the BI, it made sense to capture this behaviour using behavioural anchors. For example, if a response scale for an item 'providing reassurance' was 0= not implemented, 1=partially implemented and 2=fully implemented, the behaviourally anchored response options were defined as: 0= did not provide any reassurance, 1= listens and answers in yes and no but does not provide any constructive advice, and 2= gives constructive advice. For the flip-chart used in BI, this was done using the 3 steps (presented in Behavioural Intervention 'Delivery', page 79) for interaction style. For the patient assessment questions these anchors were defined based on the way the questions were asked.

Efforts were made to define scale response options that were exhaustive and mutually exclusive, to avoid any gaps or overlaps between alternative options. Similarly, middle alternatives of the response options were kept at equal intervals between the low and high end of the scale, to truly represent its ordinal value for rating items.

Hence, each item was assigned a three point ordinal scale, with behaviourally anchored response options based on provider interaction style forming the 'functional' features of the fidelity index (Figure 6.5). The 'response scale' represented the 'actual practice'.

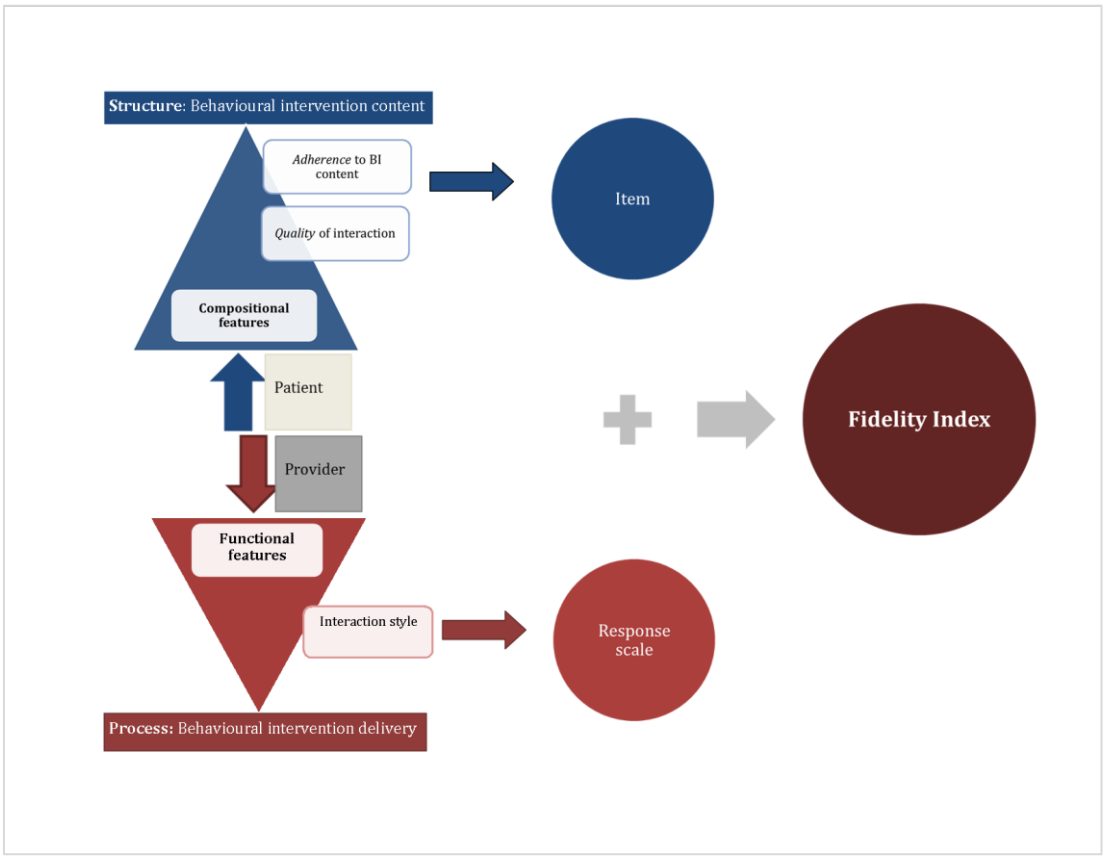


Figure 6.5: Functional features of the index forming response scales for corresponding items

6.7.4 Content validation

A modified Delphi procedure was carried out to validate the content of the fidelity index. The list of items with response scales was submitted via email (Appendix C.1) to three experts: a behavioural scientist, a smoking cessation provider and a smoking cessation expert.

Content rating for each item was requested on importance using a Likert scale (1- very important to 4- not at all important). A maximum of two rounds were to be executed to conclude individual judgements and reach consensus, given the number of experts was small. In view of similar studies (McGrew and Bond, 1995), consensus on dropping an item from the list was considered if two of the three experts rated an item as "not at all important" or "not very important".

6.7.5 Content pilot

Field piloting of the index for 'face validity' was carried out in Islamabad in a tertiary care hospital housing a TB clinic, in December 2013. Field piloting involved observation of two patient-provider BI sessions, during which the coders rated each item on the fidelity index (for details on coders see Chapter 7, 'Selection and training of coders', page 115).

6.8 RESULTS

6.8.1 Items generated for the fidelity index

Thirty-seven BI ingredients were identified from the ASSIST intervention resource (including flip-chart messages, the patient assessment questionnaire and the training manual), using the critical components technique. The list of these ingredients of BI is presented in Table 6.2.

Table 6.2: List of ingredients or (items) identified from Behavioural Intervention

5As quit Model	Ingredients of Behavioural Intervention	Item #
Ask (about tobacco use behaviour and history)	Assessing current and past smoking behaviour i. Pattern of smoking behaviour (Types of smoking? Smokers in vicinity? Children at home?) ii. Age when started smoking iii. Amount smoked	1
Advise (about consequences of tobacco use and quitting)	Awareness about the various forms of tobacco smoked in the community	2
	High blood pressure and heart disease	3
	Lung diseases like chronic cough, asthma, TB and cancer	4
	Wastage of money, staining of teeth, gum problems and bad breath	5
	Effects on children's health: pneumonia, asthma etc.	6
	Effects on pregnancy: complications in pregnancy, low birth weight baby	7
	Decide to quit, choose a quit date and utilize the money on better things	8
	Social and economic benefits of quitting	9
Assess (Willingness to quit)	Current level of motivation to stop/willingness to quit	10
	Reasons for quitting e.g. health, cost, example for others, family's health or other reason	11
	Quit from today? If 'No', Quit within next five days? If 'No', when will you be able to set a quit date?	12
	Attempted quit in the past, Number and duration of past quit attempts, Time since last quit attempt	13
	Factors that led back to smoking including social and physical factors (Social reasons: Family problems (tension), Smokers company, any other; Physical Symptoms: Craving, Indigestion, Insomnia, Headache, any other)	14

Assist (in quitting)	Fagerstrom Test for Nicotine Dependence	15
	Setting quit date	16
	Hide reminders of smoking/ways of changing the physical environment to minimise exposure to smoking cues	17
	Declare the house as 'smoke free home'	18
	Identify individual who can help support in quitting at home	19
	Decide on telling people about stopping or keeping it private	20
	Triggers and their management	
	Trigger1: immediately after rising in the morning- Offer strategies to manage trigger 1	21
	Trigger2: defecation- Offer strategies to manage trigger 2	22
	Trigger3: eating meals- Offer strategies to manage trigger 3	23
	Trigger4: free at home or feeling bored- Offer strategies to manage trigger 4	24
	Trigger5: seeing others smoke- Offer strategies to manage trigger 5	25
	Trigger6: offered smoking by others- Offer strategies to manage trigger 6	26
	Trigger7: intense physical/mental work- Offer strategies to manage trigger 7	27
	Trigger8: tense/anxious- Offer strategies to manage trigger 8	28
	Withdrawals and their management	
	Withdrawal1: Craving smoking- Offer strategies to manage withdrawal1	29
	Withdrawal2: Restlessness/anger- Offer strategies to manage withdrawal2	30
	Withdrawal3: Headache- Offer strategies to manage Withdrawal3	31
	Withdrawal4: Insomnia- Offer strategies to manage withdrawal4	32
	Withdrawal5: indigestion- Offer strategies to manage withdrawal5	33
	Withdrawal6: Anorexia and constipation- Offer strategies to manage withdrawal6	34
	Withdrawal7: Cough- Offer strategies to manage withdrawal7	35
	Withdrawal8: Weight gain- Offer strategies to manage withdrawal8	36
Arrange (Follow-up)	Give BI information leaflet to the patient and brief about it	37

6.8.2 Mapping items to theory

The 37 items identified in the previous step represent content of the BI, and another eight represent provider competence to effectively deliver BIs, making a total of 45 items representing 'compositional features' of the index and the 'intended' practice for delivering BI.

Both types of items will be scored for fidelity; content of the BI for *Adherence* and competence to deliver the BI for *Quality*.

These 45 items were specified using 23 BCTs (Michie et al., 2011b, Michie et al., 2011c), which were further grouped in terms of the behavioural determinants (Table 6.3).

Table 6.3: Items mapped to Behaviour Change Techniques and Behavioural Determinants

Behavioural Determinants	Behaviour Change Techniques (BCTs)	Items
Adherence to content		
General aspects of the interaction (R) focusing on information gathering (I)	*RC7: Information gathering and assessment	1. Assessing current and past smoking behaviour i. Pattern of smoking behaviour (Types of smoking? Smokers in vicinity? Children at home?) ii. Age when started smoking iii. Amount smoked
Specific focus on behaviour (B) and addressing motivation (M)	BM1: Provide information on consequences of smoking and smoking cessation	2. Awareness about the various forms of tobacco smoked in the community 3. High blood pressure and heart disease 4. Lung diseases like chronic cough, asthma, TB and cancer 5. Wastage of money, staining of teeth, gum problems and bad breath 6. Effects on children's health: pneumonia, asthma etc. 7. Effects on pregnancy: complications in pregnancy, low birth weight baby 8. Decide to quit, choose a quit date and utilize the money on better things 9. Social and economic benefits of quitting
General aspects of the interaction (R) focusing on information gathering (I)	RC7: Information gathering and assessment	10. Current level of motivation to stop/willingness to quit
		11. Reasons for quitting e.g. health, cost, example for others, family's health or other reason
Specific focus on behaviour (B) and addressing	BM6: prompt commitment from the client there and then	12. Quit from today? If 'No', Quit within next five days? If 'No', when will you be able to set a quit date?

motivation (M)		
General aspects of the interaction (R) focusing on information gathering (I)	RC7: Information gathering and assessment	13. Attempted quit in the past, Number and duration of past quit attempts, Time since last quit attempt
		14. Factors that led back to smoking including social and physical factors (Social reasons: Family problems (tension), Smokers company, any other; Physical Symptoms: Craving, Indigestion, Insomnia, Headache, any other)
General aspects of the interaction (R) focusing on information gathering (I)	RC7: Information gathering and assessment	15. Fagerstrom Test for Nicotine Dependence
Specific focus on behaviour (B) and maximising self-regulatory capacity/skills (S)	BS4: Facilitate goal setting	16. Setting quit date
Specific focus on behaviour (B) and maximising self-regulatory capacity/skills (S)	BS8: Advise on environmental restructuring	17. Hide reminders of smoking/ways of changing the physical environment to minimise exposure to smoking cues
		18. Declare the house as 'smoke free home'
Promote adjuvant activities (A)	A2: Advise on use of social support	19. Identify individual who can help support in quitting at home
Specific focus on behaviour (B) and addressing motivation (M)	BM2: Boost motivation and self-efficacy	20. Decide on telling people about stopping or keeping it private
Specific focus on behaviour (B) and maximising self-regulatory capacity/skills (S)	BS1: Facilitate barrier identification and problem solving BS3: Facilitate action planning /develop treatment plan	21. Trigger1: immediately after rising in the morning- Offer strategies to manage trigger 1
		22. Trigger2: defecation- Offer strategies to manage trigger 2
		23. Trigger3: eating meals- Offer strategies to manage trigger 3
		24. Trigger4: free at home or feeling bored- Offer strategies to manage trigger 4
	BS1 & 3 ; BS11: Advise on avoiding cues for smoking BS12: Facilitate restructuring of social life	25. Trigger5: seeing others smoke- Offer strategies to manage trigger 5
		26. Trigger6: offered smoking by others- Offer strategies to manage trigger 6
	BS1 & 3;	27. Trigger7: intense physical/mental work-

	BS10: Advise on conserving mental resources	Offer strategies to manage trigger 7
	BS1, 3 & 10; BS14: Teach relaxation techniques	28. Trigger8: tense/anxious- Offer strategies to manage trigger 8
General aspects of the interaction (R) focusing on general communication(C)	RC10: Provide information on withdrawal symptoms	29. Withdrawal1: Craving smoking- Offer strategies to manage withdrawal1
		30. Withdrawal2: Restlessness/anger- Offer strategies to manage withdrawal2
		31. Withdrawal3: Headache- Offer strategies to manage Withdrawal3
		32. Withdrawal4: Insomnia- Offer strategies to manage withdrawal4
		33. Withdrawal5: indigestion- Offer strategies to manage withdrawal5
		34. Withdrawal6: Anorexia and constipation- Offer strategies to manage withdrawal6
		35. Withdrawal7: Cough- Offer strategies to manage withdrawal7
General aspects of the interaction (R) focusing on general communication(C)& Specific focus on behaviour (B) and maximising self-regulatory capacity/skills (S)	RC10; BS13: Advise on methods of weight control	36. Withdrawal8: Weight gain- Offer strategies to manage withdrawal8
General aspects of the interaction (R) focusing on general communication(C)	RC6: Offer appropriate written material	37. Give BI information leaflet to the patient and brief about it
Quality of interaction		
General aspects of the interaction (R) focusing on general communication(C)	RC1: Build general rapport	38. Build general rapport: Establish a positive, friendly and professional relationship with the smoker and foster a sense that the smoker's experiences are understood
	RC9: Explain expectations regarding treatment programme	39. Explain expectations regarding treatment programme: Explain to the smoker the treatment program, what it involves, the active ingredients, and what it requires of the smoker
	RC4: Provide reassurance	40. Provide reassurance: Give general reassurance to the smoker that his/her experiences are normal and time limited, and provide positive expectations of success based on experience with other

		smokers in the same situation
	RC2: General practitioner communication approaches	41. General practitioner communication approaches: (Elicit and answer questions)
		42. General practitioner communication approaches: (Use reflective listening)
		43. General practitioner communication approaches: (Summarising information and confirming client decisions)
General aspects of the interaction (R) focusing on the delivery of the intervention (D)	RC5: Tailor interactions appropriately	44. Tailor interactions appropriately: Use relevant information from the client to tailor the behavioural support provided/ flexible adaptation that takes into account individual patient needs
	RC3: Emphasise choice	45. Emphasise choice: Emphasise client choice within the bounds of evidence based practice
* BCTs labels used in the taxonomies. For further details refer to (Michie et al., 2011b, Michie et al., 2011c)		

6.8.3 Establishing response scale for the items

The response scale points, representing the 'functional features' of the fidelity index, and their behavioural anchors based on the interaction style of the provider, are described in Table 6.4.

Behavioural anchors were defined separately for the flip-chart and patient assessment questionnaire. For the flip-chart, the anchors were defined in accordance with the 3-steps of interaction style. Within the patient questionnaire, anchors for assessment questions were defined, based on how the question was asked and whether full response was elicited using the probes provided. The Fagerström scale used for assessing nicotine dependence of the smoker was administered as part of the patient questionnaire; for this scale the anchors were defined based on how the scale was administered and if all responses within the scale were elicited (see Table 6.4).

Eventually, the developed fidelity index comprised of two sub-indices, '*Adherence*' having 37 items with a maximum score possible of 74 and '*Quality*' having eight items with a maximum score possible of 16, both measuring two important aspects of fidelity.

Table 6.4: Behavioural anchors for the response scale

Interaction style	Response scale	Behavioural Anchoring
Flip-chart (for all slides) Step1- Asking the patient to describe	0 = not implemented	Consider not implemented if the provider '*skips all 3 steps*'

the slide Step2- Facilitating them with understanding the message Step3- Clarifying/re-emphasizing the key message	1= partially implemented	Consider partially implemented if the provider '*delivers 2 out of 3 steps*'
	2= fully implemented	Consider fully implemented if the provider '*delivers all 3 steps*'
Assessment questions (in patient questionnaire)	0 = not implemented	Consider not implemented if the provider '*skips asking the question'*
	1= partially implemented	Consider partially implemented if the provider '*asks about the assessment question without eliciting a response using the categories given for that question*'
	2= fully implemented	Consider fully implemented if the provider '*asks about the assessment question and elicits response using the categories given for that question*'
Nicotine dependence assessment scale(Fagerström) (consists of 6 questions; in patient questionnaire)	0 = not implemented	Consider not implemented if the provider '*skips the scale'*
	1= partially implemented	Consider partially implemented if the provider '*asks less than 6 questions on the scale'*
	2= fully implemented	Consider fully implemented if the provider '*asks all 6 questions*'

6.8.4 Content Validation

Out of the three experts, at least two rated all items as 'very important' or 'somewhat important' on the rating scale (see Appendix C.2). None of the experts rated any item as 'not at all important'. Therefore, none of the items were excluded from the fidelity index. The Delphi procedure was, therefore, restricted to a single round, due to convergence among experts.

6.8.5 Content pilot

The piloting of the index revealed that some of the behavioural anchors used for defining the response options were not mutually exclusive. They required a clearer discrimination between 'not implemented' and 'partially implemented'. I also found that the ordering of items was not sequential for the coders to follow easily while rating audiotapes. Therefore, I made the following changes to the fidelity index in response to the content pilot:

i. For "attempted quit in the past" (item 13); if the patient did not attempt, it was decided to code it as 2 (i.e. fully implemented); similarly for "factors that led back to smoking' (item 14) a score of 2 was to be assigned, in case of no attempted quit in the past, because there was not any "not applicable" category in the response scale.

ii. For triggers and their management (items 21 to 28); again if the patient reported not having a trigger and the provider skipped the management for that specific trigger, it was decided to give that item a score of 2 (fully implemented).

iii. In the "information gathering and assessment" (item 1), there were three main questions (1. Pattern of smoking behaviour: types of smoking, smokers in vicinity, children at home; 2. Age when started smoking; and 3. Amount smoked). It was decided that in instances where none of the three main questions were asked, a 0 score would be given (i.e. not implemented). Previously this was not factored in the response options for this item that a provider might skip these questions altogether.

iv. A part of this question (item 1) required verbalising or probing for different types of tobacco used by the patient. In an instance when the provider did not ask about the forms of tobacco use to elicit proper patient response (or rushed through the question), then it was decided to code these as 1 (i.e. partially implemented).

v. Item 7 ("effects on pregnancy") was not applicable in males; therefore the clarification was added to mention this in comments if this slide of the flip-chart was skipped by the provider.

vi. Initially, the fidelity index was created in a Word document (Microsoft office 2010). Pilot coding revealed that it was not only difficult to sum all scores obtained on the response scales for each item in order to record composite index scores, but the method was also prone to errors in calculation. Therefore, it was decided to convert the fidelity index to an Excel spreadsheet (Microsoft office 2010) with built-in formulas for summing composite scores, both for *Adherence* and for *Quality*.

vii. The ordering of the items in the index was changed to match the sequence in which the providers used the assessment questions and the flip-chart slides.

The fidelity index that was developed and refined after content validation and field piloting is given in the Appendix C.3.

6.9 DISCUSSION

6.9.1 Summary of findings

Extensive search of the literature for fidelity measurement of BIs showed that it was not possible to adapt pre-existing indices. These were designed either for training purposes or for assessing behaviour change strategies that differed from the ones being evaluated. Therefore, a new fidelity index was developed based on the BI from the ASSIST study.

Critical components technique was used to select key intervention features and theoretically informed taxonomies were used to map these features. The mapped ingredients formed the core items of the index and thus represented the 'intended practice'. Some of the items related to BI content representing *Adherence* (37 items), and others to the *Quality* of interaction (8 items).

These formed the compositional features of the fidelity index that could be adapted to patient needs, as discussed in Chapter 4.

Each item in the index was designed for rating on a 3 point ordinal scale as fully, partially, or not, implemented. The response scale was behaviourally anchored based on provider interaction style and thus represented the 'actual practice'. These form the functional features of the fidelity index that might be used for standardising the process of delivery of BIs, as discussed in Chapter 4.

Modified Delphi procedure indicated a strong agreement on item inclusion in the index by experts; therefore none of the items were dropped. Content pilot by coders after collecting data on patient-provider BI sessions resulted in re-phrasing and re-defining some of the behavioural anchors, leading to a refined fidelity index (Appendix C.3).

6.9.2 Structure of the index

The way in which individual items and the two sub-indices were designed to capture fidelity is discussed in this sub-section.

Two aspects of fidelity: the adherence to content and the quality of interaction are used in the thesis. However, the question arises whether these should represent a single linear additive index or two sub-indices. Fidelity data from other studies (Borrelli, 2011, Miller and Binder, 2002, Perepletchikova and Kazdin, 2005) show that both adherence to content and competence to deliver an intervention need to be assessed, as there are low correlations between the two behaviours. Therefore, these might represent slightly different constructs of behaviour change and need to be created as distinct sets of criteria representing each aspect of fidelity. The two concepts are discrete for complex behavioural phenomenon; the competence to deliver a BI (e.g., empathy and communication style etc.) is distinct from provider adherence to its content, and both are likely to influence outcome measures (Borrelli, 2011). Having said that, both adherence and quality, albeit having unique features, are still conceptually similar (Carroll et al., 2007, Lorencatto et al., 2013a) and are often considered together in fidelity measures (Weisman et al., 2002).

A thorough account of intervention fidelity could involve multiple measures, even when only a single intervention is delivered once and described by only one change model, because most complex interventions have multiple key features (Nelson et al., 2012). Multiple measures or constructs within an index allow the flexibility of conducting different types of analyses; the overall fidelity scores could be associated with the outcomes to see how much they co-vary, or the outcomes could be associated with the measured constructs (of adherence or quality) to determine the proportion of variance explained by each (Nelson et al., 2012). Hence, to obtain full descriptions of intended activities of a BI as delivered, it is necessary to consolidate both

adherence and quality aspects of fidelity (Lorencatto et al., 2013c), that could be considered separately or combined for an overall fidelity score (Nelson et al., 2012).

It is possible that some complex BCTs in the *Adherence* sub-index, such as 'barrier identification and problem solving' take longer to deliver than BCTs such as 'providing reassurance' in the *Quality* sub-index (Lorencatto et al., 2013a). Although weighting to account for the time taken to deliver each ingredient (or item) was not designed in the index, using the two sub-indices separately and in combination might allow for overall comparison between Adherence and Quality aspects. However, the relative weighting of each item in the index needs to be determined in future work.

6.9.3 Validation of the index

The methods used for content validation of the fidelity index are discussed in this sub-section.

In this study, the modified Delphi technique differed from the standard technique due to the approach taken to administer the procedure (i.e. individual (independent opinion) versus group approach (face-to-face consultation)). Independent opinions of the experts were obtained via email, without carrying out any face-to-face consultations for discussion and disagreement. The generation of ideas individually eliminates the potential of 'group thinking' and minimises the introduction of bias in the process of item selection (Bond et al., 2000b). However, this could have limited the qualitative understanding of the items and the corresponding response scale by the experts, forcing them to choose numbers without open discussion. Delphi was originally devised to handle opinions rather than objective facts (Bond et al., 2000b). I believe that the modified Delphi was appropriate for this study because the items were well described using the taxonomies, making the rating exercise particularly objective.

The Delphi procedure was restricted to a single round of consensus as there was rapid convergence between the experts on ratings of the index items. This might result because of the BI features being well mapped to the BCTs, leading to eliciting similar responses from the experts. Often content ratings of well-established theoretical models lead to experts rating the majority of items as "very important" (Bond et al., 2000b, Holter et al., 2004), resulting in 100% convergence in the first round, rendering a second round of ratings redundant (Mowbray et al., 2003). Forced rating methods or rank ordering of items might present a solution to experts rating items similarly (Mowbray et al., 2003). However, these methods are most suited to studies where the requisite for designing an index is reduction of items. The focus of my research was not reduction of items, and therefore an all-inclusive approach for the items was taken at this stage of development of the fidelity index.

The early convergence of expert consensus could also be due to the small group of experts who might be very similar in their approach, not generating enough variation in their ratings. Behavioural change taxonomies are a rapidly evolving but relatively new science, which is

extensively used within the UK but in early stages in other parts of the world. In this study, the three experts were similar in some respects, such as, being from the UK and being well-acquainted with the taxonomies of BCTs and the strategies of behaviour change used in the ASSIST study, which were very similar to the cessation services in the UK. One expert was a behavioural scientist, another was a smoking cessation expert involved in regional cessation services and the third was a cessation advisor. A relatively small sample comprising three or four experts is often considered sufficient when using confirmatory approaches to item generation (Bond et al., 2000b). However, convergence might have varied if I had consulted a broader range of experts with different backgrounds, necessitating more than one round of Delphi procedure.

6.9.4 Limitations and challenges

A limitation of the study, as hinted in the previous point, was conducting the modified Delphi procedure with three experts. This was partly because there were only a small number of people who were experts in both behaviour change and smoking cessation and who were familiar with the behaviour change model used in the ASSIST study. Although the experts were three in number, they were diverse in their expertise by having experience in designing BIs for smoking cessation or delivering them in practice. One of the experts, in addition to being a behavioural scientist, was also experienced in scale development science.

The major challenge encountered in Study A was conducting the patient-provider BI sessions at the TB clinic. The content pilot was planned to be conducted on at least five patient-provider interaction sessions of BI. However, conducting (30 to 50 minutes) long sessions in the busy outpatient department of the hospital incurred several challenges. Firstly, the provider did not have enough time to deliver the BI to the patients, due to the overwhelming number of patients attending the clinic. Similar findings have been reported from other smoking cessation studies in TB settings (Sereno et al., 2012). Secondly, the patients were also pushed for time, as they had been waiting for hours and were not willing to stay longer after check-up. Finally, there was the unavailability of a separate space to carry out observations of the BI sessions. The TB clinic was a single room where both the TB doctor and the paramedic sat, with some patients waiting inside and some queuing just outside the room. Due to these challenges, only two BI sessions were successfully conducted.

The broader research and practice implications of the developed fidelity index are discussed in the final chapter (Chapter 9) of the thesis.

6.10 CONCLUSION

There were not any existing indices that could be adapted for fidelity measurement of a BI of smoking cessation. Therefore, a new index to quantify the adherence and quality aspects of fidelity was devised, using a comprehensive methodology. The items of the developed fidelity index were composed of the key features of a BI (used in ASSIST study) for smoking cessation

that was mapped to behaviour change theory, giving the index its necessary structure. The response scales of the items were behaviourally anchored to the interaction style of the providers who delivered the BI, representing the functional feature of the index. This anchored response scale provides the discrimination between 'intended practice' and the 'actual practice' of a provider concerning delivery of a BI. Thus, the definitions assigned to the behavioural anchors were crucial to determine how precisely the scale would measure the construct of fidelity for each item.

The fidelity index was refined at this stage (end of study A), further to generation of items and response scale, by content validation and pilot in field. It was then ready for further validation for its psychometric properties, in study B.

Chapter 7. Study B: Psychometric properties of the Fidelity Index

This chapter describes the second study in a series of three linked studies. In the previous chapter, I described the methods used to develop a fidelity index (study A). The output was a theory-linked new measure of fidelity for BIs. The study presented here (study B) investigates how reliable the new fidelity index is, by assessing its psychometric properties.

In the previous chapter, I described the methods used to develop an index for measuring fidelity to BIs of smoking cessation (study A). The fidelity index was constructed using sound methodology, content validated and piloted in the field. In this chapter I report study B, which comprised of a primary investigation, from which the data were used for two purposes: 1) for the psychometric validation of the fidelity index and 2) for exploring any association between intervention fidelity and quit rates (in study C).

This chapter first provides the background literature on methods used for psychometric evaluation of an index to inform the study methodology, and then gives an outline of the study design, settings and methods used for the primary investigation and the statistical analyses. Finally, it presents the results of the psychometric validation of the fidelity index.

7.1 METHODS FOR PSYCHOMETRIC VALIDATION OF AN INDEX

Here I describe the methods used for carrying out a psychometric validation of an index. These methods informed the psychometric evaluation of the new fidelity index and the methodological approach suited to this purpose.

7.1.1 Potential methods of observation for fidelity measurement

Fidelity is a process measure and an intermediary outcome, the measurement of which requires multiple steps (to generate data and coding) before it is ready for use. Structured observations, self-record questionnaires and secondary extraction of routine data are some of the quantitative methods recommended for obtaining data on process variables (Moore et al., 2015). The pros and cons of these methods are briefly discussed here.

Structured observations

In process evaluation, structured observation means monitoring intervention delivery (Eames et al., 2008) on the extent to which its ingredients are delivered, using a structured coding checklist (Moore et al., 2015). Observation methods are potentially useful for assessing variability in intervention delivery practices, providing accurate and objective accounts of the intervention implementation process (Moore et al., 2015). Two types of structured observation methods are described here that can be potentially used for collecting fidelity data:

Direct observation

Direct observation methods are potentially highly objective for measuring the extent to which interventions are delivered as specified (Moore et al., 2015). These include on-site observations or recordings (audio/video) of the interaction sessions.

Direct observations are more intrusive and can introduce a 'Hawthorne' effect, that is, if the provider knows that they are being watched, it can inevitably lead to a change in their target behaviour from the norm (Moore et al., 2015). Moreover, presence of an observer may adversely affect the patient-provider rapport building in the individual face-to-face intervention sessions. In such instances, examining audio or video recordings of the consultation sessions is advisable, as this might be less obtrusive. An added advantage of this type of observation is that the session recording could be rated by multiple observers (Moore et al., 2015), which makes it a preferred method for psychometric testing. Some disadvantages of the recording methods include their being labour intensive, requiring comprehensive training of the coders using these observations and generating a lot of data (Bond et al., 2000b).

Indirect observation

Other methods of monitoring intervention fidelity include *interviews and record reviews* (Bond et al., 2000b), such as provider self-record checklists (e.g. intervention content checklists, encounter logs). These methods of observation are expected to be less reliable than direct observation methods and have low correlations with objective measures (Carroll et al., 2000), but nevertheless they have been used in evaluative research to supplement objective data (Borrelli, 2011).

Provider self-record checklists often serve the purpose of a reminder for delivering the intended content of the intervention and for standardising practices (Bond et al., 2000b). Data from these checklists might be used to supplement direct methods of observation for fidelity measurement and for triangulation during analysis. However, direct methods are preferred in process evaluations, as indirect methods have shown low agreement, when used in conjunction with each other (Bond et al., 2000b). Indirect methods are judged to be more useful in practice than in research, where the purpose of fidelity monitoring is to achieve a certain basic level of adherence to the intervention protocol (Bond et al., 2000b).

Patient self-record questionnaire

Patient self-record questionnaires (e.g. patient exit interviews) play a significant role in formative process evaluation, which precedes the effectiveness evaluation stage of an intervention. However, the structured observations appear to better capture patient-provider interaction and details on practice norms, which are more useful in interpretive process evaluation (Bond et al.,

2000b). These concepts of formative and interpretive process evaluation have been discussed previously, in chapter 1 (1.5.2, page 26).

Self-record questionnaires are inexpensive and have a potentially quick turn-around. However, these could have variable completion rates or respondents may misunderstand questions or might complete them hurriedly (Bond et al., 2000b). The use of patient self-record questionnaires in fidelity assessment often relate to non-intervention-specific process issues, such as patient perceptions of being listened to, (versus being rushed while receiving the intervention) and their satisfaction with the interaction session (Bond et al., 2000b), which are not the focus of the current study. In measuring fidelity using an index, this method could be subject to bias, as the respective patients may not provide a poor rating of their provider or they may not have sufficient knowledge and understanding of the techniques used or the response scale options for rating items (Bond et al., 2000b).

Secondary extraction of routine data

Data acquisition from routine practice can be useful for post-implementation monitoring and standardising provider practice. Record reviews generate potentially objective data, which are simpler to compile than structured observation methods. Potential pitfalls of data extraction from secondary sources include; data not being complete, data may not be accurate or up-to-date, data may not fit the research goals of the evaluator and data may be difficult to access, due to patient and provider confidentiality (Bond et al., 2000b). Although monitoring of intervention delivery using the fidelity index can be implemented post-evaluation during its scale-up, the data acquired might not be suitable for psychometric testing of the index (Bond et al., 2000b).

7.1.2 Coding and rating of fidelity data

Attention to the selection, training, and monitoring of coders for rating data using a fidelity measure is important (Schoenwald et al., 2011). Coding using a fidelity index for a BI can be complex, as it involves not only the ability to judge and score based on the functional features but also a generic understanding of the BCTs to be able to understand its compositional features.

The researchers involved in a fidelity study can act as the coders themselves, or the coders can be selected from a pool of candidates independent of the study (Bond et al., 2000b). Nevertheless, independent coders are expected to provide less biased scorings of fidelity as they are less likely to be invested in obtaining findings consistent with the preconceived notions about intervention fidelity (Bond et al., 2000b). Still, there might be individual differences in the way the independent coders assign scores, such that apparent changes in fidelity scores might actually occur as a consequence of differences between the coders, rather than the provider fidelity (Schoenwald et al., 2011). Therefore, when selecting coders, it is important to consider their

skills regarding objectivity and critical thinking, in addition to their understanding of the topic under study (Bond et al., 2000b).

Apart from the consideration of their characteristics, skills and understanding of the topic of study, coders must also receive comprehensive training on general coding procedures and on specific features of the fidelity index (Fowler, 1995). Other issues that need to be considered in training of coders include confidentiality and anonymity of data being coded. Amateur coders may be reliably trained to code the content of intervention manuals and session transcripts, using the taxonomies of BCTs, after receiving adequate training (Lorencatto et al., 2013b). However, in evaluations aimed at uniformity of providers' practice across multiple sites, additional, more formal coding is usually required (Borrelli, 2011).

A further consideration when coding fidelity data is to keep a log of important decisions made to reach consensus and to document convergence among coders (Black et al., 1999).

7.1.3 Psychometric tests for index validation

One of the two important factors to consider for psychometric evaluation of an index is the application of adequate tests depending on its stage of development. The key psychometric property of interest, at this stage of formation of the newly developed fidelity index, is its reliability. It is necessary to gather evidence that the index is measuring fidelity in a reproducible and replicable manner, before investigating whether it can measure the association of fidelity with quit rates, that is, its predictive validity (Streiner and Norman, 2008). These two stages are sequential in the validation of the fidelity index in terms of its properties, that is, whether it is reliably measuring fidelity or its structure needs further improvement before the scores obtained on fidelity could be linked to an outcome (quit rate) to assess if the index has predictive qualities.

The second factor, which is connected to the reliability testing, is the appropriate sample size for carrying out the desired statistical analyses.

Reliability testing

Reliability testing provides evidence of the value of an index, by demonstrating whether the measurements of individuals at different instances, or by different coders, or by similar tests, produce the same results (Streiner and Norman, 2008). Reliability refers to the consistency of responses for a particular item in the index. It is the ratio of the variability between individuals to the total variability in the scores. In other words, it is the measure of the proportion of the variability in scores, which is due to the true differences between individuals or the same individuals at different time points. Reliability is expressed as a number between 0 and 1, with 0 indicating no reliability, and 1 indicating perfect reliability (Streiner and Norman, 2008). Various authors have made different recommendations for the minimally acceptable level of reliability;

internal consistency measures are expected to exceed 0.8 and it might be reasonable to demand stability measures greater than 0.5 (Streiner and Norman, 2008).

For examining the reliability of fidelity indices, a few terms and concepts are broadly defined, to understand the statistical considerations (presented in 7.3.10, under Analysis), as follows:

Stability

There are various ways of examining the replicability of a measure. For example; measurement of the degree of agreement between different coders (inter-coder reliability), the agreement between observations made by the same coder on two different occasions (intra-coder reliability) or observation on the patient on two occasions separated by some interval of time (test-retest reliability), and so forth (Streiner and Norman, 2008). If the focus of psychometric assessment for a particular measure, for example, is to see whether the measure can be reliably rated by different coders without their individual influence on its measurement ability, then only inter-coder reliability can be tested; otherwise if the focus varies, then one or all of these tests can be used to establish its stability and replicability for future use.

The assessment of inter-coder reliability (also called inter-rater agreement) provides a way of quantifying the degree of agreement or disagreement between two or more coders, who make independent ratings about the characteristics of a set of subjects. This type of analysis utilises the classical test theory (described on page 108) and aims to determine how much of the variance in the observed scores is due to the variance in the true scores, after the variance due to measurement error between coders has been removed (Novick, 1966), such that

$$\text{Reliability (r)} = \frac{\sigma_T^2}{\sigma_T^2 + \sigma_E^2}$$

Where σ_T^2 (sigma-squared T) is the variance of the true score and σ_E^2 is the measurement error.

A number of statistical tests can be used to determine *inter-coder reliability*, depending on the types of measurement. These include percent agreement, inter-coder correlations (α-alpha and ICC- defined on page 44), Scott's pi (π), Cohen's kappa (κ), Fleiss' kappa (K) and Krippendorf's alpha (α). Krippendorff alpha has a number of benefits over percentage agreement and Kappa statistics methods that are popularly used for inter-coder reliability testing. This test can be used for any number of coders (and not just two) and for different kinds of variables (nominal, ordinal, interval, ratio, and more), unlike other inter-coder agreement tests. Unlike Kappa, it can be used for large or small sample sizes and has no minimum sample size requirements. Furthermore, it can be used for incomplete or missing data, as it uses a system of bootstrapping, where missing values are replaced with existing values samples from within the data set.

Reproducibility or coherence of items

If the index has a relatively large number of items addressing the same underlying concept, then it is reasonable to expect that scores on each item would be correlated with scores on all other items, or there will be coherence between items of the index. Estimates of internal consistency (e.g. Cronbach alpha) represent the average of the correlations among all the items in a measure. However, these estimates do not take into account any day-to-day or between-coders variation, and thus, are likely to lead to an optimistic interpretation of the true reliability of the test (Streiner and Norman, 2008).

Principal Components Analysis (PCA), is a statistical procedure that transforms a set of observations of possibly correlated variables into a set of values of linearly uncorrelated variables called principal components (Streiner and Norman, 2008). The number of principal components is less than or equal to the number of original variables. This transformation is defined in such a way that the first principal component has the largest possible variance (that is, accounts for as much of the variability in the data as possible), and each succeeding component in turn has the highest variance possible to the preceding components (described further in 7.3.10, page 121). PCA when conducted on items within every coder, could help identify those items that load together reliably across all coders, taking into account the between-coders variation, unlike estimates of internal consistency.

Below is the general form of the formula to compute scores on the first component extracted in a PCA (Larry and Edward, 1999):

$C_1 = b_{11}(X_1) + b_{12}(X_2) + \cdots b_{1p}(X_p)$

Where;

C_1 is the subject's score on principal component 1 (the first component extracted)

$b_{11} \ldots b_{1p}$ = the regression coefficient (or weight) for observed variables 11 to 1p, as used in creating the principal component 1

$X_{1\ldots} X_p$ = the subject's score on observed variables 1to p

In the above example I presented two components, which is only hypothetical for description purposes. In reality, the 'number of components extracted' in a PCA is equal to the number of observed variables being analysed. In most analyses, only the first few components account for meaningful amounts of variance, so only these are retained and interpreted.

Reliability in differentiating between providers based on their fidelity

In addition to defining important terms for reliability testing, there are two theories for examining reliability of the measure, in terms of its ability to differentiate between various dimensions of the construct (fidelity) under study. In the current study, the dimension of interest

is the provider practice and it is important to understand which theory best fits the purpose of differentiating between provider practice. The two theories are briefly described here:

'Classical Test Theory'

In classical test theory, the observed score (X) from the measure is thought to be composed of a true score (T) that represents the subject's score, obtained, assuming no measurement error, and an error component (E) that is due to measurement error (also called noise) (Lord, 1959, Novick, 1966) , such that:

Observed Score (X) = True Score (T) + Measurement Error (E)

Measurement error (E) prevents one from being able to observe a subject's true score directly, and may be introduced by several factors. For example, measurement error may be introduced by; imprecision, inaccuracy, or poor scaling of the items within an index (i.e., issues of internal consistency); instability of the index in measuring the same subject over time (i.e., issues of test-retest reliability); and instability of the index when measurements are made between coders (i.e., issues of inter-coder reliability) (Streiner and Norman, 2008).

Reliability coefficients (e.g. internal consistency and inter-coder reliability) in classical test theory tend to overestimate the replicable (true source of error) and underestimate the error variance in a set of scores. This is particularly a problem for measures in which the coders (e.g. trained observers) provide scores for participants based on observation of complex interactions such as those involving behaviour. With these coefficients, it is only possible to examine a single source of measurement error at a given time. Besides, it does not permit studying the interaction effects that occur among these different sources of error (Preuss, 2013).

'Generalizability Theory'

The other type of theory for testing index reliability for differentiating between various dimensions of fidelity is the 'Generalizability theory', which is often used to test reliability of a measure to enhance precision of its measurement. With coder-rated measures of observations, multiple sources of error variance (e.g. coders, items) can affect the replicability of the respective measures (Lakes and Hoyt, 2009).

In contrast to the classical test theory, generalizability theory is more flexible and accurate in quantifying the extent to which the observed scores reflect error of measurement, rather than the characteristics of the individuals under study (Lakes and Hoyt, 2009). One of the principle advantages of using generalizability theory is that it simultaneously considers multiple sources of measurement error variance and provides separate variance estimates for each facet (or source of variation) not otherwise attributed to the object of measurement.

In generalizability theory the 'true score' is equal to the 'universe score'- the hypothetical mean of all acceptable (i.e. interchangeable) observations. G-study is based on generalizability theory and computes the individual variance components, by incorporating all the plausible sources of error into a single analysis of variance. An important goal of measurement by G-study is to attempt to identify, measure, and thereby find strategies to reduce the influence of these (identified) sources on the measurement in question (Streiner and Norman, 2008). The G-study involves three sequential steps:

i. Identification of important facets (sources)
ii. Variance partitioning
iii. Computation of coefficient of generalizability (g)

In a G-study, error is equated with variance in observed scores that is attributable to the facets (e.g. the items, coders or measurement occasions) or sources of variance that are irrelevant to the dimensions of interest, as well as to unexplained variations in responding (i.e. the random error) (Lakes and Hoyt, 2009). The undifferentiated error term in classical test theory is partitioned in the G-study into components attributable to the main effect of each source, as well as interactions between the sources, and between sources and the object of measurement (which is the audiotapes nested in the providers). Finally, the *g* coefficient is computed using the variances obtained from variance partitioning analysis that represents the ratio of variance attributable to universe scores to the total observed score variance (i.e. universe variance plus variance attributable to all sources of error that contribute to variance in observed scores).

Often, psychometric testing of an index would involve a mix of methods applying classical test theory and generalisability theory. Both have their advantages, depending on the research question and the psychometric property of interest for a given index. Therefore, for evaluating the functioning of the items in the fidelity index, I would summarise the quality of the overall index (composite fidelity scores) using descriptive statistics and the G-study. For assessing the quality of the individual items, I would use three different sources of information: item descriptive statistics for variance or ceiling or floor effects (defined in the Methods section), Krippendorff alpha to identify those items that were most easily agreed upon by all coders and the PCA to identify those items that load together reliably across all coders.

Sample size considerations

Determining the minimum number of data-points necessary for conducting an appropriate analysis is the first consideration in psychometric testing of a measure. If the intent is to conduct PCA to assess internal consistency, some authorities recommend a minimum sample of at least 150 or 200 (Hinkin, 1995). However, many other psychometric studies of fidelity measures have

used samples in the range of 18-32 (Bond et al., 1997a, Bond et al., 2000a, Lucca, 2000) or fewer (Teague et al., 1995).

Two factors considered important in determining the sample size for PCA are the total number (N) of participants and the ratio of participants to items (Osborne and Costello, 2004). There is a widely-cited rule of thumb that the participant to item ratio for such an analysis should be at least 10:1, but this recommendation is not supported by published research (Nunnally, 1978). A study suggested that a sample size between 50 and 100 was adequate for PCA to evaluate psychometric properties of measures of social constructs (Sapnas and Zeller, 2002). However, this study has been criticised for failing to explain the conditions in which it might be feasible to use a small sample size (de Winter et al., 2009). In a more recent work, it was concluded that a sample size between 10 and 50 was sufficient for two dimensions (e.g. Adherence and Quality) and 20 items (Zeller and Martsolf, 2002). In psychometric analysis, the recommendations on absolute N and the N by item ratio have gradually been abandoned as misconceived (de Winter et al., 2009).

In summary, the larger the sample, the more stable will be the statistical estimates of reliability. Caution should be taken when making extreme modifications in index or drawing strong conclusions in the instances where the study sample is small.

7.2 AIMS OF STUDY

"To validate the fidelity index by assessing its psychometric properties" (described above in 5.2).

The specific objectives of study B are to:

- Test the ***feasibility*** of use in primary research
- Determine its ***replicability*** for use by people with diverse backgrounds, to inform its stability as a research tool
- Determine the ***coherence*** of items in the index, to refine its structure and inform its reproducibility in future research
- Investigate the underlying ***dimensions of the fidelity construct*** within a behaviour change framework that the index can capture, to inform its utility in explaining variations in smoking cessation outcomes

7.3 METHODS

7.3.1 Design

The fidelity study was a cross-sectional design used to objectively measure intervention fidelity for a BI.

7.3.2 Study sample

The fact that I wanted to use the fidelity scores (in study C) and same methods to ensure I could link back to the same providers and settings used in the ASSIST study, had some implications for the fidelity study settings:

- Recruiting all those TB clinics that took part in the ASSIST study to enable exploration of the association between fidelity and quit rates.
- Recruiting same providers (TB DOTS paramedics) who delivered BI in the ASSIST study, as this might ensure some consistency in the delivery of the BI.

Therefore, feasibility of recruiting the same TB clinics and providers that were originally involved in the ASSIST study was assessed, by conducting an initial scoping exercise to ensure:

- Whether all of the originally involved TB clinics continued to deliver the BI in routine practice after termination of the ASSIST study and if not, what should be the strategy to re-introduce BI for smoking cessation for the fidelity study
- Identification of the same providers as were involved in the ASSIST study. A single TB DOTS paramedic is appointed at each TB clinic by the NTP in Pakistan.
- Hiring the same research officers who were previously involved in the execution of the ASSIST study
- Feasibility of using study equipment (described below) in TB clinics, as this was not previously tested in ASSIST study

The findings of the scoping exercise are presented in results section.

7.3.3 Setting and participants

The 22 TB clinics, in intervention conditions, originally involved in the ASSIST study were approached for participation in the fidelity study. These were located in two districts (Jhang and Sargodha) of the Punjab province of Pakistan (Figure 7.1).

Five out of the 22 TB clinics were secondary-care hospitals (called 'Tehsil Headquarter hospitals') while the rest were all PHCs (called 'Rural Health Centres').

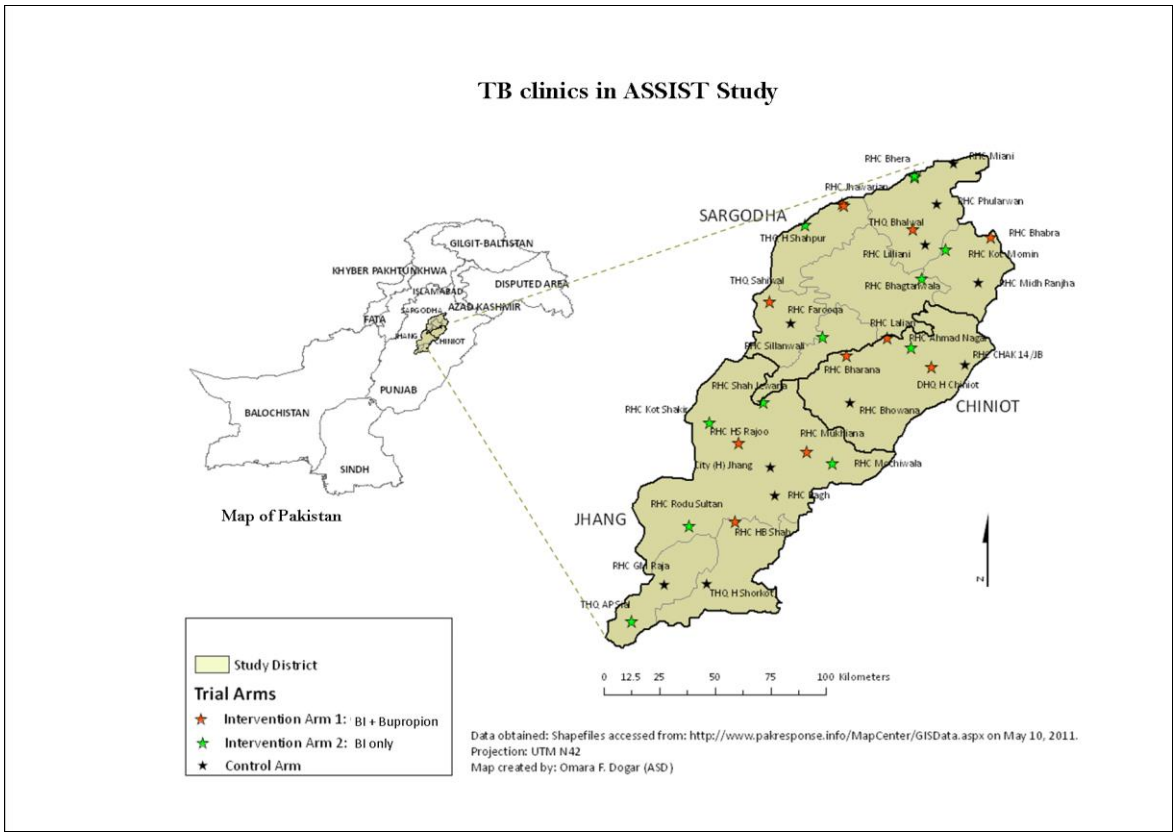


Figure 7.1: TB clinics included in the ASSIST study

Site (and provider) inclusion criteria

All TB clinics (originally in the intervention conditions of the ASSIST study) were eligible to participate in the study on invitation. TB clinics that were involved in the control condition in the ASSIST study were excluded. Only those TB clinics, out of the 22, were included where the providers were willing to take part in the fidelity study.

Patient inclusion criteria

Consenting patients aged 18 years or older with suspected pulmonary TB (cough for > 2 weeks without any other cause) or newly diagnosed with TB who were also regular tobacco smokers (>1 cigarette or hookah session/day) attending TB clinics included in the study, were enrolled. Patients requiring hospitalization or urgent medical attention were excluded.

Patient identification and recruitment

TB patients who smoked were identified, per ASSIST study protocol, by the TB physicians (one per TB clinic) and referred to the providers, who were responsible for recruiting eligible and informed consenting patients in the study.

Consent procedures

Voluntary informed consent was obtained from the patients for participation in the study by the providers, after giving them appropriate written and verbal information about the proposed study. (Providers' consent was obtained by the research officers).

An ink signature or a thumb impression, which is officially acceptable in Pakistan for those who cannot write, was obtained from those interested (patients), by going through a checklist on the consent form. The option was also given to consider the information for a week before agreeing to participate in the study.

All information sheets and consent forms were translated into the local language of communication (i.e. Urdu) for administration. These are supplied as Appendices (Appendix D.1 and Appendix D.2).

Further details on informed consent are provided in section 7.3.11, page 123.

7.3.4 Intervention

The BI used in this study was the same as ASSIST study, described above in Chapter 6 and also in the Appendix B. The TB DOTS paramedics delivered the BI for smoking cessation to the TB patients, therefore, they are referred to as 'providers' of the BI throughout this thesis.

7.3.5 Data variables

Two-sets of data were obtained from the fidelity study: the flip-chart counselling for BI and the patient assessment questionnaire. These data were used to code the fidelity index. In addition, the questionnaire also produced data on provider self-record checklists. There was some overlap in sources of data used in studies B and C, so to make this easier to follow, I describe the main types of data and its sources collected in the fidelity study in Figure 7.2.

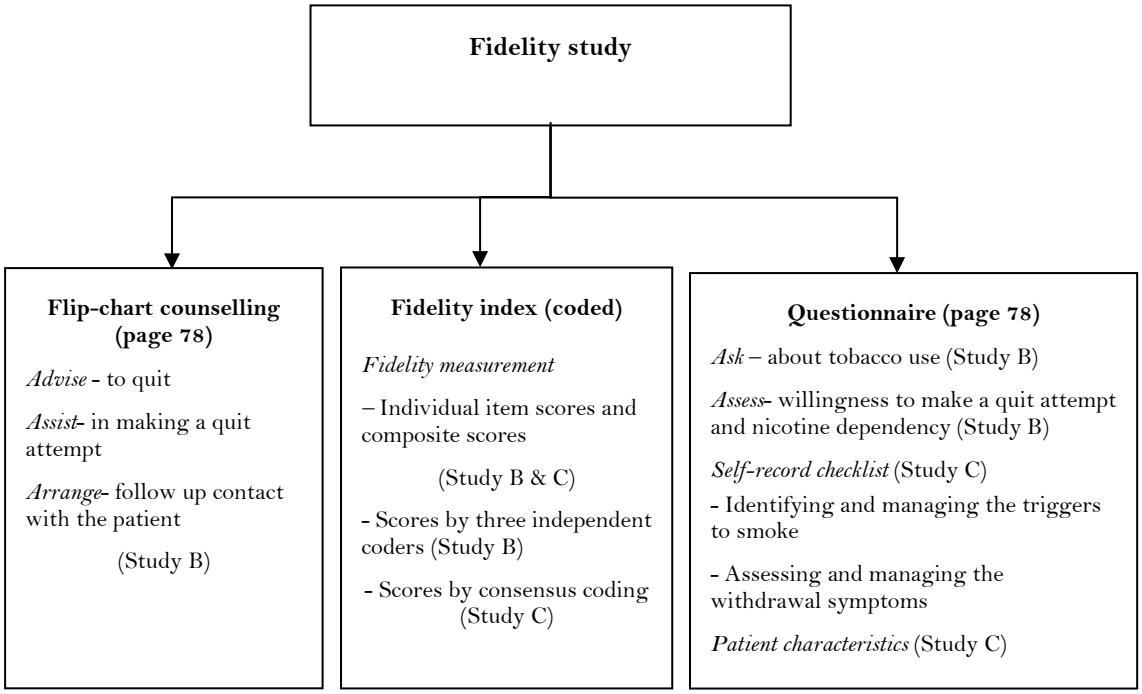


Figure 7.2: Sources of data collection in the fidelity study

7.3.6 Data collection

The data collection and coordination for this study was carried out through a local Non-Governmental partner located in Islamabad that had established linkages with the Department of Health in the two districts and the National TB control programme.

Audiotaping of the patient-provider BI interaction (included both the administration of assessment questionnaire and the flip-chart counselling) was undertaken by the providers, using a discrete device (see 'study equipment' below) during a four month period of data collection. A BI session in general was expected to last 30 to 40 minutes.

7.3.7 Tools used for data collection

Patient assessment questionnaire

The patient assessment questionnaire (Appendix D.3) was translated into the local language of communication (i.e. Urdu). It was administered by the providers to collect information on patient demographics, tobacco use behaviour, nicotine dependence, willingness to quit, past quit history, the agreed quit date, and filling in the self-record checklist (refer to Chapter 8, Table 8.1).

Study equipment

Digital recorders for audiotaping BI sessions were given to the providers, who were trained on their proper use. The digital recorder selected for use in this study was the Zoom H2n handy recorder, which had the following features:

- Five built-in microphones and four recording modes
- The technology used covers a wide area while still capturing sound sources in the centre with clarity and definition, making it perfect for all types of live stereo recording
- The H2n's built-in microphone provides two matched unidirectional microphones set at a 90 degree angle relative to one another, optimum for most stereo recording applications
- The H2n requires only 2 AA batteries, either alkaline or rechargeable NiMH. Battery life when using alkaline batteries is more than 20 hours, even during continuous recording. Alternatively, an AC adapter can be used that allows powering the H2n from any standard wall socket.
- The H2n records directly to SD cards. It supports standard SD and SDHC cards, up to 32 gigabytes.
- The H2n USB port provides a digital output of the stereo mix and allows data to be sent to and from the computer.

The selection of zoom H2n handy recorder for this study was done after much thought about the local settings of the current study and comparing field notes (described in 7.4.1, page 125) with a similar study on recording of patient-provider counselling sessions in a hospital setting, in the UK.

7.3.8 Personnel

The research officers were the focal persons in each district for monitoring and supervising study activities. They approached the TB providers to participate in the fidelity study, were involved in their training on the BI flip-chart and the assessment questionnaire, and were responsible for the day-to-day data collection, storage and its transfer tasks.

The TB DOTS paramedics (the BI providers) were auxiliary nurses who follow a physician's clinical directions on TB treatment. They record patient progress in TB registers, monitor direct administration of TB medications, and ensure clinical follow-up of patients in the TB clinics. In the fidelity study, these providers also recorded data on the patient assessment questionnaire and set up the equipment for audiotaping BI sessions.

7.3.9 Procedures

Training of the research officers

Two research officers were hired (one of whom was previously involved in ASSIST study) and trained on the study procedures to carry out the required data collection, data keeping, monitoring of the providers and transferring of the data, whilst maintaining anonymity of participants.

The research officers were trained on the BI (6.3, page 78) by the lead researcher, who then trained the providers in the two districts. The training adhered to the BI manual and procedures used in the ASSIST study to minimise any influence of factors that might have altered provider practice over time.

The research officers were involved in the content piloting (described in Chapter 6) of the fidelity index (in Islamabad). During the pilot, they carried out the audiotaping of patient-provider interactions using digital recorders, and shifted data to encrypted drive for secure transfer to the central office. This process of field testing helped identify and remedy any procedural, logistical or programmatic issues, in addition to training of the research officers on study procedures.

Training of providers

Full-day training was conducted for all providers in each district, at the district health office, on December 26, 2013 in Jhang and March 29, 2014 in Sargodha (see 'project plan' below). A provider, who was unable to attend the refresher training in Jhang, was trained in the respective TB clinic by the research officer, before starting patient enrolment.

All participating providers were given refresher training by the research officers on delivering the BI using the flip-chart and completing the assessment questionnaires with the participating patients, exactly as they would have done in the ASSIST study. The training was based on the practice manual developed as part of the ASSIST study for providing BI.

In addition, providers received orientation about the ethical issues concerning data safety, patient confidentiality and anonymity of collected data. They also received training on proper usage of study equipment (see page 114).

Selection and training of coders

Coders (three in number) were selected after interviewing from a pool of eligible candidates, who were bilingual (English and Urdu), obtained a Master's degree in their respective field of study and preferably able to comprehend Punjabi (the local language at study sites). One coder was a doctor and a public health practitioner, the second was a social anthropologist with no prior

experience of health related research and the third was a doctor and an epidemiologist (the lead researcher of the current study).

Full-day training was provided to the coders, by the lead researcher, using a BI coding manual (Lorencatto et al., 2013b). This manual is shown to effectively train inexperienced coders from multi-disciplinary professional backgrounds for reliably specifying intervention content, using the BCTs. The overall purpose of the fidelity index and each specific item within it, was also discussed in-depth with the coders during the training, along with practice exercises using two audiotapes recorded as part of the 'content pilot' (page 83) of the fidelity index in Study A: Development of the Fidelity Index.

Process of coding fidelity data

The audiotapes were played in a room where all three coders sat together, coding the BI sessions using the fidelity indices independently, but at the same time. The audiotapes of the BI sessions were in Punjabi and these were coded directly into English when scored using the fidelity index. The mean duration of each BI session delivered by each provider was also noted.

For the consensus development process, if the coders disagreed on the rating of any item, this was logged and discrepancy resolved through discussion. At instances, this process involved clarity on the part of the case-definition or '*behavioural anchoring*' (described in 6.7.3, page 87) underlying the response scale in discriminating scoring criteria of fidelity for BI ingredients. Each coder completed an independent fidelity index for each audiotape. A fourth fidelity index (consensus score) was also completed, containing scores after agreement by all coders for each item.

Data entry

The patient assessment questionnaire data was entered in a database created using SPSS 21 software package. The same codebook as used originally in the ASSIST study was referred to for data entry. Data from the coded fidelity indices were also entered in an SPSS (version 21) database, for which a codebook was created to facilitate data entry.

Data storage

Audiotapes and the completed assessment questionnaires were transported by the research officers after completion at each TB clinic and secured in a locked cabinet in the district health office. Audiotapes were stored on encrypted USB drives by the research officers (using truecrypt, free encryption software). The completed set of audiotapes from each district was sent to the Islamabad office, via registered courier.

All data (in the raw form: hard copies and audiotapes) were stored using unique identifiers and not patient names, for e.g. Clinic (per provider) number, Recording (per patient) number. The list of TB clinic codes was only accessible by the lead researcher.

Data stored in the encrypted drives and backed-up on the University of York computer is kept and analysed at the University of York. No other party would have access to the audiotapes or their content. The audiotapes are to be secured at the University of York for 5 years post-PhD study, to ensure availability of the data to answer any queries that arise from the thesis work and also the possibility of using it for post-doctoral opportunities.

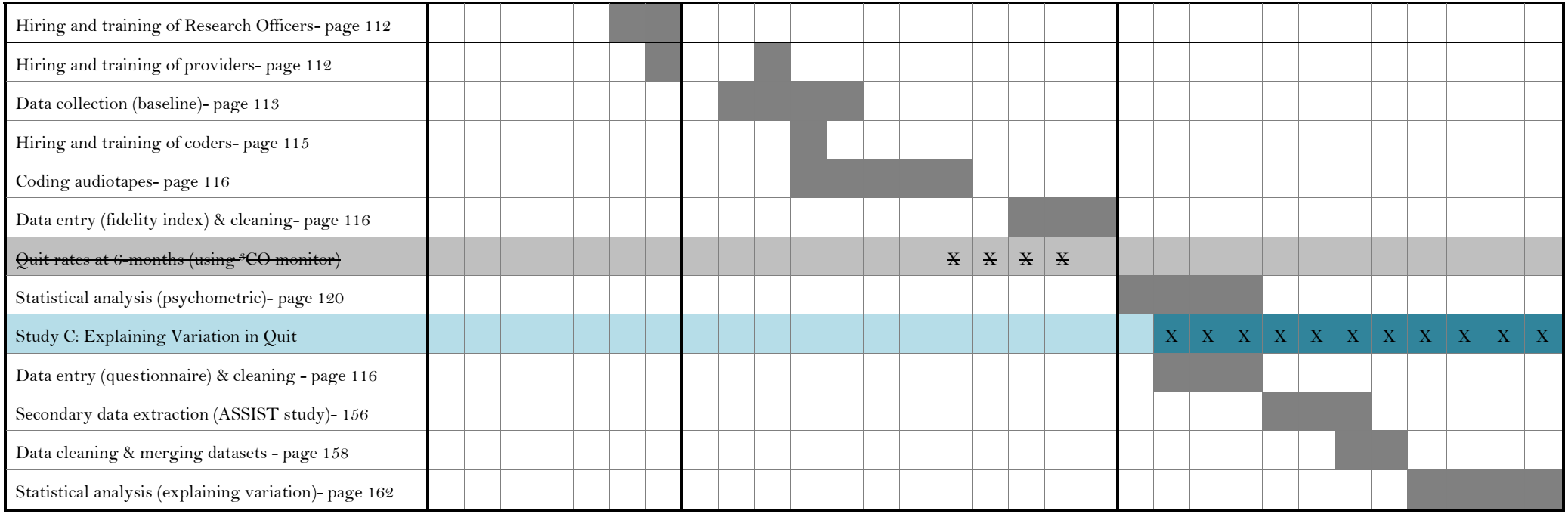
Project plan

The project plan (Table 7.1) gives the timing of the various investigations and activities for the studies, A (Chapter 6), B (Chapter 7) and C (Chapter 8).

Table 7.1: Project plan

Activities	2013							2014												2015											
	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Develop project protocol																															
Ethics application to [1]PMRC, Pakistan																															
Ethics application to [2]HSRGC, York																															
Study A: Development of the Fidelity Index			X	X	X	X	X																								
Generating index items – page 84																															
Establishing response scale - page 87																															
Modified Delphi procedure- page 87																															
Content pilot- page 83																															
Study B: Psychometric properties of the Fidelity Index	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X	X								
Developing information sheets and consent forms- page 112																															
Feasibility of proposed work (incudes obtaining and testing study equipment)- 125																															

[1] Pakistan Medical Research Council
[2] Health Sciences Research Governance Committee



[3] Carbon Monoxide breath test monitor

7.3.10 Statistical considerations

This section provides the sample size requirements for the reliability assessment of the fidelity index; these calculations were informal (considerations discussed in 7.1.3, page 108) and mainly derived from insights of expert psychometricians. Methods selected for the psychometric evaluation are also presented.

Sample size

It was anticipated that a total of at least 450 data-points would be collected, influenced by the number of TB clinics agreeing to participate in the study (as discussed above in 7.1.3, page 108). That is, if 15 TB clinics (out of the 22 approached agree to participate)* 10 audiotapes (TB patients enrolled) per TB clinic* 3 coders per audiotape (as proposed).

On average, five to ten BI sessions (depending on the feasibility and practicality of audiotaping in these settings) were anticipated for recording at each TB clinic.

Whilst in this study the sample size is driven by the participating TB clinics, it is still important to know if the sample was large enough to assess the stability and coherence of the fidelity index. Approximately 100 subjects were considered suitable for obtaining an estimate of coefficient α of 0.7 with confidence interval of $\pm$ 0.2 for an index with 45 items (Streiner and Norman, 2008). Similarly, assuming three coders rated the fidelity index on observed data, then approximately 100 subjects would be required to obtain a sufficiently powered inter-coder reliability estimate of 0.75 with standard error of $\pm$ 0.05 (Streiner and Norman, 2008).

Analysis

Three types of reliability analyses were carried out for the fidelity index: inter-coder reliability, PCA and the G-study. PCA and inter-coder reliability were used to assess 'individual item scores', while the G-study was used for assessing the 'composite index scores'. Descriptive statistics of the individual and composite scores were also calculated. All analyses were conducted using SAS (version 9.4, Cary, NC, USA); see Table 7.2, on page 123, for the SAS syntax of these tests. These methods are described below.

Descriptive statistics for individual item and composite index scores

A univariate analysis of the summary statistics for individual item scores and composite index scores (both for Adherence and Quality) was carried out for the three coder ratings, as well as for the fourth 'consensus score' (described in 7.3.4 on page 116). Mean, median and standard deviations were reported for the individual item scores.

Pearson's correlation coefficients (Pearson's *r*) were estimated for the two composite scores of the fidelity index ('Adherence' or A score and 'Quality of interaction' or Q score). Since there were only a few correlations and they were not independent, Pearson's *r* was not conducted to draw inferences on the quality of composite scores but instead it was conducted to summarise these scores. Pearson's *r* reflects the degree of linear relationship between two variables, and ranges from +1 to -1. For non-normally distributed scores, these need to be transformed before application of the test.

Transformation of Pearson's *r* for non-normally distributed scores:

The formula for Fisher's *z'* transformation is $z' = .5[\ln(1+r) - \ln(1-r)]$, where ln is the natural logarithm. Two relevantly important attributes of the *z'* statistic are that it is normally distributed and it has a known standard error, which is used for computing the confidence intervals on Pearson's *r* and the difference between the correlations (Lane, 2013).

Pearson's *r* coefficients were aggregated across coders, using Fisher's *z'* transformation to convert Pearson's *r* to the normally distributed variable *z'* across the coders, to give an overall reliability estimate. Fisher's *z'* estimates were averaged across the coders and transformed back to Pearson's *r*, using the r to z' table, for both A and Q scores.

The scatter plots from the Pearson's product moment correlation matrices were also displayed.

Inter-coder reliability

Inter-coder reliability estimates were computed for the individual item scores using *Krippendorff's alpha* across the coders (with and without the consensus scores) to identify those items that were most easily agreed upon by all coders. The estimates without the consensus scores were used to judge the quality of the items. This helped judge the items that need to be kept in the final set. The estimates, including consensus scores, were used to discuss the added value of a consensus round in scoring by independent coders.

The KALPHA macro (see Table 7.2, page 123) computes Krippendorff's alpha reliability estimate for judgments made at any level of measurement, any number of observers, with or without missing data (Hayes and Krippendorff, 2007).

Principal Components Analysis (PCA)

PCA was used to find the combination of items in the fidelity index which contained as much of the available information as possible (Motallebzadeh et al., 2007). PCA was undertaken chiefly to inform the structure of the index, by identifying those items that load reliably (together and separately for *Adherence* and *Quality*) across all coders. This was done by restricting the analysis

to load items on a single component to identify if any coder/item combinations did not load together.

The PCA was further used to discriminate the dimensions of the underlying intervention 'content' (refer to section 4.2.2, page 65) and cluster items together that correlate strongly with each other within these dimensions. This was done for the A scores, by restricting the analysis to five components, using Eigenvalues of >1 as a reference for component extraction.

PCA was conducted in SAS Proc Factor (see Table 7.2, page 123), using an orthogonal pre-rotation, which first rotates the item structure orthogonally and identifies maximal differences between these items, and then adjusts afterwards to a correlated item structure. Values of ≥ 0.4 were considered meaningful for the items that loaded on a component.

Generalizability Study (G-study)

A G-study was conducted to investigate the underlying dimensions (i.e. providers fidelity to BI) that the index measures. It involved three sequential steps: identification of important facets (sources) of variation, variance partitioning and computation of coefficient of generalisability (*g*). The G-study assessed the reliability of the fidelity index, in terms of measuring provider differences in intervention fidelity, which was explored using the generalizability theory. A two-facet G study using a mixed models factorial ANOVA with random-effects (for SAS syntax: Table 7.2, page 123) was used to partition the variance components under the generalizability study with a nested design. In addition to the variance components, variance percentage, that is, each variance component as a percentage of the total variance (sum of all variance components in the model) was computed. This is recommended for models with multiple sources (Lakes and Hoyt, 2009) and aided interpretation of variance contributed by each source, allowing comparison with crude ICC (in sensitivity analysis below). The A and Q scores were analysed independently in the G-study.

Further, three types of sensitivity analyses were explored for the *g*-coefficient:

i. Crude ICC (intra-cluster correlation coefficient) for the provider differences
ii. G-study excluding the consensus scores
iii. G-study excluding providers with very low fidelity

For the sensitivity analyses, ICC coefficients were computed, to indicate how much of the total variation in fidelity measurement was accounted for by the providers. These estimates were crude and not drawn from variance partitioning with other sources (facets) in the model. Therefore, the crude ICC would give a comparison for the amount of variation contributed by the providers with G-study estimates when other sources were accounted for in the model. As the consensus scores were considered the fourth coder in G-study, it was worthwhile to further explore how the g-coefficients changed on excluding these scores. Furthermore, it was decided to

exclude providers with very low fidelity as they showed very little variance in items to explore whether the providers with higher fidelity still varied in their practice.

Table 7.2: *SAS syntax for reliability tests*

Analysis	**SAS Syntax**
Krippendorff's alpha	The syntax (in SAS) used for the macro is: `%kalpha (data = ….,judges=ITEM1coder1 ITEM1coder2 ITEM1coder3 ITEM1 consensus, detail =1, level =2, boot =2000);` where judges is a list of variable names holding the names of the coders, level is the level of measurement (1 = nominal, 2 = ordinal, 3 = interval, 4 = ratio), detail is set to 1 if you desire to print the coincidence and delta matrices, and boot commands the number of bootstrap samples desired for inference.

PCA	The syntax used was:
	`PROC FACTOR DATA=RESTRUCT METHOD=PRIN PRIORS=SMC SCREE ROTATE=PROMAX FLAG=.30 NFACT=1;`
	`Var Item1a Item2a …. Item37a; Run;`

G-study	The syntax used was (Putka and McCloy, 2008):
	`PROC MIXED data = CONCAT ALPHA=0.05 NOITPRINT METHOD=REML ASYCOV COVTEST;`
	`class CODER_ID RECORDINGS HC;`
	`MODEL A_SCORE=;`
	`RANDOM HC RECORDINGS CODER_ID RECORDINGS(HC) HC*CODER_ID; run;`

7.3.11 Ethics considerations

The key ethical issues relating to the individuals participating in the study were dealt with as follows:

Audiotaping of BI sessions and patient interviews

Audio-taping of the BI sessions required obtaining informed consent both from the providers and the patients.

Provider informed consent

Providers were given both verbal and written information on the purpose of observing (audiotaping) their interaction sessions with the patients, before getting consent (Appendix D.1). The information sheet included the following: that taking part was entirely voluntary; the provider was not forced or obliged to allow audiotaping of any session; the provider's identity and location would remain strictly confidential; non-participation in taping the session would not prejudice their employment; they had the right to withdraw consent at any time, including during the session and after the tape had been completed (data up to the point of consent only would be used in this case); tapes would only be viewed or heard by the research team members, and external reviewers in the expert panel if required.

Patient informed consent

Patients were given both verbal and written information on the purpose of observing (audiotaping) their interaction sessions with the providers and the questionnaire led interviews, before getting consent (Appendix D.2). The information sheet included the following: that taking part was entirely voluntary; the taping of sessions was purely to assess or record the provider intervention delivery practice; the patient was not forced or obliged to allow audiotaping of any session; the patient's identity and location would remain strictly confidential; non-participation in taping the session would not prejudice their care, intervention or service offered; the patient had the right to withdraw consent at any time, including during the session and after the tape was completed, (data up to the point of consent only would be used in such cases); the patient had the right to listen to the tape after the session if desired; tapes would only be viewed or heard by the research team members, and external reviewers in the expert panel if required.

Health risks to the researchers

Assessment of the health risks to the research officers visiting the TB clinics for data collection and placement of appropriate actions for their safety was important. These research officers were employees of the government and already worked closely with the TB DOTs paramedics for disease reporting purposes, and were trained on taking appropriate precautionary steps before visiting a TB clinic. The lead researcher was not in direct contact with the TB patients during data collection.

Safe keeping and confidentiality of data

After induction, both research officers were introduced to the principles of confidentiality and informed consent and made aware of the study protocol and their obligations to follow it.

Ethics Approvals

Ethics approval for conducting the observational study was granted by the Health Sciences Research Governance Committee (HSRGC) at the University of York and locally by the National Bio-ethics Committee at Pakistan Medical Research Council (PMRC) - see Project plan (page 116).

7.3.12 Sponsorship

The Bupa foundation highly commended the ASSIST study's contribution to evidence-base in practice and awarded the research team with a small amount of prize money to be utilised for further work on the subject. This award money funded the activities of the observational study.

7.4 RESULTS

7.4.1 Feasibility findings from the scoping exercise

The research officers hired for the study approached all those providers who were originally involved in the ASSIST study. Most of these providers were working at the same TB clinic as before. Two providers in Sargodha had been transferred to nearby TB clinics, who upon request to the District Health Officer (in-charge of duty authorisation of all health care workers in the respective district), were authorised to work in their previous TB clinics (those originally involved in the ASSIST study) for the study duration. None of the providers at these TB clinics were able to sustain implementation of the BI for smoking cessation after the ASSIST study. Therefore, it was decided to run refresher training for the providers (described above in section 7.3.9 on page 115).

To elucidate the differences between digital recorders (audio and video) for practicability of use in a resource constrained setting (with frequent power cuts), a field consultation was undertaken in the UK, with a research team conducting audiotaping of patient-provider interactions. Audio-taping with a particular recorder (Zoom H2n, see 'Study equipment' page 114) was found to have advantages over other types of audio recorders, as well as videotaping. Firstly, the battery life of the video recorder was insufficient for practical use in taping an average intervention session, and secondly, there were issues of adequate room size and placement of the video recorder at an angle to capture the patient and the provider; involving complex supplementary equipment and expert help in setting it up. This could install technical challenges in using the recording equipment by the providers. In contrast, audio recorders were found to be practical for observing long and multiple sessions, as battery time was extensive (up to two days for recording an average number of moderately long sessions). Also, there were no issues of room size and placement of equipment.

7.4.2 Description of data and participants

Nineteen of the 22 approached TB clinics (and providers) agreed to participate in the study and 180 patients were enrolled to receive BI for smoking cessation (see Figure 7.3).

Fidelity data was coded for 154 audiotapes (giving 462 data-points) that were considered eligible by 3 coders. The number of audiotapes per clinic ranged between four and 10; seven clinics completed 10 sessions, five completed nine sessions, three completed eight sessions and the rest completed six, five and four sessions, respectively.

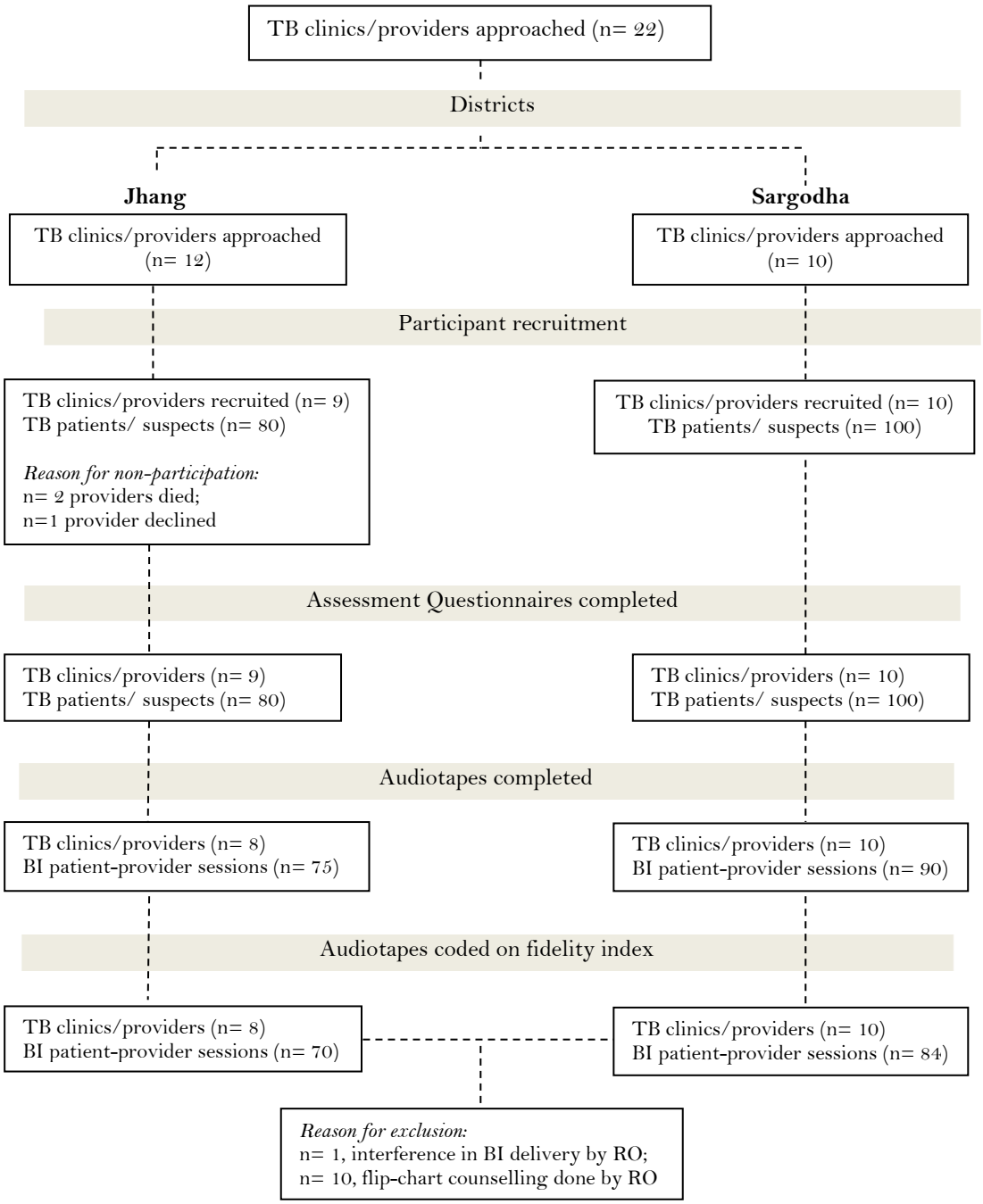


Figure 7.3: Study Flow Chart

7.4.3 Descriptive statistics for individual item and composite index scores

Individual item scores

Items 10 and 12 did not have means close to the ends of the distribution (i.e. no floor or ceiling effects) but they had very little variance (refer to Table 7.3, on page 128). Floor effect, which means very close to the lowest value, and could potentially be one possible explanation for observed low variance was seen for items 39, 40 and 44.

Note: Item 10 was 'assessing the current level of motivation to stop/ willingness to quit'; item 12 was 'Eliciting a prompt commitment from the smoker on starting the quit attempt'; item 39 was 'explaining expectations regarding the intervention programme'; item 40 was 'providing reassurance'; and item 44 was 'tailoring interactions appropriately'. For a full list of items in the fidelity index, see Appendix C.3.

Composite index scores

The Pearson's *r* value ranged from 0.88 to 0.98 for all four coders (including consensus scores) for '*Adherence*' and from 0.85 to 0.94 for '*Quality*' (refer to Table 7.4, on page 130), for composite index scores. The average correlation between different coders was 0.95 (0.92 without consensus scores) for '*Adherence*'; and 0.90 (0.87 without consensus scores) for '*Quality*'.

The scatter plots present the positive linear trends between coders for the composite scores (Figure 7.4, on page 131).

Table 7.3: Descriptive statistics and inter-coder reliability estimates for individual item scores

ITEM	Coder 1 Score			Coder 2 Score			Coder 3 Score			Consensus Score			Krippendorff's α (95% CI)	
	Mean	Median	*SD	Mean	Median	SD	Mean	Median	SD	Mean	Median	SD	Including consensus scores	Excluding consensus scores
ADHERENCE														
Item1	1.29	1.00	0.55	1.27	1.00	0.55	1.28	1.00	0.57	1.27	1.00	0.55	.888 (.82-.95)	.868 (.80-.93)
Item2	1.35	1.00	0.70	1.23	1.00	0.78	1.24	1.00	0.78	1.27	1.00	0.79	.817 (.76-.87)	.755 (.69-.82)
Item3	1.34	1.00	0.69	1.20	1.00	0.76	1.11	1.00	0.76	1.25	1.00	0.73	.782 (.71-.84)	.715 (.64-.79)
Item4	1.37	2.00	0.73	1.27	1.00	0.79	1.19	1.00	0.79	1.27	1.00	0.78	.799 (.73-.87)	.738 (.66-.81)
Item5	1.33	1.00	0.72	1.16	1.00	0.76	1.11	1.00	0.77	1.23	1.00	0.77	.771 (.70-.83)	.707 (.63-.78)
Item6	1.39	2.00	0.73	1.30	2.00	0.80	1.20	1.00	0.77	1.26	1.00	0.81	.815 (.75-.87)	.755 (.68-.82)
Item7	1.24	1.00	0.77	1.18	1.00	0.83	1.16	1.00	0.83	1.17	1.00	0.81	.788 (.72-.85)	.719 (.64-.79)
Item8	0.99	1.00	0.82	0.97	1.00	0.86	0.95	1.00	0.85	0.97	1.00	0.86	.819 (.75-.88)	.759 (.68-.83)
Item9	0.97	1.00	0.89	0.92	1.00	0.89	0.88	1.00	0.87	0.92	1.00	0.90	.893 (.85-.93)	.861 (.81-.90)
Item10	1.92	2.00	0.39	1.88	2.00	0.47	1.86	2.00	0.52	1.90	2.00	0.45	.751 (.43-1.00)	.675 (.34-.91)
Item11	1.41	2.00	0.73	1.37	2.00	0.75	1.37	2.00	0.71	1.38	2.00	0.72	.825 (.76-.88)	.782 (.71-.85)
Item12	1.89	2.00	0.45	1.82	2.00	0.58	1.82	2.00	0.58	1.87	2.00	0.49	.625 (.28-.93)	.519 (.18-.79)
Item13	1.74	2.00	0.61	1.77	2.00	0.57	1.79	2.00	0.55	1.78	2.00	0.56	.789 (.67-.89)	.734 (.61-.85)
Item14	1.35	2.00	0.86	1.37	2.00	0.85	1.44	2.00	0.80	1.44	2.00	0.78	.781 (.69-.86)	.729 (.63-.83)
Item15	1.74	2.00	0.62	1.77	2.00	0.59	1.79	2.00	0.57	1.77	2.00	0.59	.914 (.82-.98)	.885 (.78-.98)
Item16	0.75	0.00	0.97	0.75	0.00	0.97	0.67	0.00	0.95	0.69	0.00	0.95	.818 (.70-.94)	.774 (.63-.89)
Item17	0.83	0.00	0.99	0.83	0.00	0.99	0.81	0.00	0.98	0.86	0.00	0.99	.891 (.79-.98)	.856 (.73-.96)
Item18	0.77	0.00	0.98	0.70	0.00	0.96	0.69	0.00	0.95	0.71	0.00	0.96	.891 (.78-.98)	.858 (.74-.96)
Item19	0.77	0.00	0.98	0.79	0.00	0.98	0.77	0.00	0.98	0.82	0.00	0.99	.911 (.83-.98)	.890 (.79-.98)
Item20	0.76	0.00	0.97	0.81	0.00	0.98	0.75	0.00	0.97	0.81	0.00	0.98	.884 (.77-.96)	.862 (.75-.96)
Item21	1.12	1.00	0.74	1.01	1.00	0.71	0.94	1.00	0.72	0.95	1.00	0.73	.735 (.64-.82)	.678 (.58-.77)
Item22	0.97	1.00	0.76	0.88	1.00	0.76	0.70	1.00	0.76	0.76	1.00	0.76	.718 (.63-.80)	.640 (.54-.73)
Item23	0.94	1.00	0.73	0.85	1.00	0.70	0.75	1.00	0.68	0.78	1.00	0.68	.751 (.67-.82)	.681 (.59-.76)

Item24	0.99	1.00	0.72	0.91	1.00	0.71	0.70	1.00	0.66	0.77	1.00	0.67	.701 (.61-.78)	.636 (.54-.72)
Item25	0.92	1.00	0.72	0.88	1.00	0.72	0.71	1.00	0.66	0.71	1.00	0.66	.712 (.62-.79)	.645 (.54-.73)
Item26	0.84	1.00	0.74	0.81	1.00	0.71	0.64	1.00	0.65	0.66	1.00	0.66	.725 (.64-.80)	.655 (.56-.74)
Item27	0.89	1.00	0.76	0.73	1.00	0.72	0.66	1.00	0.65	0.67	1.00	0.67	.709 (.61-.80)	.638 (.53-.74)
Item28	0.83	1.00	0.75	0.73	1.00	0.71	0.56	0.00	0.64	0.57	0.50	0.62	.678 (.57-.78)	.610 (.50-.71)
Item29	0.63	0.00	0.73	0.55	0.00	0.71	0.51	0.00	0.70	0.51	0.00	0.68	.847 (.79-.91)	.812 (.75-.87)
Item30	0.58	0.00	0.70	0.51	0.00	0.67	0.42	0.00	0.62	0.47	0.00	0.65	.826 (.76-.89)	.776 (.70-.85)
Item31	0.61	0.50	0.68	0.57	0.00	0.68	0.47	0.00	0.57	0.50	0.00	0.61	.849 (.79-.91)	.810 (.74-.87)
Item32	0.60	0.00	0.71	0.51	0.00	0.68	0.41	0.00	0.63	0.47	0.00	0.65	.815 (.74-.88)	.760 (.67-.84)
Item33	0.62	0.00	0.70	0.62	0.00	0.72	0.45	0.00	0.64	0.53	0.00	0.67	.796 (.72-.87)	.740 (.66-.82)
Item34	0.60	0.00	0.67	0.56	0.00	0.72	0.49	0.00	0.64	0.49	0.00	0.64	.796 (.72-.86)	.749 (.67-.83)
Item35	0.55	0.00	0.65	0.51	0.00	0.67	0.45	0.00	0.65	0.43	0.00	0.61	.773 (.69-.85)	.720 (.64-.80)
Item36	0.50	0.00	0.62	0.44	0.00	0.64	0.40	0.00	0.60	0.40	0.00	0.58	.837 (.77-.90)	.802 (.73-.87)
Item37	0.44	0.00	0.83	0.43	0.00	0.82	0.39	0.00	0.79	0.41	0.00	0.81	.909 (.82-.98)	.879 (.76-.97)
QUALITY														
Item38	0.50	0.00	0.71	0.37	0.00	0.59	0.40	0.00	0.64	0.38	0.00	0.62	.709 (.60-.81)	.632 (.52-.74)
Item39	0.05	0.00	0.21	0.02	0.00	0.14	0.03	0.00	0.24	0.01	0.00	0.11	.157 (-.54-.72)	.052 (-.60-.64)
Item40	0.25	0.00	0.59	0.18	0.00	0.44	0.16	0.00	0.48	0.14	0.00	0.46	.566 (.40-.72)	.481 (.31-.66)
Item41	0.36	0.00	0.67	0.43	0.00	0.70	0.39	0.00	0.70	0.34	0.00	0.64	.781 (.69-.86)	.741 (.65-.82)
Item42	0.34	0.00	0.66	0.23	0.00	0.58	0.23	0.00	0.58	0.22	0.00	0.60	.769 (.66-.87)	.737 (.63-.84)
Item43	0.48	0.00	0.69	0.34	0.00	0.58	0.38	0.00	0.62	0.36	0.00	0.60	.755 (.65-.85)	.679 (.57-.78)
Item44	0.18	0.00	0.49	0.21	0.00	0.54	0.12	0.00	0.37	0.14	0.00	0.42	.765 (.63-.88)	.729 (.58-.85)
Item45	0.47	0.00	0.67	0.51	0.00	0.73	0.42	0.00	0.65	0.43	0.00	0.65	.743 (.65-.82)	.692 (.60-.78)

*SD is standard deviation
Items highlighted in grey are those with SD < .50, showing little variance
Items highlighted in mauve show low agreement
Mean krippendorff's α for Adherence was 0.80 and for Quality was 0.66

Table 7.4: Descriptive statistics and Pearson's correlation for composite (A and Q) scores

Scores	Descriptive statistics			Pearson's *α (95% CI)							
	Mean	**Median**	****SD**	**A-Coder1**	**A-Coder2**	**A-Coder3**	**A-Consensus**	**Q-Coder1**	**Q-Coder2**	**Q-Coder3**	**Q-Consensus**
A-Coder1	37.58	39.00	18.04	-	.909 (.877-.933)	.880 (.839-.912)	.899 (.864-.926)	-	-	-	-
A-Coder2	35.91	35.50	17.47	-	-	.953 (.936-.965)	.973 (.963-.981)	-	-	-	-
A-Coder3	33.55	34.00	16.80	-	-	-	.977 (.968-.983)	-	-	-	-
A-Consensus	34.66	34.00	17.50	-	-	-	-	-	-	-	-
Q-Coder1	2.61	1.00	3.59	-	-	-	-	-	.868 (.823-.902)	.847 (.796-.887)	.873 (.829-.906)
Q-Coder2	2.29	1.00	3.31	-	-	-	-	-	-	.895 (.858-.923)	.943 (.923-.958)
Q-Coder3	2.15	0.00	3.42	-	-	-	-	-	-	-	.933 (.909-.951)
Q-Consensus	2.03	1.00	3.27	-	-	-	-	-	-	-	-

*α is Pearson correlation statistic (Fisher's z transformed)
**SD is standard deviation

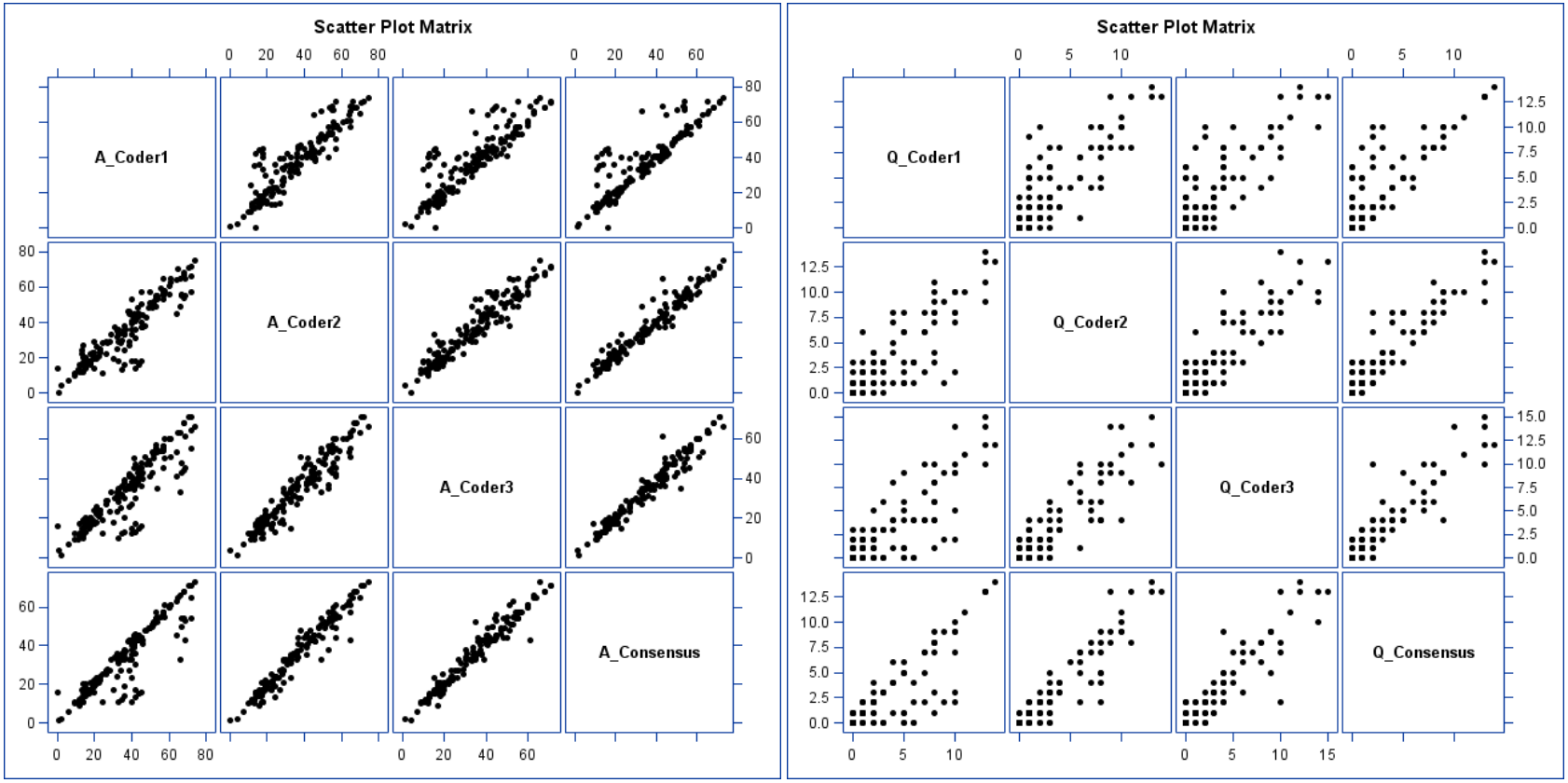


Figure 7.4: Scatterplots for Adherence and Quality

7.4.4 Inter-coder reliability

Krippendorff provides a more conservative interpretation of the α coefficient than the conventional tests of inter-coder reliability, suggesting that conclusions should be discounted for variables with values of $\alpha < 0.67$, tentatively made for α between 0.67 and 0.80, and definitely made for $\alpha > 0.80$ (Krippendorff, 2004)(p. 241). For individual items, looking at Krippendorff's α across the three independent coders, items 12, 22, 24 to 28, 38, 39 and 40 showed lower agreements, that is, $\alpha < 0.67$ (refer to Table 7.3, on page 128). The overall mean Krippendorff's α for individual items was 0.80 for '*Adherence*' and 0.66 for '*Quality*'. Therefore, the fidelity index showed moderate (for Quality) to good (for Adherence) stability in rating by different coders.

After including the consensus scores, the Krippendorff's α for items 22, 24 to 28, and 38 increased to > 0.67, showing higher agreement.

Note: Item 12 was 'Eliciting a prompt commitment from the smoker on starting the quit attempt'; item 38 was 'building general rapport'; item 39 was 'explaining expectations regarding the intervention programme'; and item 40 was 'providing reassurance'. Items 22 to 28 were about the 'identification and management of triggers'. For a full list of items of the fidelity index, see Appendix C.3.

7.4.5 Principal Component Analysis (PCA)

Item loadings for the fidelity index under a single component

Item loadings across all 45 items of the fidelity index together, showed that about 52% (Eigen value 17.8) of the variance could be explained under a single component or dimension of the index (Refer to Table 7.5). Item loading values of > 0.4 were considered meaningful in measuring a single dimension for fidelity to the BI.

Seven items (item 10, 12, 15, 16, 37, 39 and 40) showed loadings below the threshold of 0.4, consistently across the coders.

Note: Item 10 was 'assessing the current level of motivation to stop/ willingness to quit'; item 12 was 'Eliciting a prompt commitment from the smoker on starting the quit attempt'; item 15 was 'nicotine dependency'; item 16 was 'setting quit date'; item 37 was 'offering BI leaflet'; item 39 was 'explaining expectations regarding the intervention programme'. For a full list of items of the fidelity index, see Appendix C.3.

Table 7.5: Item loadings for all 45 items of the fidelity index

Individual items	Coder 1	Coder 2	Coder 3	Consensus
Item1	0.53	0.43	0.36	0.43
Item2	0.73	0.66	0.68	0.73

Item3	0.66	0.66	0.75	0.73
Item4	0.78	0.69	0.76	0.78
Item5	0.75	0.63	0.73	0.72
Item6	0.73	0.70	0.69	0.72
Item7	0.70	0.53	0.64	0.66
Item8	0.71	0.74	0.71	0.77
Item9	0.74	0.71	0.79	0.80
Item10	0.27	0.17	0.23	0.23
Item11	0.66	0.43	0.46	0.43
Item12	0.23	0.29	0.17	0.23
Item13	0.39	0.37	0.42	0.44
Item14	0.49	0.42	0.44	0.43
Item15	0.28	0.17	0.10	0.16
Item16	0.45	0.32	0.28	0.34
Item17	0.67	0.56	0.49	0.57
Item18	0.70	0.59	0.54	0.62
Item19	0.63	0.56	0.52	0.56
Item20	0.69	0.61	0.57	0.64
Item21	0.73	0.72	0.72	0.69
Item22	0.76	0.65	0.67	0.71
Item23	0.78	0.74	0.74	0.79
Item24	0.80	0.78	0.76	0.78
Item25	0.82	0.74	0.80	0.84
Item26	0.83	0.77	0.83	0.84
Item27	0.76	0.76	0.82	0.80
Item28	0.82	0.76	0.81	0.82
Item29	0.75	0.79	0.74	0.76
Item30	0.77	0.76	0.74	0.75
Item31	0.75	0.75	0.69	0.73
Item32	0.79	0.74	0.68	0.75
Item33	0.77	0.74	0.74	0.75
Item34	0.75	0.72	0.70	0.72
Item35	0.73	0.69	0.70	0.70
Item36	0.77	0.76	0.72	0.75
Item37	0.12	0.23	0.28	0.27
Item38	0.53	0.51	0.55	0.49
Item39	0.30	-0.07	0.19	0.06
Item40	0.41	0.37	0.27	0.32
Item41	0.49	0.69	0.55	0.55
Item42	0.50	0.46	0.45	0.41
Item43	0.57	0.54	0.57	0.55
Item44	0.27	0.47	0.40	0.32
Item45	0.63	0.73	0.66	0.66
Loading values $\geq$ 0.4 were taken as meaningful Proportion variance under single component = 0.52				

Item loadings for Adherence and Quality separately (single component extraction)

Item loadings across 37 items for '*Adherence*' and 8 items for '*Quality*', showed that about 60% (Eigen value 16.4) and 91% (Eigen value 4.3) of the variance was explained by a single component under each sub-index, respectively (refer to Table 7.6 and Table 7.7).

Five items (item 10, 12, 15, 16, and 37) for '*Adherence*' showed loadings below the threshold of 0.4. For the '*Quality*' item 39 did not load well.

Note: Item 10 was 'assessing the current level of motivation to stop/ willingness to quit'; item 12 was 'Eliciting a prompt commitment from the smoker on starting the quit attempt'; item 15 was 'nicotine dependency'; item 16 was 'setting quit date'; item 37 was 'offering BI leaflet'; item 39 was 'explaining expectations regarding the intervention programme'. For a full list of items of the fidelity index, see Appendix C.3.

Table 7.6: Item loadings of the 37 items for Adherence

Individual items	Coder 1	Coder 2	Coder 3	Consensus
Item1	0.53	0.42	0.36	0.43
Item2	0.73	0.65	0.69	0.73
Item3	0.66	0.65	0.74	0.71
Item4	0.77	0.69	0.75	0.77
Item5	0.74	0.62	0.73	0.71
Item6	0.72	0.70	0.70	0.72
Item7	0.70	0.54	0.66	0.67
Item8	0.74	0.74	0.71	0.77
Item9	0.73	0.70	0.78	0.79
Item10	0.27	0.18	0.24	0.23
Item11	0.66	0.44	0.47	0.45
Item12	0.27	0.29	0.18	0.22
Item13	0.39	0.38	0.43	0.44
Item14	0.52	0.44	0.47	0.46
Item15	0.27	0.18	0.12	0.17
Item16	0.47	0.35	0.32	0.38
Item17	0.70	0.61	0.54	0.61
Item18	0.74	0.65	0.60	0.66
Item19	0.67	0.61	0.59	0.61
Item20	0.72	0.66	0.62	0.67
Item21	0.71	0.70	0.70	0.67
Item22	0.76	0.65	0.68	0.71
Item23	0.78	0.73	0.74	0.79
Item24	0.80	0.76	0.75	0.78
Item25	0.82	0.74	0.80	0.84
Item26	0.82	0.77	0.83	0.84
Item27	0.76	0.75	0.82	0.80
Item28	0.81	0.76	0.80	0.82
Item29	0.76	0.79	0.74	0.76
Item30	0.78	0.76	0.73	0.75
Item31	0.75	0.77	0.70	0.74

Item32	0.78	0.75	0.68	0.76
Item33	0.77	0.75	0.74	0.75
Item34	0.77	0.72	0.70	0.71
Item35	0.73	0.71	0.72	0.72
Item36	0.78	0.78	0.73	0.76
Item37	0.06	0.19	0.24	0.22
Loading values $\geq$ 0.4 were taken as meaningful Proportion variance under single component = 0.60				

Table 7.7: Item loadings of the 8 items for Quality

Individual items	Coder 1	Coder 2	Coder 3	Consensus
Item38	0.66	0.74	0.83	0.76
Item39	0.33	-0.01	0.46	0.17
Item40	0.69	0.51	0.64	0.69
Item41	0.85	0.86	0.90	0.90
Item42	0.87	0.83	0.81	0.85
Item43	0.77	0.75	0.78	0.81
Item44	0.66	0.80	0.70	0.74
Item45	0.74	0.76	0.74	0.68
Loading values $\geq$ 0.4 were taken as meaningful Proportion variance under single component = 0.91				

Item loadings for Adherence (multiple components extraction)

Further exploration of components with Eigenvalues more than 1 for *Adherence* items, identified a total of five components (refer to Table 7.8, on page 137). At least one of the components explained 60%, the second: 11% and the third: 9% of the variance or information contained in the BI content, which formed the items of the Adherence sub-index (Table 7.9, on page 138). The five components extracted for the Adherence sub-index emerged showing the same structure across all different coders.

Items 1 and 37 did not load well on any of the five components extracted.

The five extracted components were (further discussed in section 7.5.3, page 146):

- Component 1 included items 29 to 36, which are about 'assessing and managing the nicotine withdrawal symptoms' in the BI.
- Component 2 included items 2 to 9, which are about 'providing information on the consequences of smoking and stopping to smoke' in the BI.
- Component 3 included items 21 to 28, which are about 'identifying and managing the triggers to smoke' in the BI.
- Component 4 included items 16 to 20, which comprised of 'preparing for the quit attempt' in the BI.
- Component 5 included items 10 to 15, which relate to the 'assessments of willingness to quit, nicotine dependency and quit attempt history' in the BI.

Note: Item 1 was 'smoking behaviour and history' and item 37 was 'offering the BI leaflet'. For a full list of items of the fidelity index, see Appendix C.3.

The results from the PCA indicate that item loadings for the overall index and for the two sub-indices explained considerable variance under a single component. This showed that majority of the items were measuring the same construct. The items that did not load well on PCA are discussed in the next section (7.5.3).

Table 7.8: Identifying the Principal Components for Adherence

Individual items	Loading 1				Loading 2				Loading 3				Loading 4				Loading 5			
	C 1	**C 2**	**C 3**	**C 4**	**C 1**	**C 2**	**C 3**	**C 4**	**C 1**	**C 2**	**C 3**	**C 4**	**C 1**	**C 2**	**C 3**	**C 4**	**C 1**	**C 2**	**C 3**	**C 4**
Item1	-0.02	-0.11	-0.02	-0.09	0.31	0.12	0.15	0.06	-0.1	0.3	-0.04	0.18	0.36	0.23	0.26	0.38	0.23	0.12	0.36	0.15
Item2	-0.03	0	0.04	0.08	0.28	0.53	0.71	0.67	0.57	0.16	0.05	0.03	0.11	0.14	0.09	0.1	0.02	0.05	-0.06	0.07
Item3	-0.15	-0.03	0.01	-0.02	0.28	0.76	0.84	0.76	0.59	-0.04	0.08	0.12	0.06	0.06	-0.08	-0.07	0.11	0.17	0.07	0.16
Item4	0.01	-0.04	0.04	0.01	0.33	0.82	0.87	0.81	0.68	0	0	0.09	-0.1	0.1	-0.05	0.02	0.11	0.01	0.12	0.04
Item5	-0.04	-0.03	-0.07	-0.02	0.37	0.77	0.85	0.81	0.66	-0.03	0.13	0.12	-0.04	0.09	-0.02	-0.01	-0.04	-0.01	0	-0.06
Item6	0.03	0.04	0.03	0.05	0.3	0.94	0.88	0.93	0.72	0.02	-0.05	-0.04	-0.1	-0.06	-0.01	-0.03	-0.07	-0.06	0	-0.1
Item7	-0.07	-0.09	-0.08	-0.12	0.28	0.86	0.83	0.93	0.73	0.04	0.05	0	-0.02	-0.07	0.05	0.04	-0.04	-0.06	-0.06	-0.04
Item8	0.34	0.31	0.23	0.22	-0.13	0.46	0.52	0.52	0.56	0.11	0.08	0.09	0.21	0.03	0.07	0.14	-0.02	0.06	-0.04	-0.05
Item9	0.22	0.19	0.17	0.16	-0.09	0.49	0.62	0.54	0.72	0.18	0.16	0.24	0.17	0.07	0.06	0.06	-0.11	-0.07	-0.14	-0.08
Item10	0.02	-0.1	0.09	0	0.04	-0.05	0.04	0.03	-0.01	0.02	-0.03	0	-0.12	0.04	-0.1	-0.04	0.69	0.69	0.68	0.74
Item11	-0.01	0.04	0	0.06	0.23	0.37	0.37	0.38	0.39	0.07	-0.01	-0.11	0.01	-0.12	0.07	0.1	0.38	0.43	0.39	0.34
Item12	0	-0.02	0.08	0.01	0.04	-0.02	-0.07	-0.08	-0.07	-0.06	0.04	0.13	-0.01	0.16	-0.13	-0.1	0.64	0.63	0.65	0.74
Item13	0.07	0.17	-0.04	0.07	-0.17	0.17	0.23	0.26	0.25	-0.11	0.12	0	0.01	-0.06	0.03	0	0.61	0.58	0.43	0.51
Item14	0.08	0.12	-0.13	-0.03	-0.16	0.06	0.28	0.32	0.32	0.15	0.15	-0.06	0.16	-0.05	0.15	0.2	0.5	0.6	0.39	0.43
Item15	-0.01	-0.1	-0.02	-0.01	0.18	-0.34	-0.32	-0.31	-0.35	0.37	0.03	0.03	0.25	0.16	0.3	0.34	0.47	0.39	0.53	0.48
Item16	0.01	-0.06	-0.03	-0.05	-0.07	0.1	-0.17	0.01	-0.03	0.62	0.17	-0.04	0.66	-0.19	0.52	0.63	0.15	0.16	0.02	0.03
Item17	0.09	0.14	0.05	0.07	0.01	-0.02	0	0.04	0.04	0.75	-0.05	-0.06	0.84	0.02	0.85	0.82	-0.03	-0.02	0.02	0.01
Item18	0.09	0.26	0.09	0.17	0.09	-0.05	0.03	-0.07	-0.03	0.82	-0.06	-0.01	0.89	-0.07	0.9	0.88	-0.05	-0.06	-0.04	-0.06
Item19	0.02	0.01	-0.07	-0.04	-0.05	0.09	0.11	0.1	0.12	0.84	0	-0.04	0.85	-0.01	0.9	0.86	-0.01	-0.04	-0.04	0.01
Item20	0.04	-0.1	0.06	-0.03	0.07	0.16	0.15	0.12	0.08	0.84	-0.04	0.07	0.84	0.1	0.77	0.8	-0.07	-0.05	-0.01	-0.05
Item21	0.19	0.25	0.09	0.06	0.63	0.2	0.05	0.04	0.16	0	0.68	0.75	-0.07	0.45	0	-0.05	-0.09	-0.1	-0.02	-0.06
Item22	-0.09	0.01	0.02	-0.03	0.67	0.06	-0.01	0	0.16	0.17	0.75	0.76	0.19	0.6	0.08	0.12	0.04	0.02	-0.07	-0.01
Item23	-0.07	0.08	-0.02	-0.01	0.83	0.15	-0.02	0.15	0.12	-0.06	0.84	0.81	0.07	0.69	0.03	-0.04	0	0.07	0.09	0.06
Item24	0	0.23	0.06	0.09	0.88	0.15	0.09	0.03	0.04	-0.12	0.82	0.81	0.03	0.66	-0.13	-0.02	0.04	-0.02	0.03	0.02
Item25	0.16	0.3	0.08	0.2	0.82	-0.01	0.09	0.09	0.05	-0.06	0.8	0.76	-0.06	0.65	-0.01	-0.06	0.04	-0.01	-0.04	0
Item26	0.09	0.22	0.06	0.15	0.77	0.09	0.09	0.05	0.05	-0.03	0.82	0.77	0.08	0.65	0	0.03	0	0.02	0.01	0.01
Item27	0.13	0.19	0.09	0.12	0.78	0.17	0.07	0.13	0.07	0.05	0.72	0.66	-0.04	0.56	0.13	0.09	-0.04	-0.06	-0.03	-0.05

Item28	0.12	0.23	0.16	0.16	0.88	-0.01	0.07	0.05	0.05	0.09	0.65	0.64	-0.06	0.59	0.07	0.14	-0.04	0.08	0.04	0.04
Item29	0.8	0.87	0.84	0.85	0.01	0.03	-0.05	-0.04	0.05	0.01	0.08	0.08	0.08	0.02	0.03	0.01	0.01	-0.01	0.04	0.02
Item30	0.88	0.82	0.86	0.9	0.05	0.05	0.09	0.05	0	-0.05	0.02	-0.04	0.02	0.06	-0.08	-0.01	0.02	-0.02	-0.04	-0.02
Item31	0.92	0.96	0.84	0.9	0.03	0.01	0.1	0.12	0.02	-0.05	-0.04	-0.1	-0.07	-0.05	-0.02	-0.03	0.05	0.01	-0.03	-0.03
Item32	0.89	0.9	0.81	0.92	0.02	-0.07	-0.04	-0.02	-0.01	0.04	0.01	-0.06	0.05	-0.06	0.05	0.03	0.04	0.15	0.06	0.1
Item33	0.9	1.01	0.87	0.96	0	0.11	0.07	0.11	0.13	-0.12	0.05	-0.03	-0.04	-0.15	-0.09	-0.14	-0.06	-0.04	-0.04	-0.02
Item34	0.94	0.85	0.9	0.91	0.09	-0.09	-0.03	-0.1	-0.08	0.04	0.03	0.1	-0.03	0.04	-0.04	-0.05	0.03	-0.04	-0.02	-0.03
Item35	0.83	0.83	0.8	0.8	0.09	-0.03	-0.06	-0.03	-0.07	0.06	0.05	0.05	0.05	-0.02	0.08	0.05	0	-0.01	0.08	0
Item36	0.79	0.74	0.73	0.74	0.07	-0.12	-0.01	-0.05	-0.03	0.25	-0.01	0.02	0.15	0.06	0.19	0.2	0.01	0.02	0.09	0.06
Item37	-0.02	-0.21	-0.05	-0.15	0.03	-0.07	0.06	-0.03	0.02	-0.04	0.29	0.49	0	0.55	-0.08	-0.09	0.07	0.03	0.11	0.08

C= Coder
Colour codes: Component 1=mauve, Component 2=grey, Component 3=blue, Component 4= green, Component 5= yellow
Loading values $\geq$ 0.4 were considered meaningful

Table 7.9: Variance explained by each Principal Component for Adherence

Principal components	Coder 1		Coder 2		Coder 3		Consensus	
	Eigenvalue	**Proportion variance**	**Eigenvalue**	**Proportion variance**	**Eigenvalue**	**Proportion variance**	**Eigenvalue**	**Proportion variance**
Component 1	17.37	0.61	14.99	0.58	15.37	0.58	16.41	**0.60**
Component 2	3.47	0.12	2.83	0.11	2.86	0.11	2.89	**0.11**
Component 3	2.28	0.08	2.59	0.10	2.61	0.10	2.52	**0.09**
Component 4	1.53	0.05	1.64	0.06	1.79	0.07	1.66	0.06
Component 5	1.21	0.04	1.13	0.04	1.16	0.04	1.28	0.05

7.4.6 Generalisability study (G-study)

Identification of important facets (sources) of variation

Since the fidelity index was used to measure provider Adherence to BI content and Quality of interaction, independent of patient- and coder-variance, I was interested in the share of variance that is due to true differences between providers i.e. ν_h.

The sources of variation identified for the G-Study that needed accounting for, in order to obtain the true differences between providers, were calculated: σ^2_r, σ^2_h, $\sigma^2_{i:h}$, σ^2_{rh}, $\sigma^2_{ri:h.}$, as described in Table 7.10.

Where,

- A number of audiotaped sessions *(n = 154; i)* were obtained within a number of providers *(n = 18; h)*
- The audiotaped sessions were nested within providers *(i:h)*.
- Each session was rated by three independent coders and a fourth consensus score, which makes this a *r*(i:h)* design (Brennan, 2001)(p.56).

Table 7.10: Identification of important sources of variation

Source of variability	Type of variation	Variation component
Provider/TB clinic *(h)*	Variance due to the differences in providers' actual fidelity to BI	σ^2_h
Coder *(r)*	Variance due to the coders' differences (coder leniency) in scoring	$\sigma^2_r/4$
Coder * Provider *(rh)*	Variance due to inconsistencies in different coders' scoring of the same provider, averaging over patients (audiotapes).	$\sigma^2_{rh}/4$
Patient: Provider *(i:h)*	Variance due to individual differences of patients' (audiotapes) nested within providers	$\sigma^2_{i:h}/154$
Coder *Patient: Provider *(ri:h)*	The residual or random error variance (i.e. any unanalysed facets of measurement that varied among providers)	$\sigma^2_{res}/(4*154)$

Variance partitioning

The variance partitioning yielded variation contributed by each of the σ^2_r, σ^2_h, $\sigma^2_{i:h}$, σ^2_{rh}, $\sigma^2_{ri:h}$, and the total variance presented in Table 7.11 for *Adherence* and Table 7.12 for *Quality*.

Table 7.11: Variance partitioning for Adherence

Source of variability	Variation component	Variance %
Provider/TB clinic *(h)*	σ_h^2= 235.01	98.5
Coder *(r)*	σ_r^2/4= 1.97/4= 0.49	0.20
Coder * Provider *(rh)*	σ_{rh}^2/4= 11.65/4= 2.91	1.22
Patient: Provider *(i:h)*	$\sigma_{i:h}^2$/154= 8.64/154= 0.056	0.02
Coder * patient: Provider *(ri:h)*	σ_{res}^2/(4*154)= 9.50/4*154= 0.015	0.006
Total	238.48	

Table 7.12: Variance partitioning for Quality

Source of variability	Variation component	Variance %
Provider/TB clinic *(h)*	σ_h^2= 8.27	98.6
Coder *(r)*	σ_r^2/4= 0.025/4= 0.006	0.07
Coder * Provider *(rh)*	σ_{rh}^2/4= 0.446/4= 0.111	1.32
Patient: Provider *(i:h)*	$\sigma_{i:h}^2$/154= 0.031/154= 0.0002	0.002
Coder * patient: Provider *(ri:h)*	σ_{res}^2/(4*154)= 0.838/4*154= 0.001	0.01
Total	8.39	

Computation of g-coefficient

Once the variance components were calculated, these were computed to get the *g*-coefficient of reliability, using the following formula (Brennan, 2001):

$$g = \frac{\sigma^2(h)}{\sigma^2(h)+\sigma^2(r)+\sigma^2(i:h)+\sigma^2(rh)+\sigma^2(ri:h)}$$

For Adherence:

$$g = \frac{235.01}{235.01 + 1.97/4 + 8.64/154 + 11.64/4 + 9.49/4*154} = \mathbf{0.985}$$

For Quality:

$$g = \frac{8.27}{8.27 + 0.025/4 + 0.030/154 + 0.446/4 + 0.838/4*154} = \mathbf{0.986}$$

The results of the G-study indicate that the fidelity index reliably differentiates between providers based on their Adherence to BI content and the Quality of interaction. Percentage variance contributed by the providers was significantly high (98.5 for Adherence and 98.6 for Quality) accounting for other sources of variation.

Sensitivity analyses for the g-coefficient

Crude intra-cluster correlation (ICC) of provider differences

For Adherence:

σ_h^2= 248.99, Residual= 98.27 (determined from variance partitioning)

ICC= $\sigma_h^2/\sigma_h^2 + \sigma_{res}^2$ (formula for calculating ICC)

$$\text{ICC} = \frac{248.99}{248.99 + 98.27} = \mathbf{0.72}$$

For Quality:

σ_h^2= 8.62, Residual= 3.56 (determined from variance partitioning)

ICC= $\sigma_h^2/\sigma_h^2 + \sigma_{res}^2$ (formula for calculating ICC)

$$\text{ICC} = \frac{8.62}{8.62 + 3.56} = \mathbf{0.71}$$

Note: These were calculated after excluding the consensus scores; therefore the numbers differ slightly from variance partitioning presented in Table 7.11 and 7.12 above.

These results indicate that 72% (for Adherence) and 71% (for Quality) of the variation was contributed by providers in fidelity measurement, when other sources were not accounted for in the model.

G-study excluding consensus scores

For *Adherence*:

$$g = \frac{227.14}{227.14 + 2.77/3 + 9.41/154 + 15.34/3 + 11.24/3*154} = \mathbf{0.973}$$

For *Quality*:

$$g = \frac{8.126}{8.126 + 0.008/3 + 2.736/154 + 0.578/3 + 1.008/3*154} = \mathbf{0.974}$$

G-study excluding providers with very low fidelity

Further exploration of the variation in A and Q scores by providers (TB clinics) is presented in Table 7.13 and Figure 7.5. Based on these statistics, those providers having very low or zero scores were excluded from primary analysis for this sensitivity analysis.

For A scores, the providers having a minimum score of 10 or less on any single session were excluded from the analysis (i.e. providers 4, 7, 24, 26 and 27). For the Q scores, most providers had 0 as the minimum score; therefore, selection was restricted to a maximum score of 5 or less out of the total of 16, to qualify for exclusion from the sensitivity analysis (i.e. providers 4, 7, 21, 22, 24, 26, 27, and 28). G-study estimates after excluding these providers were as follows.

For Adherence:

$$g = \frac{109.02}{109.02 + 2.008/4 + 3.07/154 + 6.88/4 + 9.71/4*154} = \mathbf{0.979}$$

For Quality:

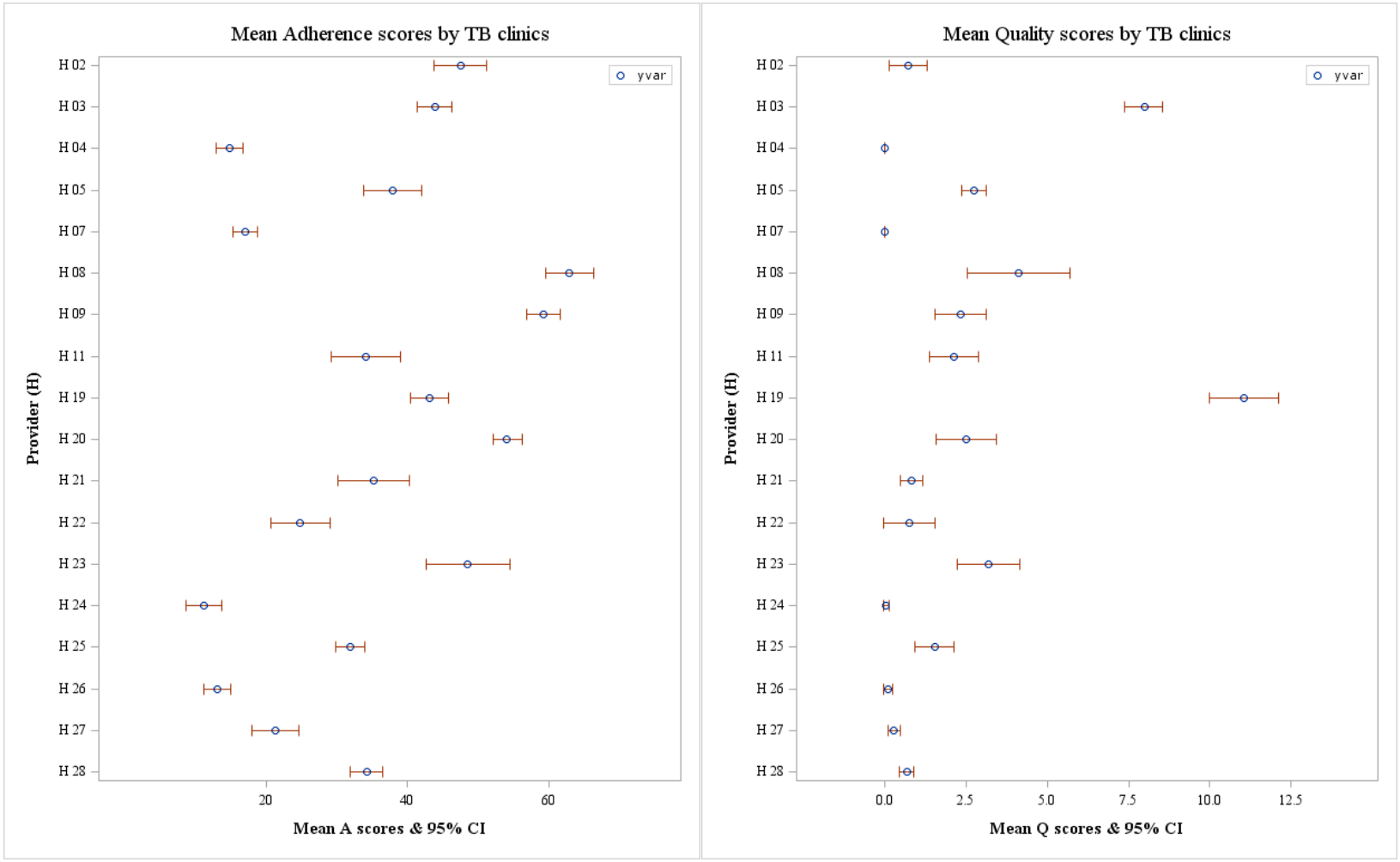
$$g = \frac{9.67}{9.67 + 0.0468/4 + 0.014/154 + 0.785/4 + 1.228/4*154} = \mathbf{0.979}$$

Table 7.13: Descriptive statistics of the A and Q scores by providers

Provider (TB clinic)	Adherence score			Quality score		
	Mean (95% CI)	**Min**	**Max**	**Mean (95% CI)**	**Min**	**Max**
2	47.50 (43.79-51.21)	33	72	0.72 (0.14-1.31)	0	8
3	43.85 (41.40-46.30)	29	61	7.98 (7.39-8.56)	4	11
4	14.83 (12.87- 16.79)	0	22	0	0	0
5	37.86 (33.74-41.98)	15	55	2.75 (2.38-3.12)	1	6
7	17.06 (15.26-18.85)	9	29	0	0	0
8	62.94 (59.51-66.36)	51	71	4.13 (2.55-5.70)	1	11
9	59.23 (56.91-61.54)	43	72	2.33 (1.53-3.12)	0	10
11	34.13 (29.20-39.05)	14	75	2.13 (1.36-2.89)	0	8
19	43.06 (40.36-45.77)	29	57	11.06 (10.00-12.13)	5	15
20	54.02 (52.11-56.29)	38	70	2.50 (1.57-3.43)	0	10
21	35.16 (30.12-40.20)	14	61	0.81 (0.48-1.15)	0	3
22	24.85 (20.72-28.98)	13	69	0.75 (-0.05-1.55)	0	10
23	48.56 (42.64-54.48)	11	72	3.19 (2.22-4.15)	0	10
24	11.17 (8.62-13.71)	1	21	0.04 (-0.04-0.13)	0	1
25	31.85 (29.80-33.90)	15	42	1.53 (0.93-2.12)	0	6
26	13.05 (11.18-14.92)	0	20	0.10 (-0.04-0.24)	0	1
27	21.33 (17.96-24.69)	10	45	0.28 (0.08-0.47)	0	2
28	34.19 (31.88-36.51)	17	46	0.67 (0.45-0.88)	0	2

Note: Grey shaded statistics show lower fidelity

Figure 7.5: Mean A and Q scores by providers



7.5 DISCUSSION

7.5.1 Summary of findings

The three types of reliability tests performed on the fidelity index found that it was a reliable tool. The inter-coder reliability showed that the fidelity index was a stable tool that could be reliably scored by coders with varying skills and expertise. The Krippendorff's alpha was 0.80 for *Adherence*, which is considered good and 0.66 for *Quality*, which is considered moderately acceptable. The items were coherent (as assessed by PCA) in measuring fidelity, overall and separately, for both *Adherence* and *Quality*. Item loadings from PCA across all 45 items showed that 52% of the variance was explained under a single construct (i.e. fidelity); 60% under *Adherence* for 37 items and 91% under *Quality* for eight items. This showed that the majority of the items were measuring a common construct. Seven items (1, 10, 12, 15, 16, 37 and 39) did not load well on single component and multi component extractions for the sub-indices on PCA. The G-study indicated that the index was highly reliable in differentiating between providers, based on their fidelity. The ICC estimates for the providers showed that more than 70% of the variation was contributed by the differences between providers' practice of delivering the BI.

Findings from each type of reliability testing are separately discussed below.

7.5.2 Inter-coder reliability

Items 12 (eliciting a prompt commitment to quit), 39 (explaining expectations regarding treatment) and 40 (providing reassurance) had an alpha score of $\alpha < 0.67$ and also showed little observed variance in univariate analysis. Low variance alone showed that there was very little variation in the item and that the evaluated providers did not really differ.

In addition to little observed variance, items 39 and 40 also showed a floor effect (very close to the lowest value). Floor and ceiling effects are indications that the item was not separating between the objects that were evaluated; all were good or bad. However, there is the potential that asking for less extreme (or more extreme) response alternatives would result in variation. That is, this could be an indication that the items were just "mis-calibrated" for the sample (too easy/ too difficult). Another alternative explanation for little observed variance with or without floor/ceiling effects could be that the item was not really relevant.

When data was explored for causes of the floor effect, it was found that all the providers scored very low on these items and the coders agreed on this (coded '0 for not implemented' in case of items 39 and 40; and '2= fully implemented' in case of item 12). Therefore, it could be interpreted for items 39 and 40 that most providers did not adhere to these ingredients of the BI. It is worthwhile noting here that the lower Kripendorff's alpha agreement between coders for these three items was because of the lack of variation (due to data being rated similarly), and not due to

the lack of agreement between the coders. Hence, these should not be excluded from the index based on the results of the inter-coder reliability, but should be explored in future analysis.

In addition, items 22 to 28 and item 38 (relating to identification and management of the triggers) showed lower agreement (Krippendorff's α ranging from 0.61 to 0.66) between the three coders. However, these items did not have little observed variance or floor/ceiling effects. When consensus scores were included as the fourth coder in analysis, these items showed higher agreement ($\alpha > 0.67$), indicating that they had the potential of being particularly amenable to consensus exercise (described in 7.3.4, page 116). This raises the question of whether a consensus round of coding fidelity indices has an added value. As there is discrepancy in agreement estimates with and without consensus scoring, I believe that these scores should not be included in any kind of psychometric assessments of an instrument (fidelity index). However, these scores might still have value for use in regression analysis to explore the relationship of fidelity with quit rates (study C). An intervention evaluation study conducting fidelity measurement, where the focus is not on assessing psychometric properties of the measure, could benefit from consensus scores, as these are highly likely to be more accurate (based on judgements of three coders), compared to scoring by a single coder.

This might be because the consensus was a process of convergence between the coders after resolving disagreements on independent ratings and when included in analysis showed higher agreement. It might also be that these items were more subjective, as they required tailoring to patient needs and were therefore more difficult to judge by the coders. Alternatively, these items might have been reflective of more than one underlying construct of the fidelity index. The rest of the items showed consistent results after excluding the consensus scores.

7.5.3 Principal Component Analysis

Item loadings for the fidelity index under a single component

Items (10, 12, 15, 16, 37, 39 and 40) did not load well overall for the fidelity index (Table 7.5) and could be explained as:

- Those items that showed little variance between the providers as these were 'not implemented' by the majority of them in this study. These (items 16, 37, 39 and 40) were rated consistently as 0= 'not implemented' by all coders. Item 16 was about setting the quit date, 37 was offering the BI leaflet, 39 and 40 were about explaining treatment expectations and providing reassurance. Out of the 154 audiotaped sessions, the numbers rated as '0' (not implemented) for items 16, 37, 39 and 40 were 101, 122, 152 and 139, respectively. It could not, therefore, be concluded that these items were coherent or not with the rest of the items in the fidelity index.

- Items not loading well in spite of being implemented by majority of the providers. These were items 10, 12 and 15, relating to assessments of patient motivation/willingness to quit and nicotine dependency. These items might not be relevant to fidelity or these might indicate that the providers did not differ in their practice based on these items; that is, they implemented them in a standard way. These items seem to be describing something different from the rest of the 42 items of the fidelity index and would be candidates for dropping from the index.

Note: For a full list of items of the fidelity index, see Appendix C.3.

Item loadings for Adherence and Quality separately (single component extraction)

Again, items 16, 37 and 39 could be explained not to have loaded well on a single construct due to being rated consistently across coders (Table 7.6) and items 10, 12 and 15 seem to be explaining something different from the rest of the 34 items in the *Adherence* sub-index.

One peculiar finding from the item loadings, when performed separately for the two sub-indices, was that item 40 'providing reassurance' (which did not load on the overall fidelity index) loaded well when extracted separately for the eight items under the *Quality of interaction.* The little data available on 15 audiotapes out of 139 consistently entered '0's (not implemented) across the coders, showed that this item measured an aspect of *Quality of interaction* similar to the rest of the seven items in this sub-index. However, this item does not have much in common with the Adherence set of items and since those items are in the majority in the full index, which could be why the loadings of item 40 in the overall index analysis were lowered.

Item loadings for Adherence (multiple components extraction)

The five components (Table 7.8) extracted for *Adherence* could be matched roughly to five of the seven key features of the BI (see section 6.3, page 78). These BI features were 'information about harms of smoking/benefits of quitting'; 'assessments of dependency'; 'preparation to quit'; 'management of triggers'; and of 'withdrawals'. The rest of the two features concern 'Ascertaining about tobacco use' and 'Offering the BI leaflet' represented by items 1 and 37.

Item 1, although implemented by the providers, did not load well on any component and this might be explained by looking at its content within the BI. This item related to the 'smoking behaviour and history', which is part of a very general BCT concerning 'information gathering and assessment'. It did not seem to measure the same thing as the rest of the items under *Adherence* and it did not explain enough variation within the sub-index to qualify as a separate component. Item 1 would be a candidate for dropping from the fidelity index on this basis.

Item 37 'offering BI leaflet', as discussed above, was rated as '0' consistently by all coders, as it was not implemented by the providers and therefore did not have enough variance (or

information) to qualify as a separate component in the Adherence sub-index. On this basis, item 37 should be kept in the fidelity index for determining item loading in other settings and overall index generalisability in future studies.

Note: For a full list of items of the fidelity index, see Appendix C.3.

7.5.4 G-study

The crude ICC (0.70) showed less variation being explained by the providers than the variance estimates (*Adherence* = 0.985 and *Quality* = 0.986) from the G-study. This could mean that accounting for other sources of variance in the G-study allowed for a better estimate of each source's contribution to the overall variation, whereas in the ICC the rest of the 30% variation was residual and could not be partitioned.

7.5.5 Refinement of the fidelity index

The focus of my study was not to produce a more parsimonious index but was to screen the items and describe their stability. However, some items did not load well on the PCA and these results can be used to eliminate redundancy in the items of the index. For index refinement, using psychometric findings, two things are considered important: items that were easily agreed upon by all coders and that also represented Adherence/Quality dimensions of the fidelity index as broadly as possible.

Based on the results from the PCA, items 1 'smoking behaviour and history', 10 'assessing the current level of motivation to stop/ willingness to quit', 12 'Eliciting a prompt commitment from the smoker on starting the quit attempt' and 15 'nicotine dependency' might be considered for exclusion from the fidelity index. These were the items that did not load well, in spite of being implemented by the majority of the providers. Items (16 'setting quit date', 37 'offering BI leaflet', 39 'explaining expectations regarding the intervention programme' and 40 'providing reassurance') that did not load well due to little observed variance (due to being coded similarly or 'not being implemented' by majority providers) showed weaker inter-coder reliabilities. These items might be useful in other contexts and therefore, worthwhile to be explored further.

7.5.6 Limitations and challenges

Limitations that cut across Study B and Study C are discussed in Chapter 9 (page 185). However, those limitations that are more specific to Study B and the challenges encountered during study set-up and data collection are presented here.

Firstly, audiotaping was chosen for observation in this study because it was less intrusive, more feasible and more affordable than using videotaping. Videotaping allows examination of non-verbal cues in patient-provider interaction (Borrelli, 2011), providing better information on

patient-provider interaction to the overall variation in fidelity. However, given that the BCTs used in the fidelity index require some degree of verbalisation (e.g., 'advise on,' 'facilitate,' 'offer'), video recording was less likely to add much information to the data collected in this study (Lorencatto et al., 2013a, Lorencatto et al., 2013b). This assumption is supported by the data from the current study, where it is evident that the Adherence items which required a higher degree of verbalisation (e.g. advise, facilitate, offer) scored better than the Quality items (e.g. build rapport, reassure, emphasise client choice and tailor interaction etc.), which were scored rather low. Videotaping might have benefited in picking up these more subjective and possibly non-verbal cues of the patient-provider interaction. This should be potentially explored in future studies using conversation analysis (an approach to studying social interaction, embracing both verbal and non-verbal conduct) on videotaped patient-provider interactions.

Secondly, audiotaped BI sessions were directly coded from Urdu/Punjabi to English on the fidelity index. The standard for coding multi-language data is to transcribe all sessions and translate into English before coding (Small et al., 1999). However, the high costs and time required for translation and transcription of 154 audiotapes ranging from 30 to 50 minutes in duration was not possible within the resources allocated for this study. Besides, the coders in my study were bilingual and the coding process was more objective using the fidelity index than coding of BCTs from audiotapes. In the latter case, there might be more subjective judgements made for coding; therefore transcription and translation would seem to add value.

Thirdly, the individual items in the index were not weighted for their relative contribution to the overall fidelity that would be measured by implementing the index. This was not possible to establish as part of the current study, due to design limitations (that is, cross-sectional data and small sample to item ratio). However, in future longitudinal studies the relative effectiveness or weight of each item might be determined where data on fidelity and longer term quit are captured. Until then, fidelity scores obtained by implementing the fidelity index should be standardised for the number of items in each sub-index, before using in inferential statistical models.

Fourthly, there are limitations to inferences drawn from the generalisability analysis. The study might have missed contextual factors that could contribute to the overall variance. Such macro- or meso-level contextual factors were not considered in this study, even though multiple sources of variance were accounted. These could be factors external to the intervention or even the intensity of training given to the coders or providers that (if included in analysis) would reduce the variance for the accounted sources. Furthermore, the coders' rating and patient-provider interaction sessions contributed very low variance compared to the providers. It is unlikely that the low variance could be due to very consistent rating among the coders or their intense training and monitoring during coding. Besides, these differences were quite large and less likely to be ruled out, even if the coders had received intense training and had been monitored strictly during coding. To strengthen the findings of the G-study, a series of sensitivity analysis were

performed. The estimates of g-coefficients remained highly reliable after excluding the consensus scores and also after excluding the providers with very low fidelity.

Finally, carrying out the fidelity study involved certain challenges, some of which led to deviations from the original study plan.

The biggest challenge encountered was the unfortunate flooding of both Jhang and Sargodha districts due to the heavy and extended rainfall in India and Pakistan in September 2014. It not only displaced people (including patients enrolled in this study) but also led to a halt in all TB related care and activities. The TB clinics (most of which were flooded) could not resume activities for another six months. In addition, all health care workers (these include the providers in this study) were instated by the District Health Officer to conduct polio and measles campaigns and other necessary activities to help recovery from the floods in the area. The original plan was to collect self-reported quit status for the enrolled patients, verified using a Carbon Monoxide (CO) breath test. This would have provided prospective data for explaining variation in quit rates for Study C, and also to further test the predictive validity of the fidelity index as part of its psychometric testing (in Study B). However, these aspects of analyses were not actualised, due to not being able to follow-up study patients.

The second challenge was that the new research officer (hired for data collection in Sargodha) conducted a few BI sessions by himself, hoping that this would expedite the data collection. This was a deviation from the protocol and led to exclusion of ten audiotaped sessions from the analysis. However, further supportive supervision prevented such events from occurring for the rest of the data collection period.

Some technical challenges in recording sessions and transferring data were experienced. Initially, the plan was to upload audiotapes from the field via Google Drive (recommended by the IT support at the University) shared with the lead researcher. However, the audiotapes were too large to be uploaded and therefore, encrypted USB drives and secure cabinets for storing these (at the District Health Offices) had to be arranged. Other technical challenges involved recharging batteries, because there was no secure area in the clinic where the providers could leave the recorders to charge. They were then given extra alkaline batteries to use when the rechargeable batteries ran out. Furthermore, background noise was an issue in initial recordings as the TB clinics were very busy, even though the type of digital recorder used was designed to minimise surrounding noise. Therefore, the providers were advised to carry out the BI sessions in a separate and quieter place, which was also important to emphasise patient confidentiality but not always feasible. Although the audiotaping and transfer of data was pilot tested before starting the study, these issues only became apparent during the actual data collection in study sites. In future studies using observation methods for intervention fidelity, it is advisable to do pilot testing of procedures and equipment in the same settings/clinics where actual study data is to be collected.

7.6 CONCLUSION

The devised index is a comprehensive and psychometrically reliable research tool for measuring fidelity to delivery of a BI for smoking cessation. Comprising 41 items under *Adherence to content* and *Quality of interaction*, (both aspects to capture intervention fidelity), it appears to be highly reliable for application by different coders with varying qualifications and skill sets. Moreover, the items (i.e. the theoretically informed ingredients of the BI) coherently measure a single construct; the intervention fidelity. In addition, the index reliably differentiates between providers capturing their variation in the practice concerning BI delivery.

Items 16 'setting quit date', 37 'offering BI leaflet', 39 'explaining expectations regarding the intervention programme' and 40 'providing reassurance' of the fidelity index need further exploration in future studies, due to not being implemented by majority of the providers in this study so they could not be investigated definitively.

Researchers are encouraged to use the fidelity index in future prospective studies evaluating BIs of smoking cessation for further validation and refinement. The fidelity index in this study was designed using taxonomies for smoking cessation; however, some of the BCTs are overarching with other healthy life-style behaviours, providing an avenue of exploring its use in similar BIs, in general.

Chapter 8. Study C: Explaining Variation in Quit Rates

This chapter describes the third study in a series of linked studies. In the previous two chapters, I described the methods used to develop a fidelity index (study A) and an observational study to test its psychometric properties (study B). The output of these two studies was a new measure of fidelity, validated for content and for psychometric properties- the fidelity index. The study presented here (study C) explores how scores obtained using the fidelity index can be used to explain variation in quit rates observed in the ASSIST study.

The psychometric properties of the fidelity index tested in study B were found to be good. The index was reliable in its ability to be used by 'amateur' coders, measure the same construct (i.e. fidelity) and differentiate between providers based on their delivery of the BI. This last property of the index suggests that its scores can be used to explain variations in quit rates between different providers. Fidelity scores are considered implementation outcome variables (Peters et al., 2013) in studies where fidelity is the primary outcome. These are also used as the intermediate variables (Brownson et al., 2012, Proctor et al., 2011) in intervention-outcome research for interpreting the differential effect of complex interventions (Brownson et al., 2012, Proctor et al., 2011). In this study, fidelity score is used as an intermediate variable for exploring variation in quit rates.

I used the prospective Fidelity study (study B) to test the psychometric properties of the index, derive fidelity scores and to explore any association between these scores and variation in quit rates (study C) observed in the ASSIST study (Appendix B), illustrated in Figure 8.1. This approach is supported by publications on refinements to the MRC framework (Campbell et al., 2007b, Craig et al., 2008), and allows resources to be used efficiently to answer several research questions simultaneously. I took Fidelity study as the starting point (T1- in 2014) as my data is retrospectively linked to ASSIST study in the past (T2- in 2010).

This chapter provides the background literature to inform study methodology, the methods used and results obtained for the secondary investigation to see if fidelity to a BI measured in providers (in study B), now explains residual variation in quit rates in the ASSIST study conducted four years earlier, but delivered by the same providers in the same settings.

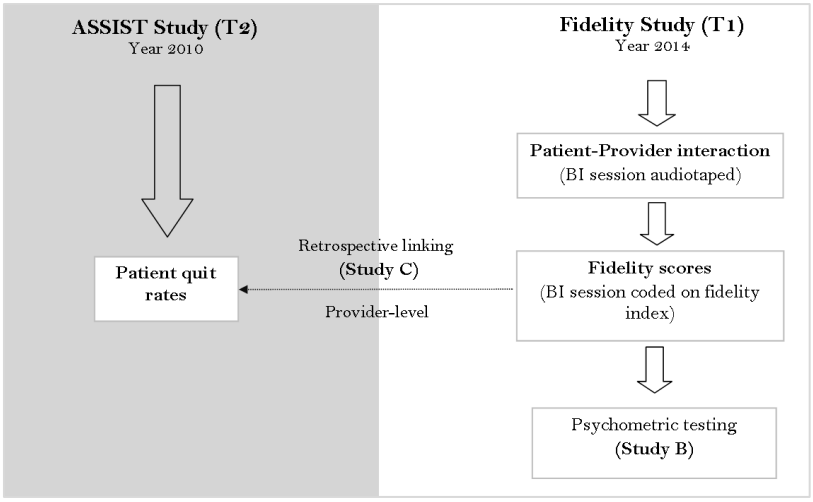


Figure 8.1: Outline of secondary investigation -study C

8.1 BACKGROUND

I first outline the literature to inform the assumptions made while combining data from two different studies and the justification for the statistical methods used to assess the effect of fidelity on quit rates.

8.1.1 Secondary analysis of existing data

Secondary data analysis is a well-established methodology (Smith et al., 2011). Broadly speaking it is the analysis of data collected by someone else (Boslaugh, 2007, Koziol and Arthur, 2011). It can include any data that are examined to answer a research question other than the original hypotheses proposed in the study (Vartanian, 2010). For the current investigation, the data were not collected by someone else but by the members of the same research team, which can confuse it with primary data analysis (Cheng and Phillips, 2014). To clarify the approach undertaken, the term 'secondary analysis of existing data' is more befitting because it avoids the confusion of trying to decide whether the data used in an analysis is 'primary data' or 'secondary data' (Cheng and Phillips, 2014).

Secondary analysis is commonly used on data from sources like national surveys, institutional records, websites and journal supplements etc. (Koziol and Arthur, 2011), which are usually large population-representative samples. In the current study, the datasets used were not population-wide samples; however, they might be representative of patient-provider interaction in routine TB care settings in LMICs. In addition, ASSIST was a cluster RCT, of which the Fidelity study was a follow-on cross-sectional acquisition of data from the same settings and providers. The data from the ASSIST study is likely to be higher quality, given its robust design and biochemical verification (CO breath tests) of quit rates. Other advantages of using existing data in this study include; the data being clean (having been analysed for RCT reporting before) and my understanding of the data (having been involved in data collection for the ASSIST study).

There are two general approaches to analysing existing data; the 'research problem driven' approach and the 'data driven' approach (Cheng and Phillips, 2014). For the current study, data from the ASSIST study were used in conjunction with Fidelity study to answer an *a priori* research question, which has advantages over unplanned post-hoc or data-driven analysis (Curran-Everett and Milgrom, 2013).

8.1.2 Analysing clustered data

Repeated measures in longitudinal studies, multiple measures on the same subject, or studies in which subjects are grouped can lead to clustered data (Kirkwood and Sterne, 2003). Clustering of data might be necessary in pragmatic intervention or educational studies, where patients or students are clustered within clinics/schools. However, the statistical precision in analysing such data decreases, because the individuals within groups tend to be more similar, giving less unique information (Koziol and Arthur, 2011). Different statistical approaches exist for analysing clustered data (Kirkwood and Sterne, 2003):

1. Calculating *summary measures* for each cluster, and analysing these summary measures using standard methods;
2. Using *robust standard errors* to correct standard errors for clustering;
3. Using *random effects* models, which explicitly model the similarity between individuals in same cluster;
4. Using *generalised estimating equations (GEE)* which adjust both for standard errors and parameter estimates to allow for clustering.

The data (described in 8.3.1) used in the current study is hierarchical, that is, individual patients are clustered within TB clinics (providers). Often such data can be analysed using multi-level models (2, 3 and 4 above), which are the recommended approach for dealing with data from complex sampling designs (Koziol and Arthur, 2011). However, in this study, individual patients were different at T1 and T2, even though the providers were the same individuals. This prevents linking of data at the individual level, allowing only cluster level (providers) linking of data between the two studies.

8.1.3 Retrospective linking of data

Secondary analysis of data merged from experimental trials and post-trial sources is often used for health services delivery research (Peters, 2009). However, it must be emphasized that these analyses are correlational and not causal (Nelson et al., 2012).

Contrary to the original plan, I was unable to collect prospective data on quit rates in the Fidelity study, due to flooding in the study area (see above 7.5.6). These data would have been used for validating the retrospective quit rates from the ASSIST study and strengthened the case for causal inference from the association of fidelity with the quit rates. Therefore, the fidelity data at

T1 (Fidelity study) had to be retrospectively linked to quit rates at T2 (ASSIST study) to investigate the set study objectives (see 8.2); assuming that the provider practice remained unchanged over time. Ideally, providers would reproduce the exact conditions in Fidelity study at T1 as implemented originally in the ASSIST study at T2. However, learning by experience or training, recall bias, and true changes may have occurred in provider practice since the BI implementation in the ASSIST study, leading to fidelity at T1 being different from T2 (Engel and Schutt, 2012). This assumption needs to be tested for scientific plausibility, of retrospectively linking fidelity scores, by comparing an indicator of providers' practice that was common between the two studies.

The patient assessment questionnaire was administered in both studies (Appendix D.3), which included a self-record checklist (on content delivered for the BI) that was filled by the providers at both T1 and T2 (see Table 8.1 for checklist content). This self-record checklist of the providers might be used to establish the consistency in their practice behaviour between T1 and T2, and strengthen the case for linking data retrospectively. Potential tests to check consistency in provider behaviour (practice scores) between T1 and T2 are discussed below.

Retrospective linking of data from multiple sources in the two studies is illustrated in Figure 8.2.

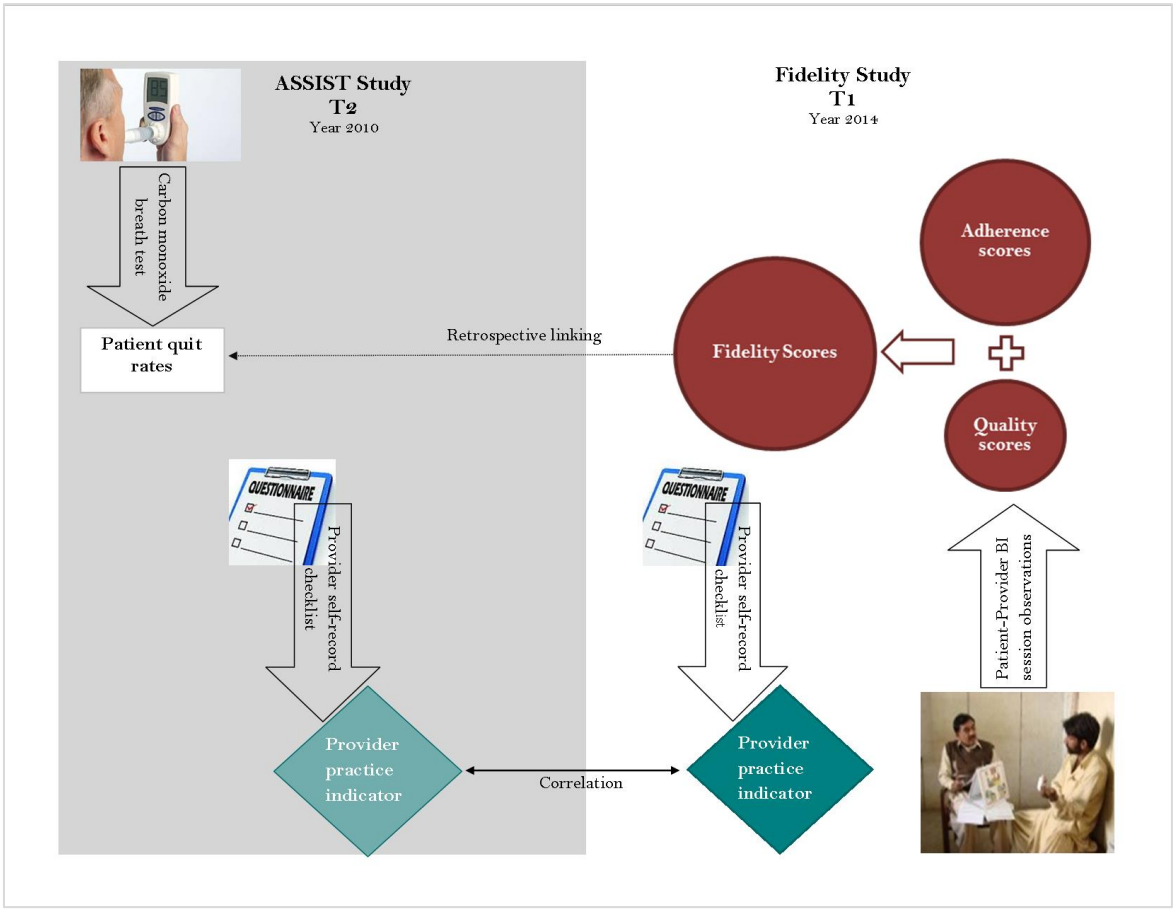


Figure 8.2: Retrospective linking of data from multiple sources

8.1.4 Consistency in provider practice behaviour over time

The degree to which these two measurements (practice scores) are related to each other could be checked by the *test–retest reliability* of the measure (Engel and Schutt, 2012). Test-retest reliability can be used to measure consistency over time or absolute agreement on a set criterion between two time-points. Decisions regarding the test for examining provider practice behaviour between T1 and T2 are dependent on whether to measure 'absolute agreement' or 'consistency' (Streiner and Norman, 2008).

'Absolute agreement' is commonly used when a group of people rank something against an external set criterion e.g. students' measurements of blood pressure when tested for accuracy against values obtained by an expert (Streiner and Norman, 2008). 'Consistency', on the other hand, has better value in determining whether the rank ordering for something among a group of people was the same when repeated (Streiner and Norman, 2008). The Kappa statistic, ICC (Streiner and Norman, 2008), and Altman and Bland method (Bland and Altman, 1986, Bland and Altman, 1999) are often used for absolute measure of agreement for test-retest reliability. However, in this study, testing whether the ordering of relative performance of the providers is preserved overtime is deemed important. Therefore, methods for agreement relative to the performance of providers at T1 and T2 are explored further.

The practice score of the providers is a continuous summary (of patients) for each provider. When there are paired measurements (same provider at T1 and T2), the Pearson (product moment) correlation coefficient '*r*' can be applied (Kirkwood and Sterne, 2003). If the variable of interest is not normally distributed, then its counterpart non-parametric tests (Kendall's tau and Spearman's rank correlations) can be used (Kirkwood and Sterne, 2003). Spearman's rank correlation calculates '*r*' between the ranks given to values of practice scores at T1 and T2, rather than between the original practice scores at T1 and T2, as Pearson's *r* does. Kendall's rank correlation (tau) compares the ranks of scores at T1 and T2 between each pair of observations for concordance (degree of similarity in a pair) and discordance (degree of dissimilarity in a pair). It simply calculates the ranks given to values of practice scores at T1 and at T2, such that higher values (and lower values) at both time-points are correlated, rather than pairing these values for the same providers. These rank tests might not be appropriate for checking concordance for the same providers paired at T1 and T2.

A slightly different version of Kendall's rank correlation is Kendall's W, also known as Kendall's coefficient of concordance (Kendall and Smith, 1939), which measures associations between ratings. It is a normalization of the statistic of the Friedman test used for one-way repeated measures analysis of variance by ranks and ranges from 0 (no concordance) to 1 (complete concordance), similar to other reliability coefficients. The procedure allows ranking of pairs of the same providers (depending on how the data is setup) and then considering the values of ranks by T1 and T2. Therefore, Kendall's W seems to be the most appropriate test for checking if the ordering of provider practice was preserved over time.

8.2 AIMS OF STUDY

The main purpose of the secondary investigation, in this study, is to assess whether fidelity index can be used to explain the variation in quit rates and to generate hypotheses for interpreting the effects of the BI on its cessation outcomes (refer to Figure 4.1: Interpretive Evaluation Model, page 55).

Therefore, the study aim is;

"To explore whether fidelity to a BI explains variation in quit rates" (described above in 5.2).

The specific objectives of study C are to:

- Describe ***intervention fidelity*** to the BI used in the ASSIST study
- Assess ***provider consistency in practice*** behaviour over time
- Explore any association between ***provider fidelity to the BI and the patient quit rates***, to explain the observed variation
- Explore the relationship between ***Adherence to intervention content and Quality of interaction***, to generate hypotheses for appropriate mediation or moderation pathways

8.3 METHODS

8.3.1 Datasets

As noted in Figure 8.2 (above), the data sources within the two primary studies were the patient assessment questionnaire, the patient-provider BI session audiotapes and the CO breath test for quit status in patients. Figure 8.3 describes the data extracted from each study before merging for analysis. The details of the datasets are given below.

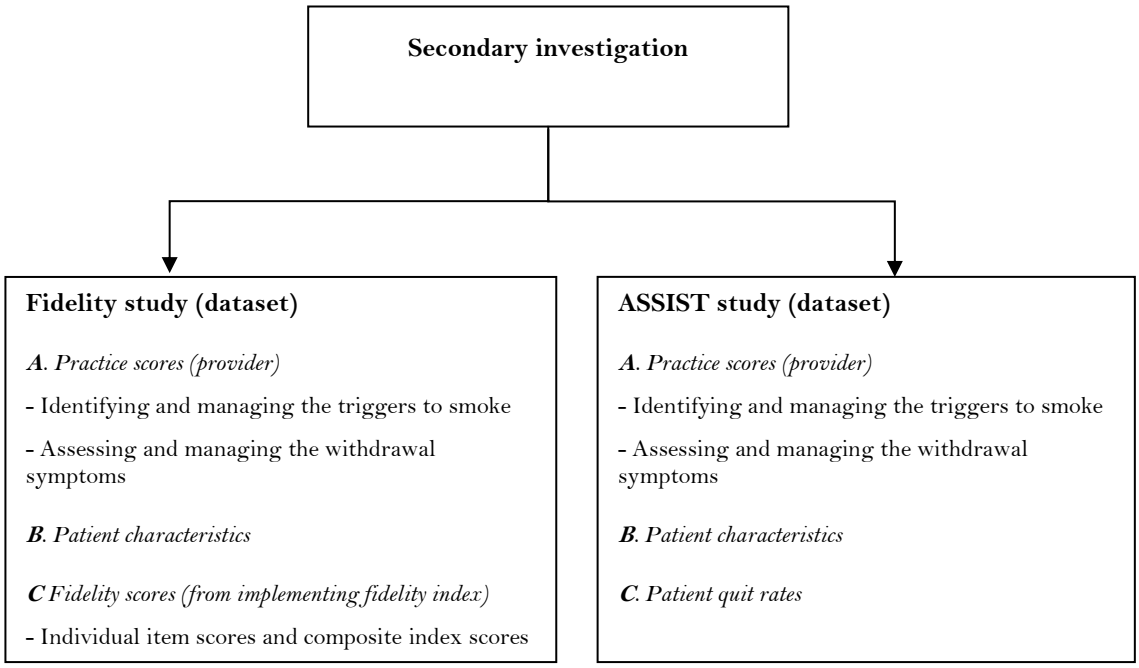


Figure 8.3: Sources of data collection for secondary investigation

ASSIST study data

- Patient assessment questionnaire: contained information (recorded by the provider) on sociodemographic, smoking history and nicotine dependency, triggers and withdrawals assessments, management strategies offered for triggers and withdrawals (self-record checklist by the provider), CO measurement (quit status) at baseline, 1 month and 6 month.

**Note:* Data were available for 22 providers (intervention arms) delivering BI for smoking cessation and 1299 patients.

Fidelity study data

- Patient assessment questionnaire: contained information (recorded by the provider) on sociodemographic, smoking history and nicotine dependency, triggers and withdrawals assessments, management strategies offered for triggers and withdrawals (self-record checklist by the provider).
- Fidelity scores (obtained by implementing the fidelity index): a) Adherence score: 37 items assessing adherence to BI content, with a maximum score possible of 74. b) Quality score: 8 items assessing essential skills or competence to deliver BI, with a maximum score possible of 16.

**Note:* Providers delivering BI were the same as in the ASSIST study, delivering the same intervention. However, data were available for 19 providers instead of the original 22 (as

described in chapter 7) and 180 patients. Fidelity scores were available for 18 of the 19 providers and 154 patients.

8.3.2 Data variables

In this study, some variables were defined based on how these were created in the original studies and some new variables were created from the existing ones, to retain their values when the unit of analysis changes from individual to cluster level. These were defined here.

Fidelity scores

These were the consensus scores (between the three coders) obtained by rating the fidelity index against the BI session audiotapes in the Fidelity study. These are composite scores for the two sub-indices; 37 items for *Adherence* with a maximum score (A score) of 74 achievable and 8 items for *Quality* with a maximum score (Q score) of 16 achievable.

For use in this study, the fidelity scores were standardised by dividing by the number of items in each sub-index:

A score (standardised) = Adherence score/37;

Q score (standardised) = Quality score/8;

The standardised A and Q scores were then summarised by computing averages of the scores obtained for all audiotaped sessions per provider, ready for merging.

Patient quit rates

This was the proportion of patients who quit smoking for each provider. The variable 'continuous abstinence' was created originally in the ASSIST study, which was re-defined for use as 'quit rate' in the current study.

Patient quit status was assessed using CO measurements at 1 and 6 month follow-up in the ASSIST study. A CO measurement of 9ppm or less classified a patient as quit from smoking (West et al., 2005). A binary abstinence measure (Yes/No) was created using this cut-off for both 1 and 6 month measurements. 'Continuous abstinence' (Yes/No) from smoking was created by combining the binary abstinence for both 1 and 6 months, such that patients who were abstinent at both 1 and 6 months were considered as 'quit' and patients who smoked at any of the time-points (1 or 6 month) were considered as 'smokers'. The unit of analysis for 'quit' was 'proportion quit' at individual level, as used in the ASSIST study analyses.

Summarising binary 'quit' outcome was not as straight forward as taking an average of all patients' quit for the respective provider, as was the case with continuous variables. Averaging binary variables involved creating dummy variables; a variable with value 1 if quit and 0 if not,

and a separate variable with value 1 if smoker and 0 if not. The summary of each dummy variable ('quit' and 'smoker') for the respective provider would yield a count when summarised, that is, the number of patients who quitted ('quit' per provider) and the number of patients who remained smokers ('smoker' per provider). These count variables could have been used for provider-level analysis if the clusters (number of patients enrolled per provider) were equal in size. However, the cluster size varied in the ASSIST study between providers, ranging from 49 to 74. If the count variables ('quit' and 'smokers') were summarised at provider-level they would be unable to retain the actual unit of 'proportion quit', due to the varying cluster sizes. Therefore, to retain the unit of analysis (proportion quit) in unequal clusters, binomial proportions were computed for all patients per provider (Shoukri and Cihon, 1998).

The binomial proportion for 'quit rate' was defined as "the proportion quit out of the total patients per provider", with values lying between 0 and 1. The numerator for 'quit rate' was the number of patients who quit in a cluster and the denominator was the total number of patients in that cluster.

Practice scores

These were the scores representing provider practice behaviour, obtained from the self-record checklists used both in the ASSIST study and the Fidelity study (see Figure 8.3). These scores were used to check the assumption that provider practice behaviour remained unchanged between T1 and T2 to lend more validity to the retrospective linking of the two studies.

The self-record checklist (see Table 8.1) was integrated for use with the BI flip-chart, by the providers, as they delivered behavioural support. It comprised multiple components, some representing the urge or symptom in the patient and others representing the management strategy for overcoming that particular urge or symptom advised by the provider. For most triggers and withdrawals, the number of strategies offered was three, but sometimes the provider tailored and offered an extra strategy, not originally part of the intended practice, which was coded as 'other'.

There were two variables for each trigger and withdrawal (16 in total) in the original datasets; however, for the purpose of the current analysis a single score representative of provider practice was needed. Therefore, a new composite indicator was created combining these variables to represent provider practice for delivering BI to the patient. Although both triggers and withdrawals relate to nicotine dependence, triggers are the (psychological) urges or desire to smoke and withdrawals are the (actual or physical) signs and symptoms an individual feels when they attempt to quit. Hence, separate practice scores were created, one for triggers and one for withdrawals, due to the slight differences in definitions of both terms.

Table 8.1: Self -record checklist of the provider

Triggers	Strategies to manage		
	1	2	3
Rising in the morning (immediately after)	Take some drink (Juice, green tea, lassi) after you get out of bed	Brush your teeth	Go for a walk
Defecation	Take some laxatives (Isbagol Husk) at night	Take newspaper to toilet instead of cigarette	-
Eating Meals	Take a chewing gum	Go for a walk after meals	Take a nap after lunch
Relaxing/boredom at Home	Take some dry fruit/grams	Do gardening	Have tea, Watch TV
Seeing others smoking	Take a chewing gum	Advocate to the friends about the smoking harms	Avoid smoker friends in the initial period
Offered smoking	Refuse it straight away	Take chewing gum	Relax and take deep breathes
Intensive Physical/ Mental work	Take dry fruit	Relax and take deep breaths	Take small short breaks and have tea
Tension/Anxiety	Chat with colleagues	Relaxing Exercises	Offer Nimaz
Withdrawals	**Strategies to manage**		
Craving/Desire for cigarettes	Take a glass of juice/ cold drink	Engage in conversation with a friend	Take deep breadths and slowly breathe out
Restlessness Irritability/ frustration/ Anger	Take a glass of juice/ cold drink	Engage in conversation with a friend	Take deep breadths and slowly breathe out
Cough	Take some throat soothers	Make a habit of morning walk	Consult doctor if cough is troublesome
Weight Gain	Take more fruits and vegetables in your diet	Exercise daily	Make a habit of morning walk
Headache	Take a pill for headache	Engage in conversation with a friend	Offer nimaz
Insomnia	Exercise daily	Offer nimaz	Consult doctor if not getting better
Anorexia & Constipation	Take plenty of fluids/vegetables/fruits	Make a habit of morning walk	If disturbing consult your doctor
Indigestion/ Heart burn	Take more fruits and vegetables.	Make a habit of morning walk	Consult doctor if not getting better

The following steps were involved in generating the practice scores.

Step 1

Number of strategies offered per patient (see Table 8.1) per trigger and per withdrawal was combined to create the composite 'indicator', such that:

- If one strategy was offered for a trigger (or a withdrawal) it was coded as 1,
- If two strategies were offered for a trigger (or a withdrawal) it was coded as 2, and,
- If three or more strategies were offered for a trigger (or a withdrawal) it was coded as 3.
- If a patient did not have a particular trigger (or a withdrawal), this was coded as 'missing' so as not to be counted in calculations.
- If the patient had a trigger (or a withdrawal) but no management strategies were offered to them then it was coded as '0' to indicate no management given.

Step 2

The sum of strategies offered per patient for all eight triggers and for all eight withdrawals was calculated for the indicators, such that:

- Number of strategies (triggers) = ∑ strategies for Trigger1 to Trigger8;
- Number of strategies (withdrawals) = ∑ strategies for Withdrawal1 to Withdrawal8

Step 3

This was the final step for generating the practice scores. The number of triggers and the number of withdrawals reported by the patients varied, therefore computing a score that was simply a summation of the number of strategies offered to each patient would not have captured the provider practice appropriately. As the strategies offered to a patient by the provider were relative to the number of triggers or the number of withdrawals that the patient had, it was sensible to standardise the indicator created in step 2 by using these numbers.

Standardising the indicator also had clinical significance because the higher the number of triggers or withdrawals the patient had, the less likely they were to succeed in quitting. By standardising the indicator, the practice scores that were generated were better representative of provider practice, as they measured the self-recorded practice in terms of strategies offered to each patient relative to the number of triggers/withdrawals that the patient had. The standardisation process is as follows.

- Total number of triggers (and of withdrawals) was calculated per patient:

  Number of triggers = ∑ Trigger1 to Trigger8

  Number of withdrawals = ∑ Withdrawal1 to Withdrawal8

- The practice scores were calculated per patient:

  Practice score (T- trigger) = $\frac{\text{Number of strategies for triggers}}{}$

$$\text{Practice score (W- withdrawal)} = \frac{\text{Number of strategies for withdrawals}}{\text{Number of withdrawals}}$$

Definition of the practice scores

The numerator for the practice scores (both for T and for W) was any combination of 0 to 23 or 24; that is, the sum of all those strategies offered to the patient for the triggers/withdrawals (see Table 8.1) and the denominator was any range of 0 to 8 (number of triggers/withdrawals of the patient). The practice scores ranged between 0 and 3 and can be interpreted as;

- '0' - when the patient had one or more triggers (or withdrawals) but was offered nothing by the provider;
- '1' - when the patient was offered on average the same number of strategies as the number of triggers (or withdrawals) they had;
- '2' - when the patient was offered on average twice the number of strategies than the number of triggers (or withdrawals) they had; and,
- '3' - when the patient was offered on average three times the number of strategies than the number of triggers (or withdrawals) they had.

The practice scores (both for T and for W) were summarised by computing averages of the scores obtained for all patients per provider.

8.3.3 Data extraction and merging

The data extracted from the ASSIST study included the binary (Yes/No) patient quit variable for 1299 patients (for 22 providers) and the 16 variables from the self-record checklist. The data extracted from the Fidelity study (as part of my PhD research) included the consensus scores for fidelity (both for *Adherence* and for *Quality*) for 154 audiotaped sessions (for 18 providers) and the 16 variables from the self-record checklist for 180 patients (for 19 providers). In addition, patient characteristics (e.g. age, gender etc.) collected by administering the patient assessment questionnaire were also extracted from both studies.

Merging data requires a unique characteristic common between the two sources. The variables of significance in this study were hierarchical (patients clustered by providers). It was the providers who were common between the two time periods and so data were merged by provider IDs.

8.3.4 Statistical analysis

Summary statistics of patient characteristics, provider practice scores, fidelity to delivery of BI sessions and the variation in fidelity scores (by providers) were computed. Analysis was carried out to investigate consistency in providers' practice behaviour between T1 and T2. In addition,

the association between fidelity and quit rates, as well as, the relationship between Adherence and Quality scores, were explored.

All analyses were conducted using SAS (version 9.4, Cary, NC, USA); see Table 8.2, for the SAS syntax used in analysis. The methods are described below.

Patient characteristics

Patient sample characteristics for both studies (ASSIST and Fidelity study) were summarised for variables, including; age, gender, age of initiating smoking, duration of smoking, nicotine dependence score, quantity of smoking, form of smoking and quit attempt in the past. Mean (with Standard Deviation-SD) and median (with Inter-Quartile Range-IQR) were reported for continuous variables and percentages for categorical variables.

Description of practice scores

Simple descriptive statistics of the provider practice scores from both T1 and T2 were calculated as mean, minimum, median, maximum and variance. Scatterplots of the practice scores from the two studies were also generated, both for the triggers and the withdrawals.

Description of fidelity

Implementing the fidelity index (as outlined in Figure 8.3) generated data on provider fidelity for delivering the BI sessions. Summary statistics of these data to describe the variation in fidelity scores and the delivery of the 45 items (BI content) by the providers were reported as follows:

Fidelity to BI delivery

The proportion of patient-provider interaction sessions in which different BI ingredients (represented by the 45 items of the fidelity index) were delivered was calculated, showing the dose of BI delivered.

Variation in Fidelity

Intra-cluster variability was computed by taking averages of fidelity scores (both A and Q scores) and standard deviations (SD) per provider. An average of the mean fidelity scores across providers was taken to calculate the inter-cluster variability.

Note: Summary statistics for fidelity scores were also presented in Chapter 7 (see Table 7.13 and Figure 7.5, pages 142-143) as part of Generalisability study, but these were not standardised scores. The fidelity scores used here were standardised for the number of items per sub-index (see 'Fidelity scores', page 158).

Examining consistency in provider practice over time

For measuring concordance (degree of similarity in a pair) in provider practice scores between T1 and T2, Kendall's W statistic was used, which can be interpreted in a similar way to a correlation coefficient *r* (see Table 8.2, on page 165, for the SAS syntax).

If there is sufficient concordance between T1 and T2, this would give some empirical support to the hypothesis that the way in which the smoking cessation BI was delivered, in the ASSIST study, and the Fidelity study four years later, was similar. This would then justify explaining the extent to which fidelity measured in the Fidelity study could be used to explore and possibly explain variations in quit rates in the ASSIST study (as noted above in 8.1.3).

Examining association between fidelity and quit rates

Mean, minimum, median, maximum and variance were calculated for the fidelity scores and the quit rates. 2D and 3D plots of quit rates and fidelity scores were also presented.

Binomial regression was carried out to analyse quit rates (dependent variable) for which fidelity scores (independent or explanatory variables), both for Adherence and for Quality, were used to explore the variation between providers; logit models with binomial link were considered appropriate (Zhao et al., 2001). Crude and adjusted estimates were computed. Odds Ratios (OR) and 95% confidence intervals were reported. Unlike logistic regression, estimated relative prevalence can be reported from binomial regression. Interaction between Adherence and Quality was also explored. However, only a model without considering effect modification yields adjusted relative prevalence (see Table 8.2, for the SAS syntax). Therefore, the odds of quitting in patients receiving BI were computed for provider Quality of interaction at varying levels of their Adherence to content (the A scores), to explain the interaction between these two covariates.

Examining association between Adherence and Quality scores

Further, simple linear regression was performed to look at the relationship between the Adherence scores and the Quality scores, to hypothesise possible mediation or moderation pathways. Non-normally distributed variables were log transformed for a better fit and for simpler interpretation of regression coefficients (see Table 8.2, for the SAS syntax). Log transformation codes values that are recorded '0' as missing, therefore, Quality scores having '0' values for five providers were inputted with '0.01' to keep them in the regression model.

Table 8.2: SAS syntax for analytical statistics

Analysis	SAS syntax

Kendall's W	The syntax (in SAS) used for the MAGREE macro is:
	`%inc data A (ASSIST study);`
	`  %magree(data=A,`
	`  items=S,`
	`  raters=R,`
	`  response=Y);`
	
	`%inc data B (Observational study);`
	`  %magree(data=B,`
	`  items=S,`
	`  raters=R,`
	`  response=Y);`

Binomial Regression	`/*Crude regression estimates*/`
	`proc genmod data=;`
	`model Number Abstinent/Number Total=Ascore/link=logit dist=binomial;`
	`run;`
	`proc genmod data=;`
	`model Number Abstinent/Number Total=Qscore/link=logit dist=binomial;`
	`run;`
	
	`/*Adjusting for Quality*/`
	`proc genmod data=;`
	`model N ABS/N TOTAL= ASCORE QSCORE/link=logit dist=binomial;`
	`ESTIMATE 'OR OF ADHERENCE' ASCORE 1/EXP;`
	`ESTIMATE 'OR OF QUALITY' QSCORE 1/EXP;`
	`run;`
	
	`/*Interaction term*/`
	`proc genmod data=;`
	`model N_ABS/N_TOTAL= ASCORE QSCORE ASCORE*QSCORE/link=logit dist=binomial TYPE3;`
	`run;`

Simple Linear Regression	`/*Log transformation of scores*/`
	`DATA;`
	`SET;`
	`LnASCORE= log(ASCORE);  /* The natural logarithm (base e) */`
	`LnQSCORE= Log(QSCORE);`
	`RUN;`
	
	`/*Crude Regression coefficients*/`
	`PROC REG DATA=;`
	`MODEL LnASCORE= LnQSCORE/ CLB;`
	`RUN;`

8.4 RESULTS

8.4.1 Patient characteristics

The majority of patient characteristics were similar in the two studies (Table 8.3), with higher numbers of past quit attempts seen in the Fidelity study.

Table 8.3: Participant characteristics in primary studies

Characteristics	Fidelity study	ASSIST study
Number of patients	180	1299
Men, n (%)	179 (99.44)	1217 (94.78)
Mean age, years (SD) *	42.92 (13.96)	40.55 (13.13)
Mean age when started smoking, years (SD)	19.53 (7.87)	20.39 (6.64)
Mean duration of smoking, years (SD)	23.25 (13.80)	20.22 (12.49)
Median nicotine dependency score, n (IQR) **	6.00 (4.00)	6.00 (3.00)
Median cigarette and hookah smoked per day, n (IQR)	20.00 (10.00)	20.00 (18.00)
Forms of smoking, n (%)		
Cigarette smokers	109 (62.64)	840 (64.67)
Hookah smokers	23 (13.22)	145 (11.16)
Dual smokers (Hookah + Cigarette)	42 (24.14)	314 (24.17)
Attempted quit in past, n (%)	92 (51.11)	359 (27.74)

*Standard Deviation
** Inter-Quartile Range

8.4.2 Description of practice scores

The practice scores (for both triggers and withdrawals) from T1 and T2 concentrated near 1 (Figure 8.4; Table 8.4), indicating that most providers offered one management strategy to their patients for each trigger (or each withdrawal) that the patient reported.

Table 8.4: Descriptive statistics of practice scores at T1 & T2

Practice scores	N	Min	Median	Max	Average	Variance
T1 (Fidelity study)						
Triggers	18	1.00	1.06	1.85	1.17	0.05
Withdrawals	15	1.04	1.11	1.67	1.18	0.04
T2 (ASSIST study)						
Triggers	22	1.00	1.07	2.77	1.19	0.14
Withdrawals	22	0.93	1.11	2.73	1.26	0.15

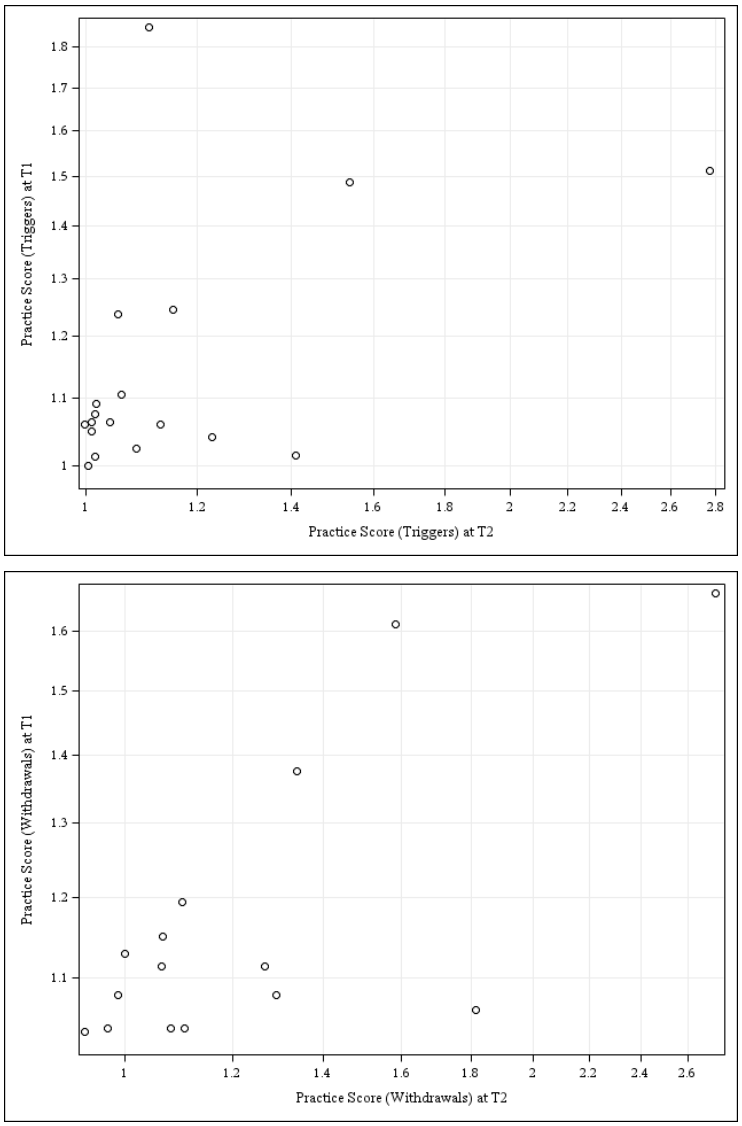


Figure 8.4: Scatterplots of 'practice scores' at T1 and T2

Note: 18 and 15 providers completed the self-record checklist for triggers and withdrawals, respectively in the Fidelity study.

8.4.3 Description of fidelity

Fidelity to BI delivery

Figure 8.5 is a graph illustrating the fidelity to the BI ingredients delivered overall. On the y-axis are the BI ingredients (represented by the 45 items of the fidelity index) and on the x-axis are the percentages of patient-provider BI sessions. Green bars represent the percentage 'fully implemented', amber 'partially implemented' and red 'not implemented'; the three response scale alternatives (described in 6.8.3) representing functional features of the fidelity index.

BI ingredients relating to assessment of nicotine dependence, readiness to quit and past quit history were fully implemented in more than 80% of the sessions (Figure 8.5). Information about harms of smoking, preparation and planning for the quit date were fully implemented in 30 to 40% of the sessions. Triggers management was fully implemented in less than 20% of the sessions, and withdrawal symptoms management was fully implemented in less than 10% of the sessions. However, both triggers and withdrawals management was partially implemented in 30 to 40% of the sessions. Ingredients relating to the Quality of interaction were the least implemented; fully implemented in less than 10%, and partially implemented in 10 to 20% of the delivered sessions.

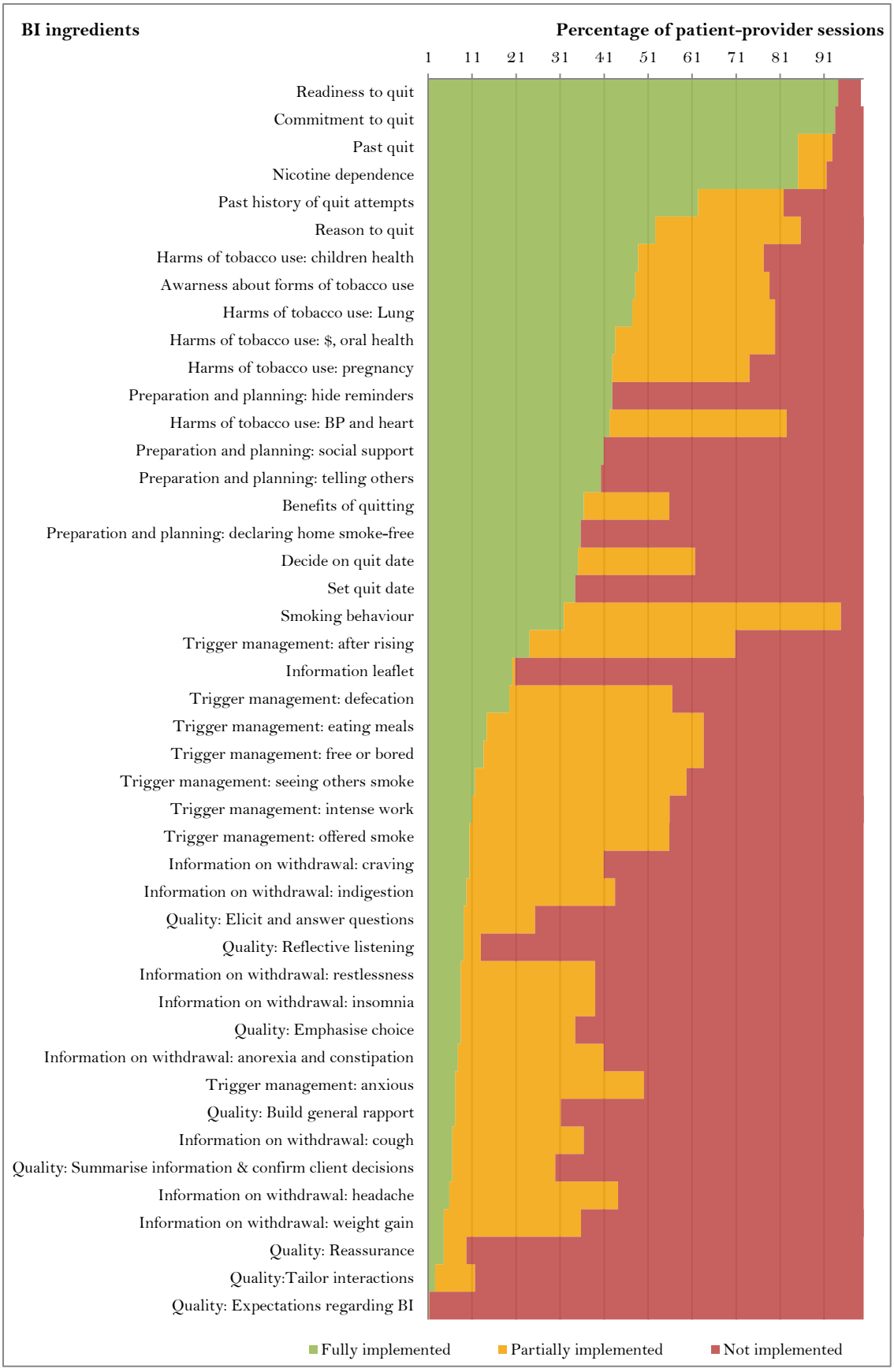


Figure 8.5: Fidelity of delivery - percentage of patient-provider sessions delivered by BI ingredient

Variation in Fidelity

The average provider *Adherence* scores ranged from 0.31 (SD: 0.19) to 1.72 (SD: 0.16); a maximum score of 2 was possible (Table 8.5). Table 8.5 is arranged by the descending order of Adherence scores. The average provider *Quality* scores ranged from 0 to 1.38 (SD: 0.42); a maximum score of 2 was possible.

The between-provider variability was 0.94 (SD: 0.44) for Adherence and 0.26 (SD: 0.38) for Quality.

Note: Description of the providers was given in study B (see section 7.3.8).

Table 8.5: Average fidelity scores within and between providers

Provider ID	**Fidelity scores** Mean (SD*)		
	Number of patients	Adherence	Quality
Intra-provider variability			
24	6	0.31 (0.19)	0 (0)
26	5	0.36 (0.07)	0 (0)
4	9	0.39 (0.16)	0 (0)
27	10	0.41 (0.11)	0 (0)
7	9	0.43 (0.14)	0 (0)
22	10	0.62 (0.22)	0.03 (0.08)
25	10	0.82 (0.17)	0.03 (0.05)
11	10	0.93 (0.43)	0.26 (0.27)
28	9	0.94 (0.19)	0.08 (0.06)
21	8	0.96 (0.38)	0.09 (0.11)
5	9	1.04 (0.36)	0.33 (0.11)
19	8	1.14 (0.18)	1.38 (0.42)
3	10	1.19 (0.22)	1.00 (0.22)
23	8	1.20 (0.41)	0.23 (0.10)
2	9	1.30 (0.28)	0.07 (0.17)
20	10	1.48 (0.15)	0.31 (0.36)
9	10	1.61 (0.18)	0.29 (0.29)
8	4	1.72 (0.16)	0.50 (0.35)
Inter-provider variability			
		0.94 (0.44)	0.26 (0.38)
*SD is the standard deviation of the mean			

8.4.4 Consistency in providers practice over time

There was moderate to high consistency in provider practice concerning management of patient triggers (Kendall's W: 0.69, p=0.06) and withdrawal symptoms (Kendall's W: 0.75, p=0.03), between the two studies time-points. Kendall's coefficient of concordance was statistically significant for withdrawal management practice and borderline significant for trigger management practice, indicating stronger agreement than can be expected by chance alone (Table 8.6).

Table 8.6: Kendall's W for concordance between providers practice at T1 and T2

Coefficient of concordance	Practice score (Triggers)	Practice score (Withdrawals)
Kendall's W	0.69	0.75
(p-value for Friedman test)	(0.06)	(0.03)

8.4.5 Association of fidelity scores with quit rates

The median quit rates were 0.48 (range: 0.07-0.72), while the median fidelity scores were 0.95 (range: 0.31-1.72) for *Adherence* and 0.09 (range: 0.00-1.38) for *Quality* (Table 8.7). Figures 8.6 and 8.7 both do not show a linear relationship between the quit rates and the fidelity scores.

Table 8.7: Descriptive statistics of quit rates and Fidelity (Adherence & Quality) scores

Variable	**N**	**Min**	**Median**	**Max**	**Average**	**Variance**
Proportion quit	22	0.07	0.48	0.72	0.43	0.04
Adherence score	18	0.31	0.95	1.72	0.93	0.20
Quality score	18	0.00	0.09	1.38	0.26	0.14

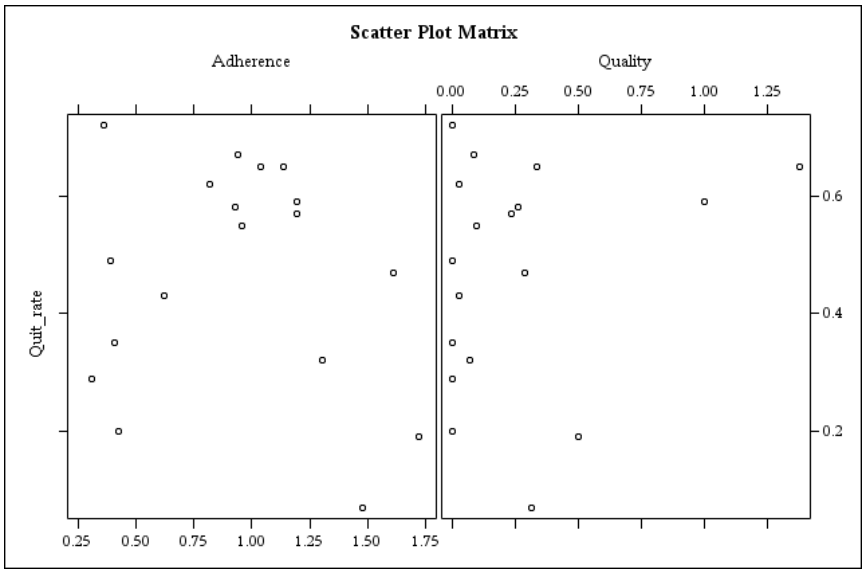


Figure 8.6: 2D plot of Quit rates by Adherence and Quality scores

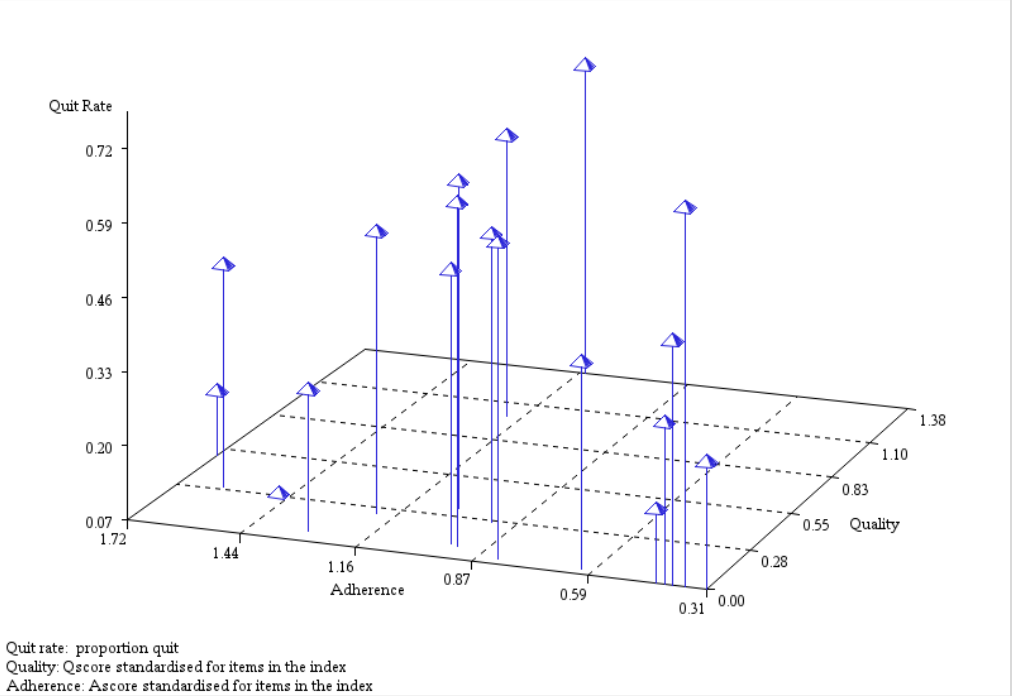


Figure 8.7: 3D plot of Quit rates by Adherence and Quality scores

The results of the binomial regression (model 1, Table 8.8) indicate that an increase in the *Quality* of interaction with the patient led to a 50% increase in the quit rates by provider and the effect estimate was statistically significant ($p < 0.05$). An increase in the level of *Adherence* to BI content delivered to the patient led to a 24% decrease in the quit rates by provider, although the effect estimate was of marginal statistical significance ($p = 0.055$).

When adjusted for Quality (model 2, Table 8.8), the association of *Adherence* with quit rates became statistically significant, though remaining negative. When adjusted for Adherence the association of *Quality* with quit rates became stronger (model 2, Table 8.8).

Table 8.8: Effect of provider fidelity (Adherence & Quality) on patient quit rates

Independent variable	Odds ratio (95% CI)	*p*-value (Wald Chi-Sq)
Model 1: Crude effects		
Adherence score	0.76 (0.57-1.01)	0.0552
Quality score	1.50 (1.06-2.12)	0.0226
Model 2: Adjusted effects		
Adherence score (adj. for Quality)	0.55 (0.40- 0.77)	0.0005
Quality score (adj. for Adherence)	2.15 (1.43- 3.24)	0.0002
Model 3: Introducing the interaction term		
Adherence * Quality	*-	<0.0001
* The interaction parameter estimate was -7.23 (-9.68 to -4.79), showing a negative interaction of high magnitude between Adherence and Quality in explaining quit rates		

A negative interaction was found between Adherence and Quality ($p<0.0001$) when the interaction term (Adherence*Quality) was introduced (model 3, Table 8.8). This means that the association between one of the two co-variates (i.e. Adherence and Quality) and the quit rates decreases, if the other co-variate increases (Szklo and Nieto, 2007).

This interaction was shown for Quality at varying levels of Adherence (Table 8.9), as computed from model 2 regression estimates (Table 8.8). These calculations found that at a good Quality of BI interaction with the patients, the patient quit rates decreased with increasing levels of Adherence for the provider.

Table 8.9: Odds Ratios of quit for *Quality* of interaction at six levels of *Adherence* to BI content

Level of Adherence (A score)	Odds of quitting for good Quality at different levels of Adherence	Odds Ratio
0.31	$e^{0.2225 + (-0.595*0.31) + 0.765}$	2.23
0.59	$e^{0.2225 + (-0.595*0.59) + 0.765}$	1.89
0.87	$e^{0.2225 + (-0.595*0.87) + 0.765}$	1.60
1.16	$e^{0.2225 + (-0.595*1.16) + 0.765}$	1.35
1.44	$e^{0.2225 + (-0.595*1.44) + 0.765}$	1.14
1.72	$e^{0.2225 + (-0.595*1.72) + 0.765}$	0.97
Computation formula for Odds Ratio: $e^{\text{intercept} + (\text{log odds of Adherence} * \text{level of Adherence}) + \text{log odds of Quality}}$ e is exponentiation of the log odds		

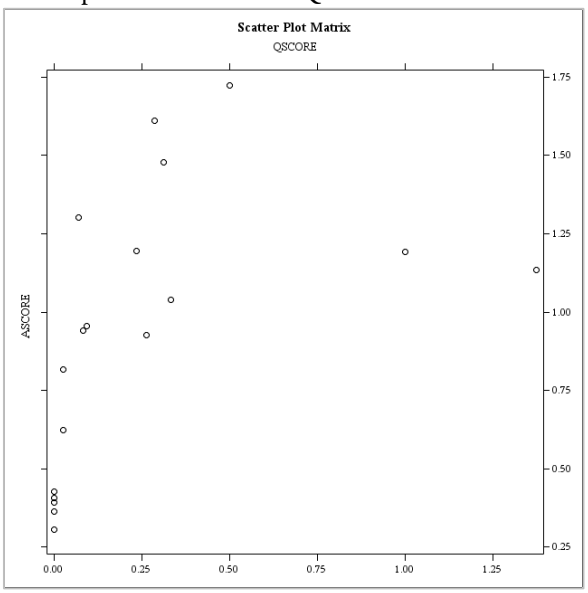
8.4.6 Relationship between Adherence and Quality scores

The scatter plots of the A and Q scores after log transformation showed better linear fit than without the log transformation (Figure 8.8).

Further exploration of the relationship between Adherence and Quality indicated that a one percent increase in *Quality* of interaction with the patient was associated with a 2.7% increase in the level of provider's *Adherence* to the BI content (Table 8.10).

Figure 8.8: Scatter plots of Adherence and Quality with and without log transformation

Scatter plot of A score with Q score



Scatter plot of log transformed A score with Q score

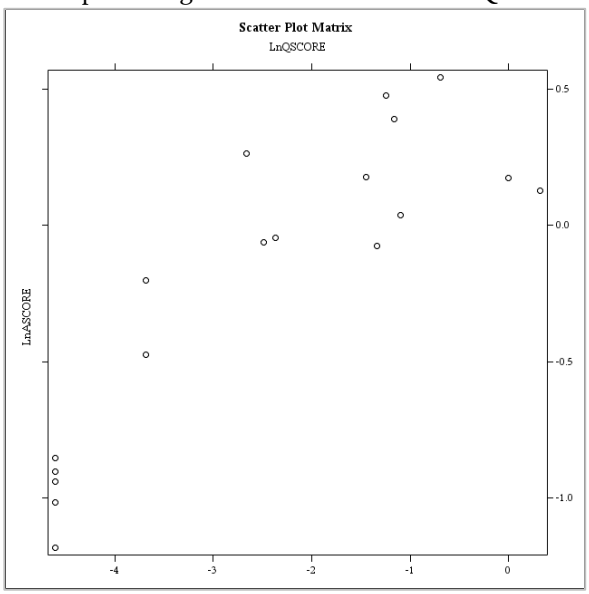


Table 8.10: Association of Quality with Adherence

Variable	*β*- estimate (95% CI)	*p*-value	R-Square	Standard Error
Quality score*	2.70 (1.90- 3.50)	<0.0001	0.76	0.37

* Both Adherence Scores and Quality scores were log transformed for regression analysis, therefore, *β*- estimate was interpreted as percentage points.

8.5 DISCUSSION

The findings specific to study C are discussed here, however, any overarching issues between the three linked studies (Study A, B and C) are discussed in the final chapter of the thesis (Chapter 9).

8.5.1 Summary of findings

Overall, in the BI sessions recorded for the patient-provider interactions, the assessments of dependency and motivation to stop (patient questionnaire) were implemented in more than 80% of the sessions and the information about the harms of smoking and benefits of quitting in 30 to 40%. The least implemented ingredients were related to triggers and withdrawals management and the Quality of interaction. Intra-cluster variability (i.e. between patients seen by the same provider) in fidelity scores was lower than the inter-cluster variability (i.e. between providers). Provider consistency in self-recording triggers and withdrawals management showed moderate to strong concordance between the Fidelity study and ASSIST study time-points.

The quit rates observed in the ASSIST study were positively associated with the *Quality* of interaction, but negatively associated with *Adherence* to BI content observed in the Fidelity study.

An interaction was also found between *Adherence* and *Quality* in their effect on quit rates, which was negative in direction. In addition, provider *Adherence* to BI content was found to be positively associated with their *Quality* of interaction with the patients.

8.5.2 Fidelity to BI

Looking at the overall fidelity of the BI delivered by the providers, it is evident that some of the BCTs were better implemented than others. The four behavioural domains that the provider 'competences' taxonomy targets (Michie et al., 2011b) were: (1) addressing motivation; (2) maximising self-regulatory capacity or skills; (3) adjuvant activities; and (4) general aspects of the interaction: delivery of the interventions, information gathering, or general communication.

In the current study, the BCTs focusing on maximising self-regulatory capacity or skills and those focusing on general aspects of the interaction (including general communication and delivery of intervention) were least implemented. These BCTs broadly covered the BI ingredients on triggers, withdrawals, the management strategies offered and the quality of interaction. The BCTs that were most frequently implemented focused on information gathering (general aspects of the interaction) and addressing motivation. These BCTs covered mainly the BI ingredients on providing information about the harms of smoking and assessments of nicotine dependence, readiness to quit and past quit history.

These results correspond to the most frequently observed BCTs, in England's Stop Smoking Services (Lorencatto et al., 2013a). In a fidelity study, Lorencatto et al. (2013a) observed that information gathering and assessments and addressing motivation were the most commonly implemented BCTs, whereas, maximising self-regulatory capacity or skills and delivery of intervention were implemented the least. This suggests a correspondence between the implementation of BI by providers in TB clinics in Pakistan with the stop smoking advisors in the UK.

Some of the BCTs in the least implemented group in Fidelity study, (i.e. emphasise choice, confirm client decisions, advise on changing routine, and facilitate on relapse prevention) have been found to be associated with higher quit rates in the English Stop Smoking Services (Michie et al., 2011b). Therefore, these would be considered important BCTs, which if not implemented in a smoking cessation BI could have implications on patient outcomes. Most of these BCTs were implemented in less than 10% of all BI sessions in the Fidelity study, indicating the need to further investigate the factors affecting their implementation. Such factors may include the patient's preferences, or provider competence and motivation.

It is also a possibility that some of these BCTs were used selectively by the providers depending on the needs of the patient. A quantitative measure of BCTs (that need tailoring to patient needs) using an adherence scale may portray a simpler image of a rather complex interaction. Therefore, it is possible that the provider was delivering a good quality intervention, which did not

necessarily comply with the ingredients outlined in the fidelity index, leading to a lower fidelity score of the delivery being noted. Future interventions designed for behaviour change should consider specifying the BCTs that must always be delivered and those that can be selectively delivered. Furthermore, distinguishing between the fixed vs. variable intervention ingredients in BI design (Hawe et al., 2004, Durlak and DuPre, 2008, Firpo-Triplett and Fuller, 2012), as discussed in 4.1.4 (chapter 4), might address this issue in quantifying BCTs using a fidelity index, in future.

Another factor promoting provider adherence to certain ingredients more than others might be related to the recording and reporting of BI ingredients as these were delivered. For example, the providers might have considered that the information gathering and assessment questions were more 'important' because they were required to record this information in the patient questionnaire. Although they also self-recorded their delivery of the triggers, withdrawals and the management strategies (the self-record checklist), they might still have believed that this was relatively not as important as the responses recorded from the patients. It is also possible that most of the providers in the Fidelity study found the least implemented BCTs 'difficult' to deliver. Effort or motivation is determined by both the importance of a task and the difficulty of achieving it (Wilson et al., 2015, Brehm and Self, 1989, Brehm et al., 1983). Importance of an action determines potential motivation (i.e. the amount of effort the provider is willing to put into intervention delivery), whereas, difficulty determines actual motivation (i.e. the amount of actual effort provider put into delivering the intervention) (Wilson et al., 2015).

Interestingly, offering written materials (information leaflet) would be considered an easy to implement BCT; however, it was only implemented in about 20% of the BI sessions in the Fidelity study. Similarly, 'setting a quit date' was often not discussed (only in 35% of the BI sessions) which is an important behaviour change strategy for smoking cessation. This may partly be explained by these activities being discussed prior to or post audiotaping of the BI session, thus limiting their inclusion in fidelity index at the stage of coding. It is likely, given the setting of intervention delivery that when the provider identified a smoker TB patient s/he confirmed their willingness to participate in the study including if they were willing to quit within a week. They might have noted down the quit date prior to starting the audiotaping of the session because they mainly perceived the flipbook as the BI session and not the entire consultation from greeting the patient to their exit from the clinic. To confirm this finding, the data from 'patient questionnaire' (Fidelity study) pertaining to setting the quit date was explored and it was found that actually 58% of the patients set their quit date as 'quit from today', 31% as 'quit within 5 days', 5% did not set a quit date and information was missing for 6% of the enrolled patients. The same would apply to the information leaflet (which was not recorded in the patient questionnaire) but it was designed to be given at the end of the flipbook interaction and the providers might have already stopped the audiotaping by that point, therefore, not capturing the

patient-provider interaction in its entirety. Post-analysis interviews with the providers might help understand some of these observations better.

An important observation from these findings is that some of the BCTs delivered by the majority of the providers were not stable items in the fidelity index (as discussed in Chapter 7) and were found to be redundant on PCA. These items (1 'smoking behaviour and history', 10 'assessing the current level of motivation to stop/ willingness to quit', 12 'Eliciting a prompt commitment from the smoker on starting the quit attempt' and 15 'nicotine dependency') relate broadly to information gathering and assessments. These items were unable to describe enough variation under the fidelity construct and therefore might not be good indicators to represent fidelity of a BI. Whether these ingredients are also not effective BCTs in the interplay of causal mechanisms of behaviour change for smoking cessation, should be explored in future research.

8.5.3 Consistency in providers practice over time

Providers' relative performance concerning BI delivery was found to be broadly consistent between the time-points of the two studies, to allow retrospective linking of fidelity scores with quit rates. This was based on the analysis of their self-record checklists completed in the two studies at T1 and T2. Although their practice was found to be consistent between the two time-points, a few limitations of using data from a self-record checklist merit discussion here.

First, the self-record checklists are not objective measures and are less reliable than direct methods of observation of provider behaviour (Carroll et al., 2000). Data obtained from such measures show low correlations with objectively captured data. However, these are considered to be good tools for supplementing objective data in evaluative research and for triangulation in analysis (Borrelli, 2011). In this study, the practice scores from the self-record checklist were used to test the degree of consistency, in terms of the relative ranking (number of strategies offered per trigger/withdrawal) of intervention options by the same providers at the two time-points. As the scores are repeated measurements from the same providers, there are less chances of occurrence of a systematic error even if the providers knowingly self-recorded better practice than in the past.

Second, self-record checklists could have variable completion rates and, when completed in patient-provider interaction settings, they have the potential of being completed hurriedly (Bond et al., 2000b). Particularly, looking at the self-record checklist (Table 8.1) in this study, providers checked (tick marked with a pen) those triggers or withdrawals that the patient had experienced, along with the strategies advised to them. When delivering a BI using an interactive flip-chart and completing the checklist at the same time, or after session completion, could potentially lead to missing information. Moreover, if the provider did not tick an item on the checklist (or forgot to do so), it wasn't possible to confirm whether that item was not delivered or that it was missing information. However, in testing the concordance in providers' relative performance, no

systematic differences were found between the two time-points that could have potentially invalidated the retrospective linking of data.

8.5.4 Association of fidelity scores with quit rates

Patient sample characteristics for the two studies were broadly similar, apart from 'attempted quit in the past', which was higher in the Fidelity study. This might be because these patients were on average 2 years older than the patients in ASSIST study and were smoking for a longer duration (on average 3 years) so potentially had more opportunity to try quitting smoking. The higher attempts at quitting could also be due to the change in the target population's awareness about tobacco use during the four years in-between. However, at the time of study there were no tobacco cessation services being offered at the respective TB clinics or any known (to the providers) tobacco-related campaigns in the area. This could have been explored better through patient exit interviews or context analysis, which was outside the scope of the current research.

Better Quality of interaction with the patients could be a non-specific effect, similar to the placebo effect, where higher patient quit might be due to the higher satisfaction on part of the patient. However, it is unsurprising that the effectiveness of the BI (higher smoking quit rates) was strongly influenced by the *Quality* of interaction with the patients in this study, given its development in counselling and clinical practice (Hardcastle et al., 2016). Quality of interaction, as measured using the fidelity index, was based on the inter-personal (Roth and Pilling, 2008) competences noted in section 4.1.3 (chapter 4), for delivering a BI for smoking cessation (Michie et al., 2011b). A recent systematic review and meta-analysis of intervention fidelity for Motivational Interviewing (D O'Halloran et al., 2014), found that interventions with higher implementation of inter-personal BCTs produced larger effects on behaviour change than those with lower implementation of these BCTs. These data support my study findings and provide initial indication that the Quality of interaction with the patient might be paramount to the success of a behaviour change intervention for smoking cessation.

The second important main effect in my model was the level of *Adherence* of the provider to the BI content; the increase in which had a negative influence on patient quit, even when adjusting for Quality of interaction. This unexpected finding might be due to any of the several reasons: it might be a spurious finding given the deficiencies in study design, or if it was internally valid it might not be externally valid, for example, it might just be attributed to an anthropological difference in the study population. Alternatively, it might be true that increase in the levels of Adherence, which means higher number of ingredients delivered (as measured using the fidelity index), actually reduced the effect of BI on patient quit.

It might be hypothesised that due to the complexity of the behaviour change interventions, asking the providers to do too many things and deal with too many dimensions all at the same time, might possibly make the intervention longer and a tedious task (Boutron et al., 2008,

Sereno et al., 2012). Interventions implemented with higher levels of adherence have been shown to be associated with improved outcomes than those with poorer adherence, in other areas (Durlak and DuPre, 2008). A fidelity study of smoking cessation behavioural support (Lorencatto et al., 2013a) suggested no association of intervention adherence with the patient quit; it was a small sample from two smoking cessation service practices in the UK. However, a recent meta-analysis of intervention theory suggests that "***more is sometimes less***" when it comes to promoting health behaviour change (Wilson et al., 2015), including smoking, diet and physical activity. The review (of 150 reports) found that a moderate number (as opposed to low or high number) of ingredients in a BI produced the highest level of change in behaviour. It suggests that an optimal number of ingredients would be low enough to prevent disengagement while being high enough to ensure the necessary level of motivation and effort to maximize compliance, and ultimately, change behaviour. Furthermore, the review estimates that intervention efficacy begins to decrease when interventions include a large number of behaviour change elements.

This could also potentially make the intervention complex for the patients to understand and absorb the main (or important) take home messages, which should ideally be simpler. A large number of BCTs in an intervention are undoubtedly more demanding (Nigg and Long, 2012) and may possibly push the human limits of cognitive capacity and self-control (Baumeister et al., 1994, Muraven and Slessareva, 2003); intervention effect, in such cases, may either plateau or decrease when more behaviour changes are required (Ornstein et al., 1993). In future studies, it might also be interesting to know the extent to which factors such as patient educational level might affect this phenomenon, given the patients in Fidelity study were from rural parts of Pakistan with lower than average global literacy rates (UNESCO Institute for Statistics).

Finally, a negative interaction was found between Adherence and Quality. This interaction was explored out of interest to see whether the two co-variates interact because of the positive correlation between them (Table 8.10) and it was found that higher levels of Adherence reduces the effect of Quality on patient quit. While the interaction between content-based BCTs and inter-personal BCTs in the effectiveness of BIs has been implied in behaviour change research (Hardcastle et al., 2016), this negative relationship between the two, in terms of smoking cessation outcomes, has not been suggested in the past. As noted above, it might be argued that increase in Adherence or higher number of ingredients being delivered, reduces the ability of high Quality of interaction with the patient to positively affect their smoking quit, that is, volume or complexity 'crowds out' the effect of a good Quality interaction on quitting. Nevertheless, this does not mean that implementing higher number of BI ingredients compromises the Quality of interaction but that it attenuates the impact of Quality on patient quit, even though the Quality might have remained the same. Furthermore, 100% adherence to BI content has been contested, in the past, as a desirable goal for achieving high fidelity (Lorencatto et al., 2013a, Borrelli, 2011). A sample of English Stop Smoking Services, with high and moderate quit rates respectively, showed similar levels of adherence to BI content across the services (Lorencatto et al., 2013a).

The same study suggested that the quality of delivery does not necessarily mean higher adherence to content. The service with lower adherence to content (56%), as opposed to the service with higher adherence to content (69%), was potentially more effective in achieving higher patient quit. This was because the service with lower adherence to content actually had better quality of delivery (implementing 10 BCTs on average compared to 7 in the other service).

8.5.5 Relationship between Adherence and Quality scores

Providers Quality of BI interaction with the patient was positively associated with their Adherence to BI content. This suggests that the providers with better Quality of interaction (based on inter-personal BCTs) were more likely to Adhere to the BI content. However, this finding is contrary to the reports on these two aspects of fidelity (Adherence and Quality), showing low correlations between the two behaviours (Borrelli, 2011, Miller and Binder, 2002, Perepletchikova and Kazdin, 2005). This might be due to the way competence to deliver the intervention (presence or absence) was defined and assessed in these studies; compared to the use of taxonomies to specify content and the behaviourally anchored three point response scale used to quantify Quality in the fidelity index, in the current study. Empirically adherence to content and competence to effectively deliver an intervention might be considered independent dimensions of fidelity but, in practice, these were found to correlate quite highly in the Fidelity study. The extent to which these two aspects of fidelity correlate and their intermediary role in the mechanism of action of a BI needs to be examined in future research, with appropriate study designs and a representative sample.

The hypotheses (research questions) generated from the preliminary analysis in this study that could help direct future research, in connection with smoking cessation, are outlined here and discussed further in Chapter 9.

- What is the optimal number of ingredients (hence the BCTs) to target in a behavioural intervention for smoking cessation to achieve successful patient quit?
    - To determine whether briefer versions of the behavioural intervention tailored to the patient's needs can achieve similar results to strictly adhered versions of the behavioural intervention
    - To determine the variable (content that might be tailored) and the fixed (process of delivery) features of the behavioural intervention
- Whether quality of interaction can be improved to an extent that it is effective in achieving successful patient quit, independent of the number of ingredients delivered?
    - To determine the balance between adherence to behavioural intervention content and the quality of interaction where both complement each other's effect

- To investigate whether adherence is a mediator or a moderator in the behaviour change intervention-outcome model
- To investigate whether quality is a mediator or a moderator in the behaviour change intervention-outcome model

- Whether adherence to behavioural intervention content is effective in achieving higher quit rates, independent of the duration of the patient-provider interaction?
  - To determine if the duration of the patient-provider interaction predicts patient smoking quit
  - To determine the relative weighting of each item (ingredient) based on the time taken to deliver each ingredient and its complexity (i.e. the number of BCTs included).

8.5.6 Limitations and challenges

This study had several limitations that merit discussion in this chapter; some cross-cutting issues on internal and external validity between studies A, B and C are discussed in the next chapter.

Firstly, there were very few data-points (18 provider records), as cluster-level data were used in study, reducing the power of analysis. Moreover, the analysis was secondary and should be considered as exploratory in nature, rather than predictive of the association between fidelity for BI and patient smoking quit. Lastly, the study used data from a cross-sectional design (Fidelity study), and retrospectively linked it to an RCT design, thus limiting the establishment of causality and temporality of the derived effects. These points are discussed below.

Few data-points

The study was restricted to a small sample of providers as I wanted to investigate fidelity in the same providers as the ASSIST study to explore patient quit, which risked introducing Type II error in the analysis results. As noted above (see 7.5.6), I was unable to collect prospective data on quit rates in the Fidelity study, which could have been analysed at the individual level, accounting for clustering; giving the analysis more power and less chance of Type II error. In addition, I would have used the prospective quit rates to validate the retrospective quit rates (from the ASSIST study) by providers and analysed if provider fidelity contributed to the wide variation seen in quit rates, by teasing out those ingredients that worked and those that did not.

Exploratory nature of the analysis

Study C used data on observing fidelity in one set of patients (Fidelity study) and assumed (based on concordance in practice scores of providers) it is the same on a different set of patients from the past (ASSIST study), although the providers were the same individuals. Such analysis cannot

establish causality, and interpretations should be restricted to exploratory purposes, because this assumes that there was a manipulation of the independent variable (i.e. fidelity). This is an issue in general with cross-sectional designs, where the same individuals are not followed up longitudinally to allow assessment of direct manipulation of the independent variables, limiting the establishment of causality and temporality. I chose this design, because at this stage, there were outstanding questions to be resolved about the reliability of the new fidelity index (and its refinement), and the generation of hypotheses about the contribution of fidelity (whether it moderates or mediates) to intervention effect. The larger resources required for a prospective study could not, therefore, be justified. Moreover, a strength of the study was the use of a range of methods to extract different types of information, and to corroborate findings; allowing increased confidence in the findings from weaker study designs (Campbell et al., 2007b).

Establishing causality and temporality

Drawing causal inferences from the secondary investigation was further complicated by the retrospective linking of fidelity to quit rates. Whilst retrospective data linkage requires fewer resources, it limits the options for measurement and for selection of comparison groups.

As briefly noted before (in 8.5.3), the practice scores that were used to test this assumption were self-recorded by the providers and these might be subject to reporting bias. Providers may misrepresent their past behaviours and choices in order to appear more consistent with their current practice. Such retrospective linkage can introduce bias in the study estimates, especially when there is state dependence in respondent's choices (Shachar and Eckstein, 2007). Specifically, observed persistence in retrospective data may be due to (a) true state dependence, (b) unobserved heterogeneity, and (c) retrospective bias in reporting previous practice (Heckman, 1981).

True state dependence, where behaviour relevant to future choices is altered as a consequence of experiencing an event, is unlikely in the case of providers filling the self-record checklist, as they were not intervened for skills training or professional capacity building on BIs between the two time-points. The unobserved heterogeneity is seen when individuals vary in certain unmeasured variables which could lead to a conditional relationship between future and past experience. However, in my study this type of state dependence is also very unlikely, given that the providers are the same individuals delivering the same intervention at the two time-points. Any unmeasured variables are not likely to be time dependent. Even though self-reporting bias could lead to non-concordance or lower concordance in ordering of the relative performance of the providers between the two time-points, it is less likely to be an effect of being observed (audiotaped) for intervention delivery (Adams et al., 1999).

8.6 CONCLUSION

The 'fidelity index', developed (in study A) and validated (in study B) using suitable methods, was implemented to obtain fidelity scores for exploring variation in quit rates (study C). It is possible to reliably assess the fidelity (i.e. both Adherence and Quality) for a BI of smoking cessation and relate this to patient quit by providers.

The negative association of Adherence (to BI content) with quit rates, which means delivering a higher number of BI ingredients reduced patient quit, implies that BI content might be simplified to maximise effective quit by optimising the number of behaviour change ingredients included in it. The positive effect of Quality of interaction of a BI suggests that it might be crucial to patient quit and may be as important as the number of BI ingredients delivered. My work suggests, and I must quote Trevor Sheldon, "***it ain't what you do but the way that you do it***" (Sheldon, 2001), that acknowledges the increasing complexity of variation in outcomes of effectiveness studies.

PART IV- DISCUSSION

In this final section, I discuss the strengths and limitations of the three studies. I also describe the methodological issues relevant to the development of a new index and its use for process evaluation. I end the thesis with implications of my work for research and practice and the future research directions.

Chapter 9. Discussion and Conclusion

In this chapter, I discuss the overarching issues with the design and the methods used in the three studies (study A, B and C). I provide a brief synopsis of key findings from these studies, and discuss their strengths and limitations using literature from previous pertinent work in the field. I then outline the crosscutting implications of developing a reliable measure of fidelity for a behavioural intervention and using it to explain variation in outcome for policy, practice and future research.

In Appendix A, I summarised the evidence on association between smoking and TB and appraised the quality of this evidence based on GRADE scoring. At the beginning of my research, my focus was on epidemiology of smoking in TB patients. However, as my work progressed and I started gaining in-depth understanding of the behaviour change interventions for smoking cessation, I realised their significance in TB patients in LMICs and the knowledge gaps in optimal implementation of proven effective treatments. This is why I put the review on 'smoking in TB patients' in appendix to this thesis because my focus has changed to intervention research and fidelity to BIs and it is better read as standalone evidence synthesis.

In the first half of the thesis I described the potential significance of offering low cost smoking cessation interventions in TB patients and the relevance of integrating these in TB care in LMICs. I explored the literature and identified gaps in the evaluation of BIs for smoking cessation and limitations in their description of what works and how it works. I also provided the justification for developing a method that could help identify the 'active ingredients' of BIs, by capturing a complete and a clear account of patient-provider sessions. Identification of the ingredients of a BI that were delivered and how these were delivered to the patients and estimating their influence on the patient quit can be vital to the potential uptake and implementation of successful BIs for smoking cessation, in routine TB practice.

My overall research goal was to investigate whether smoking quit rates were associated with the fidelity with which a BI for smoking cessation was delivered. Three studies were carried out to achieve the following specific research aims:

- Development of an index to measure fidelity to a BI for smoking cessation
- Validation of the fidelity index by assessing its psychometric properties
- Exploration of whether fidelity to a BI explains variation in quit rates

Study specific findings were discussed in previous chapters (chapter 6 - 8), and in the Appendix A for the systematic review. In this chapter, I describe the strengths of my research and discuss the overarching issues in the design and methods across the three studies and their limitations and then present my overall conclusions.

9.1 SUMMARY OF RESEARCH AND FINDINGS

In the absence of pre-existing tools to quantify fidelity for BIs, I brought together the latest advances in behaviour change science and psychometrics to produce a reliable measure of fidelity and investigated its ability to accurately capture intervention fidelity. Key findings of this research are summarised here.

9.1.1 Smoking in TB patients- evidence synthesis

The evidence synthesis on 'smoking in TB patients' indicates that tobacco smokers were twice as likely to develop TB disease, suffer recurrent TB and die from TB, compared to non-smokers. However, the evidence on the association between tobacco smoking and acquiring TB infection was found to be weak.

9.1.2 Study A- Development of the Fidelity Index

Critical components technique was used to generate items of the fidelity index from the BI used in the ASSIST study that were mapped using the taxonomies of BCTs. Forty-five items- 37 for Adherence, and eight for Quality- were identified to represent 'intended practice' of the provider. A three point ordinal response scale that was behaviourally anchored rated each index item on fully, partially, or not, implemented, to represent 'actual practice' of the provider. Content validation by three experts showed a high level of agreement on all items for inclusion in the index. A pilot of the index resulted in re-phrasing and re-defining some of the behavioural anchors for the items scale. This was then consolidated into a new index for measuring intervention fidelity for a BI, which consisted of two main dimensions of 'Adherence to content' and 'Quality of interaction'.

9.1.3 Study B- Psychometric properties of the Fidelity Index

The index was validated psychometrically and shown to be a good method for capturing significant variation in provider fidelity. The newly-developed fidelity index was found on inter-coder reliability testing to be highly reliable for application by different coders. Moreover, the items of the index were coherent in measuring fidelity, as assessed by Principal Components Analysis. The two sub-indices (Adherence and Quality) could be used in isolation or in combination, as assessed by PCA performed separately for the two sub-indices and together under the overall fidelity construct. The Adherence sub-index clustered around five domains that map to five of the seven key features of the BI ('information about harms of smoking/benefits of quitting'; 'assessments of dependency'; 'preparation to quit'; 'management of triggers'; and 'withdrawals'), showing its ability to cluster BI content into its key aspects of behaviour change. In addition, Generalisability study showed it to reliably differentiate between providers by capturing the variation in their practice concerning BI delivery.

9.1.4 Study C- Explaining Variation in Quit Rates

Overall in the BI sessions audiotaped for the patient-provider interaction, information gathering and assessments of nicotine dependence and past quit were implemented to a much higher degree than the triggers and withdrawal management and the Quality aspects. Based on the analysis for the assumption of consistency in provider practice (between Fidelity study and ASSIST study time-points), fidelity scores were retrospectively linking to quit rates.

Regression analysis indicated that patient quit from the ASSIST study was positively associated with the same providers' Quality of interaction but negatively associated with their Adherence to BI content, measured four years later. A negative interaction of high magnitude was also found between Adherence and Quality showing that an increase in one of these two covariates would in turn decrease the other's effect on the quit rate and vice versa. I hypothesised that too many BI ingredients implemented in an intervention might negatively influence patient quit. Further regression analysis of Adherence with Quality suggested that the providers with better Quality of interaction (based on inter-personal BCTs) were more likely to Adhere to the BI content.

9.2 STRENGTHS OF THE RESEARCH

The strengths of this research are the development of an index for measuring fidelity of BI delivery that did not exist before, and its validation and application to explore variation in quit rates for the first time. This research uses a sound methodology by bringing together knowledge of behaviour change science and psychometrics to quantify the fidelity of BIs. This is the first time that the taxonomies of BCTs have been applied to develop an index for measuring fidelity. This is also the first time that provider fidelity for delivering a BI for smoking cessation has been studied in a LMIC, as part of routine TB care practice. Finally, this research informs the methods that can be used for process evaluation of BIs as a part of effectiveness evaluation, where the end-point is not the process outcomes (i.e. fidelity) but is the impact of these process variables on patients' quit rates. These strengths are discussed below.

9.2.1 Originality

TB and tobacco review

Due to the lack of direct evidence of smoking cessation on health outcomes (like TB), I conducted a systematic review on the association of TB outcomes with smoking (as noted at the beginning of chapter; Appendix A). I used the GRADE criteria, which enables a systematic and transparent process to assess the quality of evidence and strength of association for a particular outcome across studies and reviews (GRADE Working Group, 2004). As the review included observational studies, the GRADE checklist for quality assessment was applied to each TB outcome using the Bradford Hill's criteria (Meader et al., 2014).

To my knowledge, this is the first systematic review to provide strong evidence (of good quality) for the association of smoking with TB treatment outcomes including TB recurrence and mortality. It supports evidence from previous reviews on the strong association between smoking and developing TB disease. However, in contrast to the findings of the previous reviews, it suggests a weak association of smoking with latent TB infection (based on the quality of the studies).

New fidelity index

I was not able to identify an existing fidelity index or a scale that could be adapted or adopted for use in my investigation. Therefore, I conducted phased research to first develop a measure of intervention fidelity and establish its psychometric properties and then to implement the index to obtain fidelity scores for explaining quit rates.

Whilst methods for labelling and coding of BIs using standard practice manuals for smoking cessation (Lorencatto et al., 2013a, Lorencatto et al., 2013b, Michie et al., 2011c, West et al., 2010) and describing fidelity in general (Dane and Schneider, 1998, O'Donnell, 2008) existed, a measure that would quantify fidelity based on intervention theory did not exist. It is one thing to assess whether or not BI ingredients have been delivered and another to assess how well they have been delivered (Lorencatto et al., 2013b). So far, the focus on assessing fidelity in health services research and behaviour change science has mainly been on adherence to intervention ingredients or BCTs (presence or absence) (Lorencatto et al., 2013a, Hagger and Hardcastle, 2014, Hardcastle et al., 2016). Recently, efforts have been made to develop methods for quantifying how well BI ingredients are delivered and move closer to measuring fidelity to delivery rather than only assessing adherence to content (Hardcastle et al., 2016, Farmer et al., 2012, Tober et al., 2008).

In existing indices (Young and Beck, 1980, Moyers et al., 2005, Lane et al., 2005), the focus has been on the overall adherence to BIs as stipulated in intervention protocols or practice manuals; these fidelity tools did not break BIs down into specific techniques (of behaviour change) and were, therefore, not fit-for-purpose means for the isolation and classification of BI ingredients. Furthermore, these tools were not designed to capture enough variation in the delivery of each intervention ingredient because of having nominal response scale options rather than ordinal that can be used to quantify the technique using behavioural anchors. The index developed in this study brings together two important aspects of fidelity (Adherence to BI content and Quality of interaction) in a single tool and teases out the specific behaviour change function of each ingredient using the taxonomies of BCTs (content- and competence- related). It is the first of its kind to enable quantification of these aspects of fidelity by using the behaviourally anchored response scales created for each individual item in the index.

Exploring variation in outcomes

To understand how complex interventions achieve certain effects (insight into the 'black box' of the intervention) is possible by evaluating process variables with the outcomes (Hulscher et al., 2003). Measuring fidelity is especially important in multi-site studies, where findings may vary in the magnitude of effect between sites (Bond et al., 2000b), as they did in the ASSIST study.

Interventions implemented with higher fidelity have been shown to be associated with better outcomes than those with poor fidelity in other fields (Durlak and DuPre, 2008). However, a small sample of two English Stop Smoking Services with high and moderate quit rates showed similar levels of fidelity across the services (Lorencatto et al., 2013a). That study assessed fidelity using treatment manuals against patient-provider audiotapes and defined fidelity according to the criteria: <50% adherence to manual as 'low' and 80-100% adherence as 'high'. My study is the first to investigate whether fidelity (level of Adherence and Quality of interaction) is associated with quit rates. It indicates that the Quality aspects of fidelity might be as or more important in the causal mechanism of behaviour change for successful quit in the patients. The extent to which differences in fidelity may help explain variation in quit rates needs to be examined in future research.

My work can pave the way for future studies that tap into the mechanisms of behaviour change by identifying the individual BI ingredients that work and the likely candidate mediators. The fidelity index could potentially be used to identify the active ingredients of a BI. In addition, it provides a tool that researchers can utilise to measure fidelity for predicting outcomes of complex BIs.

9.2.2 Rigour

With guidance from the MRC framework for process evaluation, I brought together existing fidelity literature and the BI specification taxonomies, to inform the methods of my research. I used an interpretive evaluation model (Figure 1.2) to first capture the active ingredients of a BI for smoking cessation in a format that would allow scoring based on fidelity and then apply these scores to explore variations in patient quit rates.

Methods used to develop the fidelity index

The main strength of this work is the rigour of methodology to combine behavioural sciences and quantitative scale development methods in designing a reliable new measure of intervention fidelity. My work was broadly based on behaviour change for smoking cessation; it builds on and extends previous research on fidelity assessment of BIs (Lorencatto et al., 2013a, Lorencatto et al., 2013b, Michie et al., 2011b, Michie et al., 2011c).

The taxonomies of BCTs (Michie et al., 2011c, Lorencatto et al., 2013b, Michie et al., 2013), although used in previous studies (Lorencatto et al., 2013a, Lorencatto et al., 2013b) to describe BIs 'compositionally' have not been used to score the fidelity of BIs in relation to the 'functionality' of its features. This later concept surpasses the simplicity in coding 'present vs. absent' for fidelity assessment and is a step forward in establishing a method for assessing the quality with which BCTs are delivered. In this study, I used the BCTs taxonomies to specify the BI content and therefore integrated the scale responses, based on process of delivery with each specified ingredient. Such integration or assimilation of an ordinal scale in a well-grounded behaviour change theory enabled the capture of the full account of fidelity's contribution to behaviour change. Further, psychometric testing of the index supports this concept and suggests that the items of the index seem to measure fidelity, as captured via the response scale.

Methods used for data collection

The primary strength of the Fidelity study was that it generated data grounded in direct observation of patient-provider interaction sessions (Bell and Kravitz, 2014). Survey interviews and record reviews or exit interviews (as proxy measures of patient-provider interaction) often do not agree with each other or with direct observations of the actual visit (Carroll et al., 2000).

A related strength is that patient-provider interaction was studied *naturalistically* because the BI sessions were delivered by the TB providers in their clinics during routine TB care of the patients, offering the opportunity to capture patient-provider interactions independent of the research. This is in sharp contrast to most experimental intervention studies, in which independent variables are manipulated to create the situations studied (Bell and Kravitz, 2014).

Furthermore, the Adherence and Quality scores (from the Fidelity study) were free of reporting bias, as these were obtained by rating fidelity using audiotapes of the patient-provider BI sessions. Similarly, the quit rates (from the ASSIST study) were based on CO reading rather than self-reports giving objective measurement of patient quit status at 6 months post-intervention.

9.2.3 Importance

TB and Tobacco review

The findings from the evidence synthesis on smoking in TB patients will inform policy makers and TB managers of the importance of smoking cessation in TB patients. This might particularly be useful when drawing conclusions from evaluations of smoking cessation interventions that usually only report quit outcomes and not the differential impact on TB outcomes.

Application of behaviour change science and intervention delivery in the LMICs

Although the BI used in ASSIST study was found to be highly effective in helping patients successfully quit smoking in TB clinics (Siddiqi et al., 2013), the taxonomies of BCTs have not been used to code interventions used in the context of countries like Pakistan. Whether these taxonomies reliably code BIs outside of the English Stop Smoking Services is yet to be established. However, as noted in study C (8.5.2), the most frequently delivered BCTs (information gathering and assessments, and addressing motivation) and the least delivered BCTs (maximising self-regulatory capacity or skills and delivery of intervention) by the providers in TB clinics in Pakistan correspond to the implementation of the BI by stop smoking advisors in the UK (Lorencatto et al., 2013a, Lorencatto et al., 2013b). This suggests that the implementation of BIs for smoking cessation, the provider fidelity to delivery and related challenges in the LMICs (like Pakistan) might be very similar to the HICs.

Use of process variables as intermediate factors to explain variation in outcomes

Fidelity has been referred to as the 'methodological strategy' used to enhance the reliability and validity of complex interventions (Bellg et al., 2004). The revised MRC guidance on development and reporting of complex interventions (Craig et al., 2008) recommends integrating process and outcome evaluations (Oakley et al., 2006a), tailoring to local contexts (Campbell et al., 2007a), and making greater use of the insights provided by theory of complex adaptive systems (Shiell et al., 2008). The guidance recommends achieving high fidelity in a pilot phase before proceeding to a full RCT, which however, does not happen in real-life, given the limited resources (for a full pilot phase); resulting in variations in delivery of the intervention or infidelity in RCTs of complex interventions (Knittle, 2014).

The variation in outcomes, as seen in the ASSIST study, are common in evaluations of complex interventions and most researchers attempt to explain this variation by factors related to study design and measures of effectiveness, paying little attention to further investigation of the process of intervention delivery (Hardeman et al., 2008, Bellg et al., 2004, Borrelli et al., 2005). Process evaluation, which means exploring the way the intervention was delivered, could provide valuable insights into variable impact. The MRC framework for process evaluation (Moore et al., 2015) outlines formative or developmental research approaches (Carroll et al., 2007, Linnan and Steckler, 2002, Grant et al., 2013), which focus mainly on refinement of novel interventions, considering fidelity as a process outcome. My research, however, focuses on informing a method for using process variables, such as fidelity, as the intermediate factors in explaining outcomes of effectiveness of complex interventions.

In effectiveness evaluations, fidelity is considered as an intermediate factor with the sole purpose of explaining other important outcomes like smoking cessation, which are the prime target of the intervention (Brownson et al., 2012, Proctor et al., 2011). I therefore adapted the key functions of

process evaluation defined by MRC (as described in Chapter 1, Figure 1.2 on page 29) and applied interpretive rather than formative research approaches to develop a mechanism to quantify fidelity and use it for investigating variation in quit rates. These concepts are described in chapter 1 (see 1.5.2), and further recount is given here.

Interpretive research illuminates the mechanism of action of intervention theory and enhances understanding of its impact or worth (Stetler et al., 2006). This is the first step that enables evaluators to move beyond description of fidelity (Moore et al., 2015) and use it as an independent variable that contributes to the effectiveness of the intervention (Peters et al., 2013). I conducted a post-trial process evaluation of the BI for smoking cessation used in the ASSIST study to interpret the wide variation in quit rates, and proposed a system beyond describing intervention fidelity that could potentially identify 'active ingredients' and measure the contribution of fidelity to intervention effectiveness.

9.3 LIMITATIONS OF THE RESEARCH

The study specific limitations have been discussed in previous chapters (for study A, B and C). The crosscutting limitations affecting the internal and the external validity of the overall research are discussed here.

9.3.1 Internal validity

The index was based on an extensive review of the literature on psychometrics and concepts of behaviour change. In choosing which areas of the literature to review (and which to exclude); I pre-determined the form the index was likely to take. For example, although I considered that patient responsiveness and receipt in a BI was important, I restricted my review to elements of fidelity directly targeting provider practice and attitude in delivering the intervention, as that was the focus of my research. Moreover, I did not consider the wider literature on the theory of the psychological processes involved in changing behaviour that might not be covered fully by the taxonomies of behaviour change. Incorporation of these additional theories could further refine the fidelity index (in future) and improve its practical value. Particularly, the 'Quality of interaction' aspects relating to provider 'competences' (Michie et al., 2011b) or 'relational techniques' (used in Motivational Interviewing) could be further improved, as the number and description of behaviours and techniques is expanded in the future (Hardcastle et al., 2016). In addition, new BCTs are emerging very quickly and more comprehensive taxonomies might be produced (Michie et al., 2014-2017) in the future that can be used for defining the items in the fidelity index more precisely and for adding new items. However, using the current taxonomies there might be a chance that some important behaviour change features are missing from the index.

A limitation in the research design overall is the inability to control for confounders contributing to the variation in quit rates. There are multiple reasons for this, including: fewer data-points not allowing multi-variable regression; observations of study subjects without contemporaneous controls; and no data on context-related factors that might influence intervention delivery. In my study, a number of potentially confounding sources could include the factors extrinsic to the intervention, like the attributes of the provider, the wider context at the meso-and-macro level or factors related to patient responsiveness. In determining the potential association of fidelity with quit rates, therefore, I could have overestimated and/or missed effects. However, some of these factors at the provider-level were explored as part of the ASSIST study (described in chapter 3; Table 3.1 and Table 3.2), which were found not to have contributed to the variation seen in the quit rates. The broader context level factors could be explored using validated tools for context evaluation (Bergström et al., 2015), in future studies.

9.3.2 External validity

Another limitation of the study deals with the generalisability of the results. The Fidelity study was conducted in TB clinics (situated in primary and secondary health care centres) in two districts of the Punjab province in Pakistan. Using a resource intensive method like direct observation of patient-provider sessions is not practically possible in a national sample of providers and patients (Bell and Kravitz, 2014). Nonetheless, I feel that the provider practice in this multi-centre study was structurally similar to TB clinic settings in other LMICs. Furthermore, the implementation pattern of BCTs indicated that provider practice behaviour in this study was quite similar to the stop smoking advisors in the UK (as discussed in chapter 8). However, the findings from this research are less likely to be representative of people's behaviour, cultural practices and attitudes towards smoking. Ultimately, this weakness can be managed through the replication of research by investigators across diverse settings.

As I was interested in variation between providers and not just a high level of fidelity, this might have been affected by providers showing low fidelity or scoring on those items that were least implemented by the majority of the providers. The results from the Generalisability study (see 7.4.6) showed that a considerable amount of variation in the fidelity measurement could be attributed to the providers. Further sensitivity analysis, excluding those providers with very low fidelity, retained the same results. Overall, the fidelity measured in this study for most ingredients was low to moderate. BI ingredients relating to 'setting quit date', 'offering BI leaflet', 'explaining expectations regarding the intervention programme' and 'providing reassurance' were negligibly implemented by the providers. However, these were not excluded from the index as these might be important in fidelity measurement (of other BIs) and should be evaluated in future research to make the approach more generalizable.

A further threat to the representativeness of the study is the possibility of reactivity (Bell and Kravitz, 2014), also called the 'Hawthorne effect' or 'observer effect'. The concern is that the

patients and the providers might alter their interaction behaviour, knowing that they were being observed. In comparison to the ASSIST study, the current study observed patient-provider BI sessions by audiotaping, which could have resulted in improved providers' fidelity (Bell and Kravitz, 2014). Therefore, these sessions may not be representative of typical practice, although research shows that the effect of being observed (passively) wears off quickly (Bell and Kravitz, 2014). In addition, the observation that BCTs as simple as 'offering a BI leaflet' were not implemented might suggest that the providers were less likely to be subject to the Hawthorne effect. However, these sessions are likely to represent a best case scenario, and therefore overestimate rather than underestimate fidelity (Lorencatto et al., 2013a). Moreover, the competences to deliver the intervention are unlikely to be affected by observation, even though adherence to BI content might improve under such circumstances. In the case where all providers improved their fidelity to BI compared to the past practice, this would lead to dilution of effect with quit rates which are from the ASSIST study, when providers were not being observed.

9.4 APPLICATIONS OF THE FIDELITY INDEX IN PRACTICE AND IN RESEARCH

A fidelity index, if designed appropriately and found to be reliable and valid has multiple potential uses, not only as a research tool but also in policy and practice. Some uses of the fidelity index to inform research and practice from defining interventions, to keeping track of future adaptations, assessing adherence, understanding the 'black box', to refining theory, are highlighted here.

9.4.1 Defining intervention

Use of a fidelity index offers numerous advantages for smoking cessation practices. The fidelity index might help optimise a BI for scale up in LMICs where, due to resource limitations, it is desirable to integrate cessation within existing programmes. In these settings, the policy makers and public health managers are often more interested in a variety of options for active intervention ingredients that can be delivered in a shorter time and yet retain an acceptable overall effectiveness of the intervention.

If used in routine practice, on TB surveillance data, it might help local practices to link implemented intervention ingredients with the patient outcomes later in their treatment course and to tailor delivery of BIs for smoking cessation for the respective settings. Although routine monitoring of cessation services cannot establish causality, descriptive information about changes in the intervention ingredients delivered and observed outcomes over time can be useful in identifying trends that co-occur.

9.4.2 Keeping track of future adaptations

In psychiatric rehabilitation research, fidelity measures have been used to map out the theoretical domains of the intervention and also to define model adaptations or modifications from the original intervention (Bond et al., 2000b). Hence, following the same model, the newly developed fidelity index for BIs of smoking cessation can also be used to define and describe adaptations to original intervention ingredients. Adaptations or deviations would be easier to identify and record using the fidelity index which links each item to the underlying model of behaviour change. For instance, the fidelity index could be used to differentiate between individuals given less intense behavioural support, compared with those given more intense treatment and take a more pragmatic stance to reporting findings, rather than controlling conditions to achieve high fidelity to an intense BI programme.

9.4.3 Monitoring and training

Fidelity assessment can be used to make the content of time-limited interventions more structured and focused (Lorencatto et al., 2013a), as opposed to practice manuals. The use of fidelity indices can be extended to the optimal scale-up of interventions nationwide, where implementation problems such as widely disparate services between administrative levels might occur, often falling short of the original effectiveness of the intervention as a result (Bond et al., 2000b). Fidelity measures can also function as self-recording tools by the providers of cessation services to monitor and document intervention adherence overtime.

In addition to the intervention ingredients, the fidelity index could be used to identify specific skills and competences of providers for delivering BIs. The Quality of interaction aspect of the index can assess the weaknesses and strengths of the providers, in terms of their behaviour change competences. This could facilitate and help direct the focus of provider training on specific skills and competences (Lorencatto et al., 2013a).

9.4.4 Understanding the 'black box' of the intervention

Behavioural change interventions often face the challenge of being translated from complex behavioural mechanisms into routine clinical practice by health care providers who are not very familiar with the science of behaviour change. Often, providers of such complex interventions are left with the resource for delivering the intervention and a list of activities. A fidelity index can potentially be used as a research dissemination strategy (Bellg et al., 2004), by giving the providers a deconstructed map of activities linked to behaviour change mechanisms that could help them understand why they are doing each activity and how they might tailor it to the patient needs, without losing focus of the mechanism of change.

Explaining variation

Assessing intervention fidelity gives researchers more confidence in their findings from effectiveness evaluations (Bellg et al., 2004). Whether the associations found are plausible or not influences causality according to the Bradford Hill approach (Schünemann et al., 2011). Determining fidelity could help establish causality for the clinically or behaviourally plausible outcomes. If an evaluation of a new intervention shows significant results but fidelity is not monitored, it is difficult to establish whether the outcomes were due to effective intervention content or other factors that may have been unintentionally added to or omitted from the intervention (Cook et al., 1979). On the contrary, if the evaluation of an intervention shows no effect and the fidelity is not monitored, it is difficult to establish if the outcomes were due to ineffective intervention content or lack of fidelity to intervention delivery (Moncher and Prinz, 1991). In the former, assessing fidelity to intervention features can prevent ineffective interventions from being scaled-up, saving governments, providers and patients from high cost, whereas, in the latter case, it can prevent potentially promising interventions from being discarded prematurely (Bellg et al., 2004).

Furthermore, complex interventions research can benefit from specifying intervention content and explaining individual differences in outcomes between various practices. An extension of the CONSORT statement for non-pharmacological treatments recommends reporting of the precise details of the intervention as it was implemented, the method of standardisation of the intervention between providers, details of assessing adherence of the providers with the intervention protocol and description of the different components of the intervention ,including the procedures for tailoring the interventions to individual participants (Boutron et al., 2008). Further work on descriptions of complex interventions (Hoffmann et al., 2014) recommend including detailed accounts of what works, what does not work, what works less and how it works etc. in reporting of effectiveness studies.

Differential effectiveness of various ingredients

The fidelity index provides a theoretical framework which can be utilised in research designed to understand the mechanisms of change for smoking cessation. An important use of the fidelity index in effectiveness studies is to identify the active ingredients of an intervention that actually predict quit rates (Bond et al., 2000b).

In the fidelity index, intervention ingredients that are theoretically mapped are represented by items of the index making its compositional features (for details see Chapter 6). How these intervention ingredients are delivered in reality represents fidelity of the providers and is captured by the response scale for each item. Demonstrating empirically that an item of the index is an 'active ingredient' (or effective in changing a particular behaviour) requires obtaining

individual item associations with a criterion measure, that is, quit rate in case of smoking cessation, while controlling for confounders and moderators (Bond et al., 2000b).

Therefore, discriminating effective parts of the intervention from those that are not likely to work in a given context is possible using the fidelity index. The current work on this fidelity index paves the way for research that taps into these mechanisms using appropriate research designs; further discussed below in section 9.5.

9.4.5 Theory testing and refinement

Fidelity measurement can help determine the extent to which the effects seen are due to specific behavioural functions or due to factors non-specific to the intervention or deviations from the intervention protocol (Nigg, 2002). Unless the intervention fidelity is measured, it is difficult to establish the extent to which intervention theory is the prime mechanism of the observed behaviour change (Bellg et al., 2004).

The fidelity index could enable researchers to identify the active ingredients of their interventions; this could give them an opportunity to refine the intervention and describe its content, in terms of the current reporting standards (Boutron et al., 2008, Hoffmann et al., 2014). In addition, it would give researchers an insight into the Quality of interaction of the BI and help them identify those skills and competences that were more or less effective. Moreover, the analysis of sub-indices scores (Adherence to content and Quality of interaction) might determine the effect of the intervention (content) apart from the effects related to the Quality aspects (Bellg et al., 2004).

9.5 RECOMMENDATIONS FOR FUTURE RESEARCH

Future research should further validate the fidelity index, develop shorter versions, develop item bank for intervention fidelity, test the hypotheses generated from this work, identify the active ingredient of BIs and investigate the effect of smoking cessation on TB outcomes.

9.5.1 Further validation of the index

Prospective longitudinal studies can be used to test the predictive validity of the fidelity index, that is, to establish whether the fidelity measured by implementing the index predicts quit rates. The analysis of the psychometric properties of the fidelity index was insufficient for items which were not implemented by the majority of the providers in the Fidelity study and thus could not be investigated thoroughly. Further psychometric testing of the index, including coherence of items using PCA, should evaluate these items (16 'setting quit date', 37 'offering BI leaflet', 39 'explaining expectations regarding the intervention programme' and 40 'providing reassurance') when data are available from other contexts and settings. These evaluations will determine

whether these items are redundant or if they are not measuring fidelity to behaviour change and should be dropped from the index.

Furthermore, the index was developed in one setting and around a particular smoking cessation BI. To improve the generalisability of the fidelity index, it should be tested in other settings and populations and validated for HICs as well as LMICs.

9.5.2 Developing a shorter version

Developing a simpler and shorter version of the fidelity index for providers (self-monitoring), managers (supervision and training), routine auditing of services and research is an important next step. The fidelity index is complex and has multiple items that cannot be used readily outside research purposes. It needs to be simplified for application in routine practice. The psychometric testing and content validations of the shorter versions of the fidelity index should be conducted in future studies.

9.5.3 Developing item bank for intervention fidelity

Developing an item bank (Hahn et al., 2010) for intervention fidelity by adding new items and dimensions to the fidelity index can help intervention evaluators and researchers use readily available items for measuring fidelity, tailored to their respective interventions. Items could be selected from the bank to form customized short indices, or can be administered in a sequence and length determined by the researcher (Cella et al., 2007). Although far from perfect, such item banks can form a common definition and understanding of behavioural mechanisms concerning smoking cessation and other health behaviour change. From a practical perspective, re-writing and re-testing an item, adding more items, re-testing a bank after some modifications, or splitting up a bank into units that are more unidimensional require time and resources. The purpose of such item banks is to have a common metric and range for measuring a construct and a shared meaning and understanding across users.

9.5.4 Identifying 'active ingredients' of a Behavioural Intervention

As described above, identification of the active ingredients of an intervention relies on individual item associations with the outcome, which is not possible in commonly used experimental and intervention research designs. Establishing such associations would require larger sample sizes and comparative designs. Prospective longitudinal studies and effectiveness studies (Piper et al., 2016), using multi-phase optimisation strategy (Baker et al., 2016), to systematically test the effect of the presence or absence of individual items (or BCTs mapped to the items) on smoking behaviour, and the Quality aspects that mediate the effect, can move the field forward in providing mechanistic explanations (Hardcastle et al., 2016).

In addition, a mere 'absent' vs 'present' distinction is inadequate in capturing the natural variance in the effect of the active ingredients and processes by which the intervention leads to behaviour change (Hardcastle et al., 2016). The ordinal response scale of each item in the designed fidelity index would allow the quantification and ordering of an active ingredient in terms of how well it was implemented (i.e. fully-, partially- or not-implemented). This can be tested in large implementation-effectiveness hybrid studies, possibly using factorial designs.

9.5.5 Hypotheses testing

Future experimental or intervention research using longitudinal designs should test the hypotheses generated in this study. These (as described in Chapter 8) are the following:

- What is the optimal number of ingredients (hence the BCTs) to target in a behavioural intervention for smoking cessation to achieve successful patient quit?
- Whether quality of interaction can be improved to an extent that it is effective in achieving successful patient quit, independent of the number of ingredients delivered?
- Whether adherence to behavioural intervention content is effective in achieving higher quit rates, independent of the duration of the patient-provider interaction?

9.5.6 Effect of smoking cessation on TB outcomes

Future studies should investigate the effectiveness of BIs for smoking cessation on TB outcomes. One issue that is vital to the design of successful smoking cessation interventions in routine TB practice is to determine the optimal number of behavioural change elements to target. To close this gap in integration of BIs for smoking cessation within existing TB programmes in LMICs, future studies could potentially build on my work and utilise the fidelity index.

An effectiveness-implementation hybrid study, 'Tobacco cessation within TB programmes: A 'real' world solution for countries with dual burden of disease' (ISRCTN43811467) is underway, which I am privileged to be coordinating. This is a four year (2015 to 2019) multi-country, multi-site RCT, funded by the European Union's Horizon 2020 research and innovation programme, which will not only assess smoking cessation effects on TB treatment outcomes but will also investigate the fidelity to delivery of the behavioural intervention. This study could potentially investigate the hypotheses generated from my work and further the research on the subject.

9.6 REFLECTION ON THE CONTEXTUAL USE OF THE FIDELITY INDEX

The fidelity index was designed based on a single BI for smoking cessation, which was mapped to 23 BCTs using the taxonomies of behaviour change. Some of these BCTs were used more than once (see Appendix C.3). The index is currently structured (from left to the right columns) as behavioural determinant, BCT, item, and the response scale. Modifications in elements on the left means less effect on the reliability and more effect on reliability as you move to elements on the

right. Contextually this means that if we group items together by the BCTs or the behavioural determinants and assess the fidelity the index might still hold good psychometric properties established in this study. However, if it comes to modifying the items for which the response scales have been specifically devised than the overall reliability of the index could potentially decline. Having said that, the two sub-indices can be used separately and the findings of the PCA analysis for the Adherence sub-index (Table 7.8 and

Table 7.9) could be used wisely to separate out the five principal components. As long as the five principal components (described under 'Item loadings for Adherence (multiple components extraction)' on page 146) are used the items within these can be dropped or modified without affecting the variation explained by each principal component too much. However, reliability needs to be re-evaluated if new items are added to the fidelity index.

Nevertheless, the fidelity index developed in this study provides a reliable strategy to be used as a guideline to develop and refine quantitative measures for BI fidelity in future.

9.7 CONCLUSION

A new index for measuring fidelity for BIs of smoking cessation was developed and found to be reliable. Fidelity, as measured using the index, was associated with quit rates; where Quality of interaction was found to positively and Adherence to BI content negatively affect quit rates.

The fidelity index could be a useful tool for exploring and possibly explaining variation in outcomes of smoking cessation. The application of scales such as the fidelity index can contribute to process evaluation of BI programmes in research and their quality improvement in practice, aiming to optimise care delivery and maximise effective outcomes. The development and refinement of such fidelity measurement methods should be the focus of future research in health behaviour change.

The granularity of measuring fidelity to delivery of a BI containing multiple ingredients and the quality aspects of the interaction with the patient demonstrated in this study provides a foundation for generating evidence that can inform targeted future training programmes, continuing professional development and efforts to integrate within existing programmes. Researchers in the field of smoking cessation and other health behaviour change are encouraged to use and build on the fidelity index work.

APPENDICES

Appendix A. SMOKING IN TB PATIENTS- EVIDENCE SYNTHESIS

This section provides a review of the literature on association between smoking and TB outcomes. It appraises the position of this evidence using a reliable grading tool that can assess the strength of association and confidence in the 'estimated effect,' as opposed to the 'true effect' across studies for a given outcome. I describe the methods I used to carry out this review of the effects of smoking on risk of acquiring TB infection, developing TB disease and on TB recurrence and mortality, and present my findings. I also discuss the limitations of the review and next steps.

It is important to understand the epidemiology of TB in light of emerging risk factors and life styles beginning to dominate LMICs, before we go to cessation. For example, in sub-Saharan Africa where smoking rates have been very low is about to hit a smoking epidemic with profound effects on TB and other diseases. In order to utilise the teachable moments in a TB clinic to help TB patients make a quit attempt and sustain it in the long-term, a better understanding of how smoking cessation interventions work is required. An integral question then arises about the epidemiological link between TB and tobacco smoking; necessary to disentangle the TB disease specifics that are influenced by smoking.

A growing body of evidence demonstrates that tobacco smoking increases the risk of acquiring TB infection and developing TB disease. In TB patients, continued smoking leads to poorer treatment outcomes and mortality. However, this evidence is not sufficient to enable policy makers and practitioners to make informed decisions on integrating effective smoking cessation interventions in TB care settings. Firstly, the evidence on the association between smoking and each outcome of TB (e.g. latent TB, active TB, TB mortality, TB relapse etc.) has not been synthesised across different reviews and latest studies. Secondly, the quality of evidence from separate studies has not been rated overall or accounted for in interpreting results.

To my knowledge, the latest systematic reviews on smoking and TB association were conducted in 2007. The subject is in dire need of synthesising evidence from these reviews by including additional epidemiological studies that have been conducted since 2007, as well as providing clear cut conclusions based on the quality of evidence. This is not an attempt to conduct another systematic review and meta-analysis to generate new estimates for the association between smoking and TB outcomes. In fact it is a synthesis of information from existing reviews and additional observational studies to date, to provide clearer inferences of the derived estimates (of the effect of smoking on TB outcomes), based on the quality assessment of the evidence.

OBJECTIVE

My overall research aim is, to investigate whether the evidence on association between smoking and outcomes of TB (latent TB, active TB, TB mortality, and TB recurrence) is of sufficient quality.

The specific objectives are to:

i. Synthesize evidence across systematic reviews and epidemiological studies

ii. Critically appraise the quality of this evidence

METHODS

Criteria for considering studies for this review

Type of studies

This evidence synthesis includes reviews as well as primary epidemiological studies on the association of smoking with TB.

Type of participants

Tobacco smoking primarily affects the lungs and therefore the focus of this review will remain pulmonary TB. Studies included in this review were conducted in adults with pulmonary TB.

Types of outcome measures

The following outcomes are of interest in the association between active smoking (exposure) and TB:

Latent TB infection

Individuals with latent TB infection do not feel unwell and do not have any symptoms. They are infected with *Mycobacterium tuberculosis*, but do not have TB disease (Centers for Disease control, 2014). The only sign of TB infection is a positive reaction to the tuberculin skin test or TB blood test. Such individuals are not infectious and cannot spread TB infection to others. Overall, without treatment, about 5 to 10% of those infected will develop TB disease at some time in their lives (Centers for Disease control, 2014). About half of those who develop TB will do so within the first two years of infection.

Active TB disease

In some people, TB bacteria overcome the defences of immune system and begin to multiply, resulting in the progression from latent TB infection to TB disease (Centers for Disease control, 2014). Some people develop TB disease soon after infection, while others develop TB disease later,

when their immune system becomes weak. Individuals with TB disease are considered infectious and may spread infection to others. TB disease is a serious condition and can lead to death if not treated.

TB mortality

This is defined as the cause of death designated as being due to TB (Slama et al., 2007).

TB recurrence

Recurrent TB, often called TB relapse, is an indicator of TB treatment outcome. Recurrent TB cases are defined as those previously treated for TB, that were declared cured or treatment completed at the end of their most recent course of treatment, and are again diagnosed with an episode of TB (either a true relapse or a new episode of TB caused by reinfection, also known as 'recurrence') (World Health Organisation, 2014a).

Type of exposure

Smoking

For the scope of this particular evidence review, the following case definition for smoking is considered: active tobacco smoking (including cigarettes, water-pipe, hand-rolled cigarettes, cigars or pipes), exposure (including current, past or ever smoking) versus non-exposure; either self-reported smoking status or bio-chemically verified (via CO or cotinine measurement) or both.

Search methods for identification of studies

Relevant studies were identified by searching Medline via Ovid. Database searching was conducted in two subsequent steps.

Step 1- for the synthesis of evidence from reviews- involved systematically searching and identifying published reviews of smoking and TB association in literature.

Step 2- for updating the evidence from these reviews- involved systematically searching and identifying additional primary epidemiological studies on smoking and TB association, published since the last date of the database search mentioned in the reviews identified in step 1.

Bibliographies of selected articles were also searched for relevant studies. The main keywords and phrases used included; tobacco use, smoking, tuberculosis, mortality, infection, disease, treatment outcomes, relapse and recurrence, which were sourced from previous reviews (Bates et al., 2007, Lin et al., 2007, Slama et al., 2007). Searches (for review articles) were limited to Abstracts, in English, and from January 1985- July 2015. For the primary studies, the searches were restricted to

Abstracts, in English, and from 2005- July 2015. The detailed search strategies are provided in Tables A.1 and A.2 below.

Given the limitations of resources and time for this work, article searching was restricted to English language, full text availability and Medline database only.

A.1: Search strategy output for reviews on Tobacco and TB (1985-2015)

Step	Search syntax	Results
1	"Tobacco Use"/ or Tobacco/	24975
2	Smoking/	126901
3	Tuberculosis, Pulmonary/ or Latent Tuberculosis/ or Tuberculosis/ or Mycobacterium tuberculosis/	143038
4	1 or 2	147686
5	3 and 4	627
6	limit 5 to (abstracts and English language and "review articles" and yr="1985 -Current")	35

A.2: Search strategy output for additional studies on Tobacco and TB (2005- 2015)

Step	Search syntax	Results
1	"Tobacco Use"/ or Tobacco/	24975
2	Smoking/	126901
3	Tuberculosis, Pulmonary/ or Latent Tuberculosis/ or Tuberculosis/ or Mycobacterium tuberculosis/	143038
4	1 or 2	147686
5	3 and 4	627
6	limit 5 to (abstracts and English language and full text and yr="2005 -Current")	56

Data collection and analysis

Selection of studies

The criteria for identifying the studies for evidence synthesis are provided separately for the reviews and primary epidemiological studies on the smoking and TB association, followed by the information that was extracted.

Inclusion criteria for Reviews on smoking and TB:

- Systematic reviews published between 1985 and Jul 2015

- Systematic reviews focusing primarily on smoking forms of tobacco
- Systematic reviews reporting on at least one of the (above) defined outcomes (of TB)

Inclusion criteria for Primary studies on smoking and TB:

- Studies published between 2005 and Jul 2015
- Epidemiological studies; cohort, cross-sectional or case-control studies
- Studies focusing primarily on smoking forms of tobacco
- Studies reporting a risk estimate (e.g. relative risk, odds ratio or hazard ratio) and confidence intervals (CI) for at least one of the (above) defined outcomes (of TB)

Exclusion criteria:

Studies/reviews primarily focusing on smokeless tobacco or second-hand smoke or with no clearly defined TB related outcomes or not reporting a risk estimate or reporting on subjects with comorbidities (e.g. diabetes or HIV or lung cancer) were excluded.

Data extraction and management

Information from all studies (including the systematic reviews) was extracted in a standardised way. A uniform template was used to extract data on: category of TB outcome, study design, location, population/setting, measures of outcome and exposure, number analysed, effect estimates, adjusted confounders and dose-response gradient. For the systematic reviews, data on heterogeneity effect and publication bias were also extracted.

Synthesis of evidence and quality appraisal

The Grading of Recommendations, Assessment, Development, and Evaluation (GRADE) approach was taken to synthesise effects of smoking on TB outcomes from identified studies and draw sound conclusions based on the quality of this evidence (GRADE Working Group, 2004).

GRADE enables a systematic and transparent process to assess the quality of evidence and strength of association for a particular outcome across studies and reviews. Initially, GRADE was restricted for use in clinical trials with outcomes for key treatment interventions. In recent years, however, its use has been extended to non-randomised studies, by including considerations arising from the Bradford Hill's criteria for causation (Schünemann et al., 2011). In the current evidence synthesis, a GRADE table is presented for each outcome of interest that identifies the basis of judgements made about the quality and the score assigned. An initial four points are awarded to evidence that is largely based on RCTs, and 2 points to evidence based on observational studies. The points are then upgraded or downgraded based on quality assessment criteria (the GRADE checklist, discussed below). An overall GRADE score (from 4 to 0) based on the assessment of the quality of evidence for a particular outcome is achieved by this process.

The GRADE checklist for quality assessment was applied to each TB outcome using the Bradford Hill's criteria (where applicable) within each of its five items (Meader et al., 2014). The five items (detailed in Appendix A.9) of the checklist are:

i. Study limitations (study design, temporality, and risk of bias);
ii. Inconsistency (overlapping confidence intervals, direction of effect, heterogeneity);
iii. Indirectness (objective vs. subjective outcome measure, direct comparisons, coherence, biological plausibility and specificity);
iv. Imprecision (strength of association, median sample size, number of included studies);
v. Publication bias and other considerations (dose-response gradient, adjustment for key confounders).

The GRADE approach for assessing quality of evidence for the association between smoking and TB infection, disease, mortality and recurrence involved the following procedures:

1. *A-priori* ranking of "high (4)" to randomized controlled trials and "low (2)" to observational studies

2. "Downgrading" initial ranking based on limitations of study design, inconsistency, indirectness, imprecision or publication bias

3. "Upgrading" initial ranking based on large effect size, dose-response gradient and adjustment for main confounders

4. Final grade assigned for the quality of evidence as "high", "moderate", "low" or "very low" for each outcome of interest. Interpretation of the ranks of the quality of evidence is provided in Table A.3 (Schünemann et al., 2011).

5. Conclusion based on the quality of evidence and consideration of other factors that impact on the strength of association

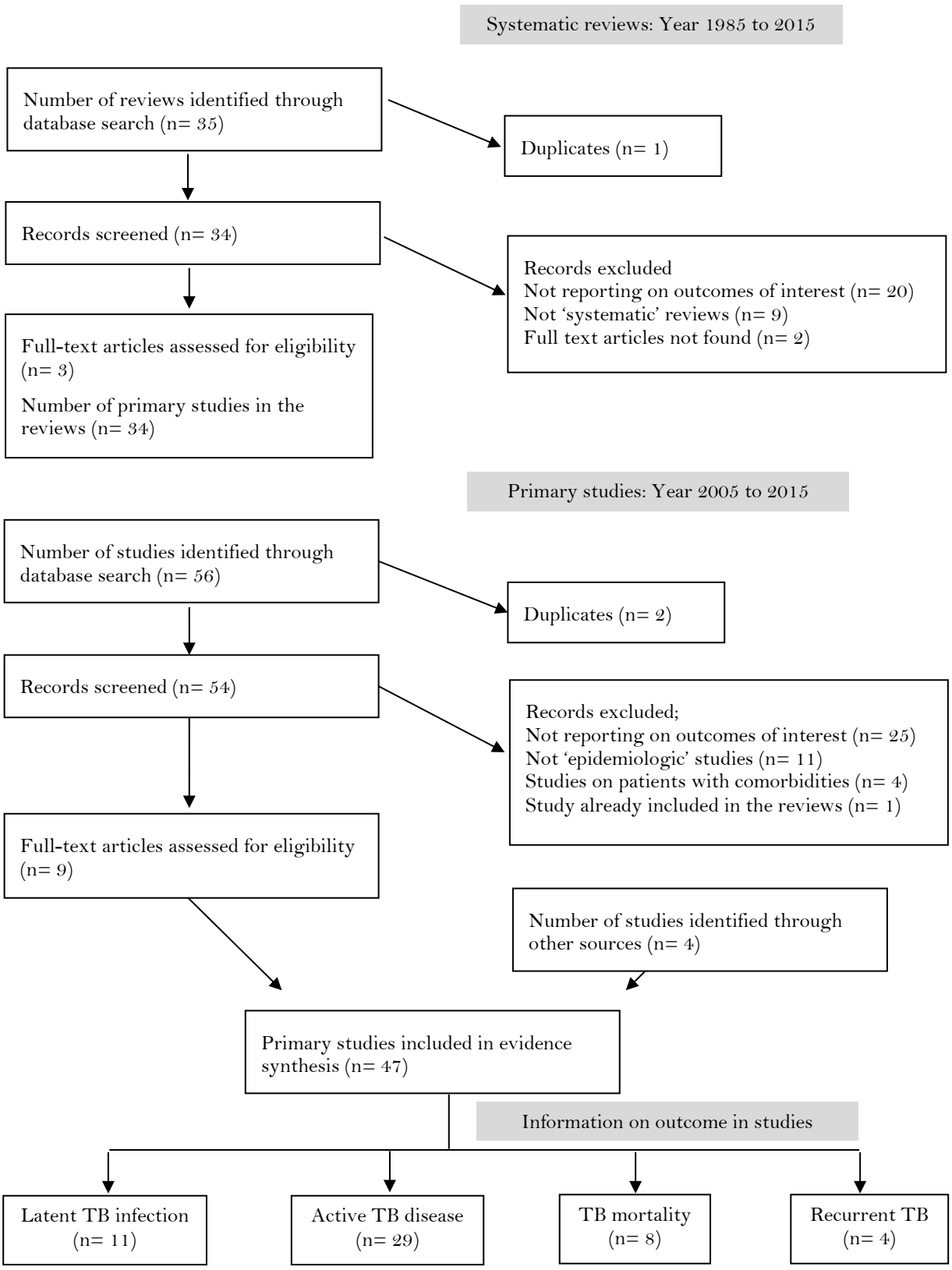
A.3: Final GRADE ranking

High ⊕⊕⊕⊕	There is high confidence that the true effect lies close to that of the estimate of the effect
Moderate ⊕⊕⊕	There is moderate confidence in the effect estimate: the true effect is likely to be close to the estimate of the effect, but there is a possibility that it is substantially different
Low ⊕⊕	There is limited confidence in the effect estimate: the true effect may be substantially different from the estimate of the effect
Very low ⊕	There is little confidence in the effect estimate: the true effect is likely to be substantially different from the estimate of the effect

RESULTS

Description of studies (reviews and additional primary studies)

Database searching, for Step 1, led to the identification of three relevant systematic reviews (Bates et al., 2007, Lin et al., 2007, Slama et al., 2007) (Figure A.4; Appendix A.10). Thirty-four studies were included in these systematic reviews collectively on TB outcomes and active (current or ever) smoking (Appendix A.10 & Appendix A.11). Step 2, database searching for additional primary studies, published after 2005, led to identification of another nine studies (Lindsay et al., 2014, Horne et al., 2012, Rao et al., 2012, Brunet et al., 2011, Wen et al., 2010, Jee et al., 2009, Wang and Shen, 2009, Gajalakshmi and Peto, 2009, d'Arc Lyra Batista et al., 2008) (Appendix A.10 & Appendix A.11). Further bibliography searching of the included studies identified four relevant studies (Thomas et al., 2005, Al-Darraji et al., 2015, Adib et al., 1999, Pednekar and Gupta, 2007) for inclusion in the evidence synthesis (Appendix A.10 & Appendix A.11).



A.4: Flow diagram of identified reviews and additional primary studies

Evidence for Latent TB infection

There were 11 studies on the association between smoking and acquisition of TB infection; one of which was case-control (Anderson et al., 1997) and remaining were cross-sectional in design (Table A.5). Smokers were more likely to have latent TB infection compared to the non-smokers; odds ratio (OR) of latent TB infection ranged between 1.73 (95%CI: 1.46-2.04) and 1.83 (95%CI: 1.49-2.23) for the systematic reviews (Bates et al., 2007, Lin et al., 2007, Slama et al., 2007).

There was no evidence of heterogeneity between these studies by looking at the I-squared statistics in the three systematic reviews and the overlapping of confidence intervals (CI) around the effect estimates for the additional primary studies. The only outlier was Brunet et al. (2011), a cross-sectional study done in South Africa that reported OR of 0.64 (95%CI: 0.14-2.79); the study used objective measures, of TB infection (Interferon Gamma Release Assays- IGRA) and smoking (serum cotinine) as opposed to Tuberculin Skin Test (TST) positivity for TB infection and self-reported smoking in other studies. However, the number analysed for TB infection was very small i.e. 108 individuals. The largest effect (OR: 2.8; 95%CI: 1.6-5.2) was seen in Hussain et al. (2003), which was a cross-sectional study conducted in prisoners, in Pakistan, using TST for TB infection and self-reported smoking status.

The magnitude of the median sample size (> 300 participants) and the number of included studies (> 10 studies) was considered high, according to the GRADE checklist. Publication bias was not observed for the studies included on TB infection in the three systematic reviews.

A dose-response gradient was observed for acquisition of TB infection with both the quantity and the duration of smoking (Hussain et al., 2003, Den Boon et al., 2005, Anderson et al., 1997, Horne et al., 2012).

The true effect of tobacco smoking on acquisition of TB infection may be substantially different from the estimate of the effect, which is approximately double the odds of acquiring TB infection among smokers than non-smokers, according to the systematic reviews. This evidence for the association between smoking and TB infection remains rather limited, especially considering temporality could not be established, due to the lack of longitudinal observational studies.

A.5: Synthesis of evidence for the association between smoking and TB infection

			Rating	Adjustment to score
Quality assessment	No. of studies/ starting score		11 observational	2
	Factors decreasing confidence	Limitation in study design (Temporality)	Serious (10 cross-sectional, 1 case-control, studies)	-1
		Inconsistency (Heterogeneity in point estimates, CI's, direction of effect)	None serious[1]	0
		Indirectness (Biological plausibility, Generalizability, outcome not a surrogate measure, direct comparison)	Serious[2]	-1
		Imprecision (Strength of association, magnitude of median sample size, magnitude of included studies, outcome a common event)	None serious	0
		Publication bias	None serious	0
	Factors increasing confidence	Large effect	Weak evidence[3]	0
		Dose-response	Moderate evidence	+1
		Mitigated bias and confounding	Moderate evidence[4]	+1
	Final GRADE score of quality of evidence			2
Summary of findings	Quality of evidence			Low ⊕⊕
	Conclusion			Due to the lack of temporality data (weak study designs) the credibility of observed association remains limited.

[1] Heterogeneity between primary studies and within systematic reviews was considered

[2] Both TB infection and smoking in majority of the studies are assessed using surrogate measures or self-reporting

[3] Weak evidence- not all effect sizes >2 or < 0.5 and significant

[4] Adjustment for main confounders (i.e. age, gender and SES) in majority studies

Abbreviations: SES, Socio-economic status; OR, Odds ratio

Evidence for active TB disease

There were 29 studies on the association between smoking and developing active TB (Table A.6); two of which were prospective cohort (Leung et al., 2004, Jee et al., 2009), five cross-sectional and remaining were case-control designs (for details refer to Appendix A.10). Smokers were more likely to develop TB disease compared to the non-smokers; the effect estimates of developing active TB

disease ranged between 2.01 (95%CI: 1.63-2.48) and 2.33 (95%CI: 1.97-2.75) for the systematic reviews (Bates et al., 2007, Lin et al., 2007, Slama et al., 2007).

There was moderate evidence of heterogeneity (I^2: 40-60%) between these studies by looking at the heterogeneity statistics of pooled studies in the three systematic reviews and the overlapping of CIs around the effect estimates for the additional primary studies. The effect estimates of three studies (Shetty et al., 2006, Brunet et al., 2011, Brown and Campbell, 1961) showed negative association, but were statistically non-significant. One of these studies was a cross-sectional design and the other two were case-controls, with sample sizes ranging between 200 and 400. A large effect of smoking on TB disease was seen in two case-control studies; Tekkel et al. (2002): OR 4.62 (95%CI: 2.44-8.73), and Gupta et al. (2001): OR 4.42 (95%CI: 2.55-7.66). Both cohort studies showed moderate to high effect size estimates; OR of 2.87 (95%CI: 2.00-4.11) in Leung et al. (2004), and Hazards Ratio (HR) of 1.4 (95%CI: 1.3-1.4) in Jee et al. (2009).

The magnitude of the median sample size (> 300 participants) and the number of included studies (> 10 studies) was considered high, according to GRADE checklist. Publication bias was not observed for the studies included on TB infection in the three systematic reviews.

A dose-response gradient was observed for the increase in risk of developing TB disease both with increasing quantity (Leung et al., 2004, Gajalakshmi et al., 2003, Gupta et al., 2001, Dong et al., 2001, Wang et al., 2005, Alcaide et al., 1996, Adelstein and Rimington, 1967, Lowe, 1956, Wang and Shen, 2009, Jee et al., 2009) and duration (Yu et al., 1988, Buskin et al., 1994, Ariyothai et al., 2004, Lienhardt et al., 2005, Kolappan and Gopi, 2002) of smoking.

From grading evidence on the association between tobacco smoking and active TB disease, there is high confidence that the true effect lies close to that of the estimate of the effect, which is approximately double the odds of developing TB disease among smokers over non-smokers, according to the systematic reviews. Therefore, strong evidence supports that tobacco smoking increases the risk of developing TB disease.

A.6: Synthesis of evidence on the association between smoking and TB disease

			Rating	Adjustment to score
Quality assessment	No. of studies/ starting score		29 observational	2
	Factors decreasing confidence	Limitation in study design (Temporality)	None serious (5 cross-sectional, 22 case-control, 2 cohort studies)	0
		Inconsistency (Heterogeneity in point estimates, CI's, direction of effect)	Serious[1]	-1
		Indirectness (Biological plausibility, generalizability, outcome not a surrogate measure, direct comparison)	None serious[2]	0
		Imprecision (Strength of association, magnitude of median sample size, magnitude of included studies, outcome a common event)	None serious	0
		Publication bias	None serious	0
	Factors increasing confidence	Large effect	Moderate evidence[3]	+1
		Dose-response	Moderate evidence	+1
		Mitigated bias and confounding	Moderate evidence[4]	+1
	Final GRADE score of quality of evidence			4
Summary of findings	Quality of evidence			High ⊕⊕⊕⊕
	Conclusion			Strong evidence that tobacco smoking increases the risk of developing TB disease

[1] Moderate statistical heterogeneity (i.e. 40- 60%) between studies estimates

[2] Diagnosis of TB disease (Sputum smear microscopy or culture) is a direct measure while smoking status remains self-reported in most studies

[3] Moderate evidence- effect sizes >2 or < 0.5 for majority studies/meta-analysis included and significant; OR ranged between 2.0 and 2.3, statistically significant for the systematic reviews

[4] Adjustment for main confounders (i.e. age, gender, alcohol use, region, BCG scar and SES) in majority studies

Abbreviations: SES, Socio-economic status; OR, Odds ratio; BCG, Bacillus Calmette–Guérin

Evidence for TB mortality

There were 8 studies on the association between smoking and TB mortality; four prospective cohort and four case-control studies (Table A.7). Smokers were more likely to die from TB disease compared to the non-smokers; the OR of dying from TB ranged between 2.00 (95%CI: 1.14-3.49) and 2.24 (95%CI: 1.34-3.73) for the systematic reviews (Bates et al., 2007, Lin et al., 2007, Slama et al., 2007).

There was high heterogeneity (I^2: >60%) between the mortality studies pooled in the systematic reviews. However, all studies had statistically significant (positive) associations in the direction of effect between smoking and death due to TB. The highest effect estimates were observed in Gajalakshnmi et al. (2003): OR 4.5 (95%CI: 4.0-5.0), Wen et al. (2010): HR 4.19 (95%CI: 1.8-9.7), and Gupta et al. (2005): OR 3.31 (95%CI: 1.34-8.16).

The magnitude of the median sample size was high (> 300 participants) and the number of included studies was moderate (5-10 studies), according to the GRADE checklist. Publication bias was not assessed for the studies included on TB mortality in the three systematic reviews.

A dose-response gradient was observed for TB mortality with both the quantity (Jee et al., 2009, Lam et al., 2001, Liu et al., 1998, Pednekar and Gupta, 2007) and the duration of smoking (Lin et al., 2007). Cause of death was established in two studies by death certificates; one relied on verbal autopsy and the rest on records from hospitals and thenational statistics database. In a cohort of British doctors, a Relative Risk (RR) of 2.8 with a dose-response gradient for quantity smoked (RR for those who smoked more than 25 cig/day was 5 (Slama et al., 2007)) was observed for TB mortality, compared to lifetime non-smokers (Doll, 1999). The study was not included in the evidence synthesis as it did not provide the CIs around the risk estimates and was not a direct report of the cohort study.

From grading evidence on the association between tobacco smoking and TB mortality, there is high confidence that the true effect lies close to that of the estimate of the effect, which is approximately double the odds of dying from TB among smokers over non-smokers. Therefore, strong evidence supports that tobacco smoking increases the risk of dying from TB.

A.7: *Synthesis of evidence on the association between smoking and TB mortality*

			Rating	Adjustment to score
Quality assessment	No. of studies/ starting score		8 observational	2
	Factors decreasing confidence	Limitation in study design (Temporality)	None serious (4 cohort, 4 case-control studies)	0
		Inconsistency (Heterogeneity in point estimates, CI's, direction of effect)	Serious[1]	-1
		Indirectness (Biological plausibility, generalizability, outcome not a surrogate measure, direct comparison)	None serious[2]	0
		Imprecision (Strength of association, magnitude of median sample size, magnitude of included studies, outcome a common event)	None serious	0
		Publication bias	None serious	0
	Factors increasing confidence	Large effect	Moderate evidence	+1
		Dose-response	Moderate evidence	+1
		Mitigated bias and confounding	Moderate evidence[3]	+1
	Final GRADE score of quality of evidence			4
Summary of findings	Quality of evidence			High ⊕⊕⊕⊕
	Conclusion			Strong evidence that tobacco smoking increases the risk of TB mortality

[1] High statistical heterogeneity (i.e. > 60%) between studies estimates

[2] Moderate evidence- effect sizes >2 or < 0.5 for majority studies/meta-analysis included and significant; OR ranged between 2.0 and 2.2, statistically significant for the 3 systematic reviews

[3] Adjustment for main confounders (i.e. age, gender and SES) in majority studies

Abbreviations: SES, Socio-economic status; OR, Odds ratio

Evidence for TB recurrence

None of the systematic reviews reported pooled estimates for the association between smoking and TB recurrence. There were 4 primary studies (all cohort in design) included in the evidence synthesis for this association (Table A.8).

There was no evidence of heterogeneity between studies; effect estimates for all studies were statistically significant and in the direction of the effect between smoking and the risk of recurrence. The OR reported in three of the studies were, 3.1: 95% CI, 1.6-6.0 (Thomas et al., 2005), 2.53: 95%CI, 1.23-5.21 (d'Arc Lyra Batista et al., 2008) and 2.48: 95%CI, 1.04-5.89 (Leung et al., 2004). One study (Jee et al., 2009) reported a Hazards Ratio of 1.3 (95%CI: 1.2-1.4), which can be considered high as HR is a risk ratio and could only be approximated to OR when the event is rare. However, TB recurrence is not a rare event; relapse rates of 18.1% among smokers and 7.3% among non-smokers are observed, of which 77% of cases occur in the first six months of follow-up after treatment completion (Thomas et al., 2005).

The magnitude of the median sample size was high (> 300 participants) and the number of included studies was small (< 5 studies), according to the GRADE checklist. Publication bias could not be assessed as the only systematic review (Slama et al., 2007) reporting on TB recurrence studies did not conduct meta-analysis of these studies. A dose-response gradient was observed for the quantity smoked and risk of TB recurrence in Jee et al. (2009). The study by Thomas et al. (2005) used sputum smear microscopy and culture to confirm TB recurrence at 6, 12 and 18 months post-TB treatment and cure of the patient, while the rest of the studies used retreatment after a completed treatment for active TB in the past.

From grading evidence on the association between tobacco smoking and TB recurrence, there is high confidence that the true effect lies close to that of the estimate of the effect, which is approximately double the odds of recurrence of TB among smokers than non-smokers. Therefore, strong evidence supports that tobacco smoking increases the risk of re-infection and/or relapse from TB.

A.8: Synthesis of evidence on the association between smoking and TB recurrence

			Rating	Adjustment to score
Quality assessment	No. of studies/ starting score		4 observational	2
	Factors decreasing confidence	Limitation in study design (Temporality)	None serious (All prospective cohort designs)	0
		Inconsistency (Heterogeneity in point estimates, CI's, direction of effect)	None serious	0
		Indirectness (Biological plausibility, generalizability, outcome not a surrogate measure, direct comparison)	None serious	0
		Imprecision (Strength of association, magnitude of median sample size, magnitude of included studies, outcome a common event)	Serious[1]	-1
		Publication bias	None serious	0
	Factors increasing confidence	Large effect	Moderate evidence[2]	+1
		Dose-response	Moderate evidence	+1
		Mitigated bias and confounding	Moderate evidence[3]	+1
	Final GRADE score of quality of evidence			4
Summary of findings	Quality of evidence			High ⊕⊕⊕⊕
	Conclusion			Strong evidence that tobacco smoking increases the risk of TB recurrence

[1] Less than 5 studies

[2] Moderate evidence- effect sizes >2 or < 0.5 for all studies

[3] Adjustment for main confounders (i.e. age, gender, alcohol and SES) in majority studies

Abbreviations: SES, Socio-economic status

DISCUSSION

Summary of main results

Tobacco smokers are twice as likely to develop TB disease, suffer recurrent TB and die from TB, compared to non-smokers. However, the evidence on the association between tobacco smoking and acquiring TB infection was found to be weak. According to the GRADE scores, the evidence for active TB disease, TB mortality and TB recurrence was ranked as high quality, while the evidence for latent TB infection was ranked as low quality. This study supports the findings from previous

reviews on the association of smoking with TB disease. Unlike previous reviews that were unclear about the association of smoking and TB mortality (Bates et al., 2007), my study provides concrete evidence on this association and fills the gap in knowledge. An important finding of this review is that the association of smoking with TB infection is weak, which has been judged on the meta-analysis estimates of Odds Ratios ranging between 1.5 to 2 in previous reviews (Bates et al., 2007, Lin et al., 2007). This study also provided the evidence for the association between smoking and TB recurrence for the first time, which was found to be strong and of good quality. TB recurrence has not been evaluated in previous reviews of smoking and TB outcomes.

Overall completeness and applicability of evidence

The high grade of the quality of association between smoking and TB disease, mortality and recurrence across all available evidence allows a higher confidence in interpreting the estimates of effect. For the given outcomes, high quality of the effect estimate means that the true effect of smoking on these three outcomes lies close to the observed effect, which is found to be moderate to strong (an OR of 2) (Craun and Calderon). The longitudinal studies (especially prospective cohort) on the effect of smoking on these three TB outcomes lend credibility to the already established temporal and causal associations. On the other hand, the low quality of evidence on acquisition of TB infection gives limited confidence in the observed effect estimates, as they might differ substantially from the true effect. This is mainly because there was a single longitudinal study (case-control) and no prospective cohort studies that could establish the temporal association between smoking and TB infection. However, this finding should not be taken as evidence of no association, because four of the studies in the review found a dose-response gradient (i.e. a higher risk of acquiring TB infection with larger exposure), which makes a causal association more plausible.

Smoking adversely impacts a variety of other TB-related outcomes that were not covered in the current review. Smoking affects individuals with latent TB infection by triggering their progression to develop active TB disease (Alcaide et al., 1996) or reactivation of TB. Furthermore, smokers encounter more severe TB disease, with numerous and bigger infiltrates (cavitation) (OR 1.9; 95%CI: 1.6 to 2.3) and greater likelihood of hospitalization (OR 1.8; 95%CI: 1.5 to 2.2) than non-smokers (Altet-Gomez et al., 2005). In addition, the duration of conversion of sputum smear (from positive to negative for mycobacterium TB) is prolonged in TB patients who also smoke tobacco (Onyebujoh et al., 1999). These outcomes are also of significance from a public health point of view and suggest opportunistic cessation treatment by targeting all of the patients attending TB clinics (e.g. TB suspects) and not just those diagnosed with TB.

The current evidence synthesis was restricted to active smoking and its effect on TB outcomes. However there is now evidence that passive smoking (aka Second-Hand Smoking (SHS)) also influences TB disease and worsens its outcomes (Patra et al., 2015). The association between SHS

and the risk of TB infection and disease was systematically reviewed and published in a peer review journal (Dogar et al., 2015) by myself and colleagues.

Potential biases of the review

There are several potential limitations of this evidence synthesis study. Firstly, limited database searching and language restrictions might have led to exclusion of important primary studies. However, to minimise this bias, an extensive bibliography searching of all the identified studies was carried out. Second, there is possibility of misclassification of both exposure (tobacco smoking) and outcome status (TB outcomes of interest). Active smoking assessment relied majorly on self-reports in almost all of the studies, which may not be an accurate account of smoking status in some individuals (West et al., 2007). Furthermore, some studies only measured 'current smoking,' without accounting for the duration of smoking to date, which may be subject to reverse causation. TB patients are often diagnosed months after they first experience the respiratory symptoms, leading many of them to stop smoking even before they are diagnosed. This might be explained by the fact that several studies (McCurdy et al., 1997, Shetty et al., 2006, Crampin et al., 2004) showed a stronger effect of 'former smoking' compared to 'current smoking' on TB. The use of a surrogate measure (the TST) for all but one study of TB infection, might also lead to misclassification of outcome status, depending on the cut-offs of induration used.

Research implications

Prospective cohort studies, with objective measures of TB infection, are required to establish its true causal association with smoking. Moreover, future studies should utilise highly sensitive and specific objective measures (e.g. serum cotinine testing) for smoking assessment. There is strong evidence that tobacco smoking not only leads to active TB disease but also adversely affects TB treatment outcomes like recurrence and death, which can be offset by timely and appropriate action to target TB patients for smoking cessation. Although the strength of association for tobacco smoking and TB disease/mortality/recurrence was found to be moderate to high, the implication for population health in LMICs is critical. Given the high prevalence of smoking in these countries, which has been consistently rising over time, a considerable portion of the global burden of TB may be attributed to smoking (Pai et al., 2007).

CONCLUSION

There is strong evidence that tobacco smoking increases the risk of developing active TB disease, the risk of its recurrence and the risk of death from it. However, its association with the risk of latent TB infection is found to be weak and the quality of evidence remains limited. Prospective longitudinal studies can strengthen the evidence on association between smoking and TB infection, in future.

More importantly, integration of smoking cessation services within TB care and management (targeting TB disease, recurrence and mortality) is an urgent consideration. This can only be achieved if the need to integrate smoking cessation services into TB care and management services is realised and incorporated into current policies and initiatives rolled out by LMICs burdened with TB.

A.9: GRADE checklist for Quality Assessment

Study limitations (Risk of Bias)

1) Was random sequence generation used (i.e. no potential for selection bias)?

- Yes
- No
- Unclear

2) Was allocation concealment used (i.e. no potential for selection bias)?

- Yes
- No
- Unclear

3) Was there blinding of participants and personnel (i.e. no potential for performance bias)?

- Yes
- No
- Unclear

4) Was there blinding of outcome assessment (i.e. no potential for detection bias)?

- Yes
- No
- Unclear

5) Was an objective outcome used?

- Yes
- No

6) Were more than 80%[4] of participants enrolled in trials included in the analysis (i.e. no potential reporting bias)?

- Yes
- No
- Unclear

7) Were data reported consistently for the outcome of interest (i.e., no potential selective reporting)?

- Yes
- No

[4] 80% drop out is given as an example here a different proportion can be used depending on the context of the systematic review area

- Unclear

8) No other biases reported? (i.e. no potential of other bias)

- Yes
- No

9) Did the trials end as scheduled (i.e not stopped early)?

- Yes
- No

Inconsistency[5]

1) Point estimates did not vary widely?

- Yes
- No

2) To what extent did confidence intervals overlap?

- Substantial overlap
  (all confidence intervals overlap at least one of the included studies point estimate)
- Some overlap
  (confidence intervals overlap but not all overlap at least one point estimate)
- No overlap
  (At least one outlier: where the confidence interval of some of the studies do not overlap with those of most included studies)

3) Was the direction of effect consistent?

- Yes
- No

4) What was the magnitude of statistical heterogeneity (as measured by I^2)?

- Low (e.g. I^2 <40%)
- Moderate (e.g. I^2 40-60%)
- High (e.g. I^2 >60%)

5) Was the test for heterogeneity statistically significant ($p<0.1$)?

- Not statistically significant

[5] Reviewers may choose to use estimates from a subgroup analysis which may explain the inconsistency but should be cautious that such a explanation of heterogeneity may be due to the play of chance

- Statistically significant

Indirectness

1) Were the populations in included studies applicable to the decision context?

- Highly applicable
- Applicable
- Poorly applicable

2) Were the interventions in the included studies applicable to the decision context?

- Highly applicable
- Applicable
- Poorly applicable

3) Was the included outcome not a surrogate outcome?

- Yes
- No

4) Was the outcome timeframe sufficient?

- Sufficient
- Insufficient

5) Were the conclusions based on direct comparisons?

- Yes
- No

Imprecision

1) Was the confidence interval for the pooled estimate not consistent with benefit and harm?

- Yes
- No

2) What is the magnitude of the median sample size?

- High (e.g. 300 participants)
- Intermediate (e.g. 100-300 participants)
- Low (e.g. <100 participants)

3) What was the magnitude of the number of included studies?

- Large (e.g. >10 studies)
- Moderate (e.g. 5-10 studies)
- Small (e.g. <5 studies)

4) Was the outcome a common event (e.g. occurs more than 1/100)?

- Yes
- No
- Not applicable (i.e. not a dichotomous outcome)

Further optional question for those engaged in guideline development[6]

5) Was there no evidence of serious harm associated with treatment?

- Yes
- No

Publication Bias (other considerations)

1) Did the authors conduct a comprehensive search?

- Yes
- No

2) Did the authors search for grey literature?

- Yes
- No

3) Authors did not apply restrictions to study selection on the basis of language?

- Yes
- No

4) There was no industry influence on studies included in the review?

- Yes
- No

5) There was no evidence of funnel plot asymmetry?

- Yes

[6] This reflects GRADE guidance that guideline developers may use a less stringent threshold for judging imprecision of an intervention's benefits when there is no evidence of harm compared with when judging the benefits of an intervention where there is strong evidence of harm

- No
- Unclear

6) There was no discrepancy in findings between published and unpublished trials?

- Yes
- No
- Unclear

A.10: Data extraction from reviews and primary studies

Category	Study design	Study	Location	Population/ setting	Measure of disease outcome	Measure of exposure	Sample size	Effect estimate (95% CI)	Heterogeneity (I-sq; p-value)	Adjusted confounders	Dose response gradient	Publication bias
TB infection	Review	Lin et al. 2007	-	-	TST cut-off 5mm = 2 studies; TST cut-off 10mm = 4 studies	Current smoking= 4 studies; Ever smoking= 2 studies (all self-reports)	-	OR: TST cut-off 5mm = 2.08 (1.53-2.83); TST cut-off 10mm = 1.83 (1.49-2.23)	0%	Alcohol, SES	N/A	None detected
TB disease					Clinical TB disease diagnosed on sputum smear microscopy and/or CXR	Current smoking= 13 studies; Ever smoking= 2 studies Former smoking= 1 study; (all self-reports)		OR: 2.01 (1.63-2.48)	63.7%	Alcohol, SES, Type of study, Mode of diagnosis		
TB mortality					Reporting TB mortality- no strict definition used	-		OR: 2.00 (1.14-3.49)	98.7%	none	Yes	not assessed
TB infection	Review	Bates et al. 2007	-	-	TST cut-off 10mm= 5 studies;	Current smoking= 2 studies; Ever smoking= 2 studies Former smoking= 1 study (all self-reports)	-	RR: 1.73 (1.46-2.04)	p-value= 0.71	none	N/A	None detected
TB disease					TB disease diagnosed by sputum smear microscopy or culture or notification	Current smoking= 7 studies; Ever smoking= 4 studies Former smoking= 1 study (all self-reports)		RR: 2.33 (1.97-2.75)	p-value= 0.04			None detected
TB mortality					Death certificate or verbal autopsy	-		RR: 2.15 (1.38-3.35)	p-value= <0.001			not assessed
TB	Review	Slama et	-	-	TST (cut-off not	Current smoking=	4729	OR:	p-value=	none	N/A	Not

infection		al. 2007			specified)	3 studies; Ever smoking= 2 studies (all self-reports)		1.76 (1.47-2.12)	0.525			assessed
TB disease					Clinical TB disease diagnosed on sputum smear microscopy, culture and/or CXR and response to anti-TB drugs	Current smoking= 14 studies; (all self-reports)	159854	OR: 2.28 (1.77-2.95)	p-value= <0.001			
TB mortality					Death certificate or verbal autopsy or medical records verifying death due to TB	-	67168	OR: 2.24 (1.34-3.73)	p-value= <0.001			
TB infection	case-control	Anderson et al. 1997	USA	Prisoners	TST cut-off 10mm	Self-reported current smoking	293 (116 HIV+ve)	OR: 1.78 (0.98-3.21)	-	Age, living conditions, gender, alcohol, HIV, contact with TB patient, BMI	Yes (for both duration and quantity)	-
TB infection	cross-sectional	DenBoon et al. 2005	South Africa	High risk Urban community	TST cut-off 10mm	Self-reported ever smoking	2347	OR: 1.77 (1.33-2.35)	-	Same address clustering, age gender, SES, BMI	Yes (Dose response observed on pack years)	-
TB infection	cross-sectional	Hussain et al. 2003	Pakistan	Prisoners	TST cut-off 10mm	Self-reported current smoking	425	OR: 2.8 (1.6-5.2)	-	Education, duration of imprisonment, crowding of cell, age, SES, BCG	Yes (Quantity)	-
TB infection	cross-sectional	Plant et al. 2002	Vietnam	Vietnamese immigrants	TST cut-off 10mm	Self-reported ever smoking	1395	OR: 1.53 (1.13-2.09)	-	Age, gender, TB contact, living	No	-

										condition, SES		
TB infection	cross-sectional	McCurdy et al. 1997	USA	Migrant farm workers	TST cut-off 10mm	Self-reported current and former smoking	296	OR: 1.87 (0.73-4.8)	-	Birthplace, age, gender	Not assessed	-
TB infection	cross-sectional	Solsona et al. 2001	Spain	Homeless shelter	TST cut-off 5mm	Self-reported current smoking	447	OR: 1.72 (1.02-2.86)	-	Age, gender, alcohol, BCG	Not assessed	-
TB infection	cross-sectional	Adib et al. 1999	Lebanon	Prisoners	TST cut-off 8mm	Self-reported current smoking	3931	OR: 1.2 (1.1-1.3)	-	Residence area, age, gender, occupation, duration of imprisonment	Not assessed	-
TB infection	cross-sectional	Horne et al. 2012	USA	Population based National survey	TST cut-off 10mm	Self-reported current smoking verified by serum cotinine testing	3843	OR: 1.76 (1.06-2.94)	-	Birthplace, age, gender, SES, TB contact, Education, BCG, race/ethnicity	Yes (in ethnic subgroups)	-
TB infection	cross-sectional	Al-Daraji et al. 2015	Malaysia	Prison staff	TST cut-off 10mm	Self-reported current smoking	420	OR: 1.94 (1.17-3.22)	-	Age, alcohol, duration of work, current post	Not assessed	-
TB infection	cross-sectional	Lindsay et al. 2014	USA	Population based National survey	TST cut-off 10mm	Self-reported current smoking verified by serum cotinine testing	938	OR: 2.31(1.17-4.55)	-	Age, gender, SES, race, birthplace, household size, TB contact	Not assessed	-
TB infection	cross-sectional	Brunet et al. 2011	South Africa	TB suspects	IGRA (Quantiferon TB gold)	serum cotinine levels	108	OR: 0.64 (0.14-2.79)	-	Age, gender alcohol, SES, HIV	Not assessed	-

										status		
TB disease					Sputum smear microscopy and culture; CXR		410	OR: 0.63 (0.38-1.03)	-	Age, alcohol, SES, previous TB		-
TB disease	Cohort	Leung et al. 2004	Hong Kong	Clients of the elderly health services	TB confirmed by microscopy, CXR or histology	Self-reported current smoking	252	OR: 2.87 (2.00-4.11)	-	Age, gender, alcohol, SES, living conditions, comorbidities	Yes (Quantity)	-
TB recurrence					Self-reported active TB with history of cured TB		42659	OR: 2.48 (1.04-5.89)	-		Not assessed	-
TB disease	case-control	Shetty et al. 2006	India	TB outpatients and controls were their relatives	TB diagnosed per NTP guidelines	Self-reported current smoking	378	OR: 0.80 (0.34-1.89)	-	Age, gender, alcohol, SES, living conditions, comorbidities, biomass fuel use	Not assessed	-
TB disease	case-control	Lienhardt et al. 2005	Gambia, West Africa	TB patients and controls were from their households and community	TB confirmed by microscopy	Self-reported current smoking	2325	OR: 2.54 (1.77-3.66)	-	Age, gender, alcohol, SES, BCG, comorbidities, TB contact	Yes (duration)	-
TB disease	case-control	Wang et al. 2005	China	TB patients and neighbourhood controls	TB confirmed by microscopy	Self-reported current smoking	474	OR: 1.54 (1.16-2.04)	-	Age, gender, SES	Yes (Quantity)	-
TB disease	case-control	Aryothai et al. 2004	Thailand	TB inpatients and controls from	TB confirmed by microscopy, CXR or histology	Self-reported current smoking	200	OR: 2.70 (1.04-6.97)	-	Age, alcohol, living conditions, BCG,	Yes (for both duration and quantity)	-

				outpatient /inpatient from other departments						comorbidities, TB contact, BMI		
TB disease	case-control	Kolappan et al. 2002	India	TB patients and controls were from the community	TB confirmed by microscopy or culture	Self-reported current smoking	665	OR: 2.24 (1.27-3.04)	-	Age	Yes (for both duration and quantity)	-
TB disease	case-control	Tekkel et al. 2002	Estonia	TB patients and controls were from population registry	TB diagnosed per WHO European guidelines	Self-reported current smoking	496	OR: 4.62 (2.44-8.73)	-	Age, gender, region, SES	none	-
TB disease	case-control	Dong et al. 2001	China	TB patients and controls were from the community	TB confirmed by microscopy	Self-reported current smoking	348	OR: 1.65 (1.00-2.73)	-	Age, gender, region, SES, alcohol, living conditions, BCG, comorbidities, TB contact, BMI, dust	Yes (Quantity)	-
TB disease	case-control	Gupta et al. 2001	India	TB patients; chest clinic and healthy controls	TB confirmed by microscopy, CXR or treatment response	Self-reported ever smoking	400	OR: 4.42 (2.55-7.66)	-	Age, gender, SES, TB contact	Yes (cumulative exposure)	-
TB disease	case-control	Alcaide et al. 1996	Spain	TB patients; TST +ve	TB confirmed by microscopy, CXR or TST +ve	Self-reported current smoking	92	OR: 3.60 (1.50-7.20)	-	Age, gender, SES	Yes (Quantity)	-

				controls								
TB disease	case-control	Crampin et al. 2004	Malawi	TB patients and controls were from the community	TB confirmed by microscopy or culture	Self-reported current smoking	606	OR: 1.3 (0.7-2.4)	-	Area, gender, age, HIV	none	-
TB disease	case-control	Leung et al. 2003	Hong Kong, China	TB patients; controls from household survey	TB confirmed by microscopy, CXR or histology	Self-reported current smoking	8686	OR: 2.13 (1.46-3.11)	-	none	none	-
TB disease	case-control	PerezPadilla et al. 2001	Mexico	patients	TB confirmed by microscopy or culture	Self-reported current smoking	833	OR: 1.5 (1.0-2.3)	-	urban and rural residence, crowding, education, biomass fuel use, income	none	-
TB disease	case-control	Tocque et al. 2001	UK	TB patients; controls from GP database	TB confirmed by microscopy or culture	Self-reported current smoking	310	OR: 2.33 (1.40-3.88)	-	none	none	-
TB disease	case-control	Toledo et al. 2000	Brazil	HIV patients		Self-reported current smoking	477	OR: 1.3 (1.0-1.6)	-	none	none	-
TB disease	case-control	Adelstein et al. 1967	UK	Mass X-ray volunteers	TB confirmed by CXR	Self-reported current smoking	73287	OR: 4.55 (2.4-8.6)	-	none	Yes (Quantity)	-
TB disease	case-control	Shah et al. 2003	Pakistan	prisoners			75	OR: 1.59 (0.44-5.37)	-	none		-
TB disease	cross-sectional	Yu et al. 1988	China	Sanitary workers	TB confirmed by microscopy or CXR	Self-reported current smoking	30268	OR: 2.17 (1.29-3.68)	-	TB contact, housing area, type of work	Yes (Quantity)	-
TB disease	case-control	Lowe et al. 1956	UK	patients	TB confirmed by notification	Self-reported current smoking	2179	OR: 1.61 (1.27-2.02)	-	none	Yes (Quantity)	-

TB disease	case-control	Lewis et al. 1963	UK	TB inpatients and controls inpatient from other departments	TB confirmed by microscopy	Self-reported former smoking	200	OR: 1.01 (0.55-1.85)	-	alcohol	none	-
TB disease	case-control	Brown et al. 1961	Australia	TB inpatients and controls inpatient from surgical service	not mentioned	Self-reported current smoking	200	OR 0.95 (0.45-2.02)	-	alcohol	none	-
TB disease	case-control	Gajalakshmi et al. 2003	India	urban	TB confirmed by self-report	Self-reported ever smoking	1122	OR: 2.90 (2.60-3.30)	-	Age, SES, smokeless tobacco	Yes (Quantity)	-
TB mortality					TB patients and household controls	Death due to TB confirmed by vital statistics records and/or verbal autopsy	33220	OR: 4.5 (4.0-5.0)	-		Not assessed	-
TB disease	cross-sectional	Gupta et al. 1997	India	rural and urban	TB confirmed by microscopy, CXR	Self-reported current smoking	707	OR: 1.38 (0.80-2.39)	-	age	none	-
TB disease	cross-sectional	Shah et al. 1959	India	staff	TB confirmed by MMR	Self-reported current smoking	439	OR: 2.70 (1.37-5.29)	-	none	none	-
TB disease	cross-sectional	Rao et al. 2012	India	Marginalised tribal group	TB confirmed by microscopy or culture	Self-reported current smoking	9538	OR: 1.8 (1.3-2.5)	-	Age, gender, alcohol	Not assessed	-
TB disease	Cohort	Jee et al. 2009	Korea	Korean Cancer Prevention study-adult males	TB confirmed by CXR	Self-reported current smoking	8657	HR: 1.4 (1.3-1.5)	-	Age, alcohol, BMI	Yes (Quantity)	-
TB mortality					National statistics office records for cause of death		659	HR: 1.58 (1.27-1.97)	-			-
TB recurrence					Self-reported confirmed by prior		6218	HR: 1.3 (1.2-1.4)	-			-

					hospitalisation for active TB							
TB disease	case-control	Wang et al. 2009	China	TB patients and controls were from the community	not specified	Self-reported current smoking	1839	OR: 1.93 (1.51-2.48)	-	Age, gender, alcohol, education	Yes (for both duration and quantity)	-
TB disease	case-control	Gajalakshmi et al. 2009	India	TB patients and controls were from the community	TB confirmed by criteria of state TB clinics	Self-reported current smoking	2912	RR: 2.7 (2.2-3.3)	-	Age, education	Not assessed	-
TB mortality	Cohort	Gupta et al. 2005	India	adults from voters list	Death due to TB confirmed by municipal corporation records	Self-reported ever smoking	99570	OR: 3.31 (1.34–8.16)	-	Age, gender, SES	Not assessed	-
TB mortality	case-control	Sitas et a. 2004	South Africa	TB patients and controls who died due to causes other than smoking	Death due to TB confirmed by death notification form	Self-reported smoking 5 yeaqrs prior to death	1538	OR: 1.61 (1.23-2.11)	-	Age, gender, SES, race	Not assessed	-
TB mortality	case-control	Lam et al. 2001	Hong Kong	TB patients from death registry and live controls	Death due to TB confirmed by death certificate	Self-reported ever smoking	13251	OR: 2.54 (1.24–5.22)	-	Age, gender, SES	Yes (Quantity)	-
TB mortality	case-control	Liu et al. 1998	China	TB patients and	Death due to TB confirmed by administrative/m	Self-reported ever smoking	99481	OR: 1.42 (1.33-1.52)	-	Age, gender, region,	Yes (Quantity)	-

				controls who died due to causes other than smoking	edical records and verbal autopsies					urban/rural		
TB mortality	Cohort	Wen et al. 2010	China	Standard medical screening program	Death due to TB confirmed by death records in national database	Self-reported current smoking	90580	HR: 4.19 (1.8-9.7)	-	Age, gender	none	-
TB mortality	Cohort	Mangesh et al. 2007	India	adults from voters list	Verbal autopsy	Self-reported ever smoking	81443	RR: 2.12 (1.70-2.66)	-	none	Yes (Quantity)	-
TB recurrence	Cohort	Thomas et al. 2005	India	TB patients	Sputum smear micropscopy and culture to confirm at 6, 12 and 18 months post cure	Self-reported current smoking	503	OR: 3.1 (1.6-6.0)	-	Drug compliance and drug sensitivity	Not assessed	-
TB recurrence	Cohort	D'Arc Batista et al. 2008	Brazil	TB patients	A patient who started second treatment during their follow-up for first treatment	Self-reported current smoking	711	OR: 2.53 (1.23-5.21)	-	Year of entry in cohort, area of residence	Not assessed	-

TB Infection

(Anderson et al., 1997, McCurdy et al., 1997, Adib et al., 1999, Den Boon et al., 2005, Hussain et al., 2003, Plant et al., 2002, Solsona et al., 2001, Horne et al., 2012, Al-Darraji et al., 2015, Lindsay et al., 2014, Brunet et al., 2011)

ADIB, S. M., AL-TAKASH, H. & AL-HAJJ, C. 1999. Tuberculosis in Lebanese jails: prevalence and risk factors. *European journal of epidemiology*, 15**,** 253-260.

AL-DARRAJI, H. A. A., TAN, C., KAMARULZAMAN, A. & ALTICE, F. L. 2015. Prevalence and correlates of latent tuberculosis infection among employees of a high security prison in Malaysia. *Occupational and environmental medicine*, 72**,** 442-447.

ANDERSON, R. H., SY, F. S., THOMPSON, S. & ADDY, C. 1997. Cigarette smoking and tuberculin skin test conversion among incarcerated adults. *American journal of preventive medicine*, 13**,** 175.

BRUNET, L., PAI, M., DAVIDS, V., LING, D., PARADIS, G., LENDERS, L., MELDAU, R., VAN ZYL SMIT, R., CALLIGARO, G., ALLWOOD, B., DAWSON, R. & DHEDA, K. 2011. High prevalence of smoking among patients with suspected tuberculosis in South Africa. *European Respiratory Journal*, 38**,** 139-46.

DEN BOON, S., VAN LILL, S., BORGDORFF, M., VERVER, S., BATEMAN, E., LOMBARD, C., ENARSON, D. & BEYERS, N. 2005. Association between smoking and tuberculosis infection: a population survey in a high tuberculosis incidence area. *Thorax*, 60**,** 555-557.

HORNE, D. J., CAMPO, M., ORTIZ, J. R., OREN, E., ARENTZ, M., CROTHERS, K. & NARITA, M. 2012. Association between smoking and latent tuberculosis in the U.S. population: an analysis of the National Health and Nutrition Examination Survey. *PloS one*, 7**,** e49050.

HUSSAIN, H., AKHTAR, S. & NANAN, D. 2003. Prevalence of and risk factors associated with Mycobacterium tuberculosis infection in prisoners, North West Frontier Province, Pakistan. *International journal of epidemiology*, 32**,** 794-799.

LINDSAY, R. P., SHIN, S. S., GARFEIN, R. S., RUSCH, M. L. & NOVOTNY, T. E. 2014. The Association between active and passive smoking and latent tuberculosis infection in adults and children in the united states: results from NHANES. *PLoS ONE [Electronic Resource]*, 9**,** e93137.

MCCURDY, S. A., ARRETZ, D. S. & BATES, R. O. 1997. Tuberculin reactivity among California Hispanic migrant farm workers. *American journal of industrial medicine*, 32**,** 600-605.

PLANT, A., WATKINS, R., GUSHULAK, B., O'ROURKE, T., JONES, W., STREETON, J. & SANG, D. 2002. Predictors of tuberculin reactivity among prospective Vietnamese migrants: the effect of smoking. *Epidemiology and infection*, 128**,** 37-45.

SOLSONA, J., CAYLA, J., NADAL, J., BEDIA, M., MATA, C., BRAU, J., MALDONADO, J., MILA, C., ALCAIDE, J. & ALTET, N. 2001. Screening for tuberculosis upon admission to shelters and free-meal services. *European journal of epidemiology*, 17**,** 123-128.

TB Disease

(Leung et al., 2003, Leung et al., 2004, Shetty et al., 2006, Lienhardt et al., 2005, Wang and Shen, 2009, Wang et al., 2005, Ariyothai et al., 2004, Kolappan and Gopi, 2002, Tekkel et al., 2002, Dong et al., 2001, Gupta et al., 2001, Alcaide et al., 1996, Crampin et al., 2004, Perez-Padilla et al., 2001, Tocque et al., 2001, Toledo Jr et al., 2000, Adelstein and Rimington, 1967, Shah et al.,

2003, Yu et al., 1988, Lowe, 1956, Lewis and Chamberlain, 1963, Brown and Campbell, 1961, Gajalakshmi et al., 2003, Gupta et al., 1997, Shah et al., 1959, Rao et al., 2012, Jee et al., 2009, Gajalakshmi and Peto, 2009, Brunet et al., 2011)

ADELSTEIN, A. & RIMINGTON, J. 1967. Smoking and pulmonary tuberculosis: an analysis based on a study of volunteers for mass miniature radiography. *Tubercle*, 48**,** 219-226.

ALCAIDE, J., ALTET, M., PLANS, P., PARRON, I., FOLGUERA, L., SALTO, E., DOMINGUEZ, A., PARDELL, H. & SALLERAS, L. 1996. Cigarette smoking as a risk factor for tuberculosis in young adults: a casecontrol study. *Tubercle and Lung Disease*, 77**,** 112-116.

ARIYOTHAI, N., PODHIPAK, A., AKARASEWI, P., TORNEE, S., SMITHTIKARN, S. & THONGPRATHUM, P. 2004. Cigarette smoking and its relation to pulmonary tuberculosis in adults.

BROWN, K. & CAMPBELL, A. 1961. Tobacco, alcohol and tuberculosis. *British Journal of Diseases of the Chest*, 55**,** 150-158.

BRUNET, L., PAI, M., DAVIDS, V., LING, D., PARADIS, G., LENDERS, L., MELDAU, R., VAN ZYL SMIT, R., CALLIGARO, G., ALLWOOD, B., DAWSON, R. & DHEDA, K. 2011. High prevalence of smoking among patients with suspected tuberculosis in South Africa. *European Respiratory Journal*, 38**,** 139-46.

CRAMPIN, A., GLYNN, J., FLOYD, S., MALEMA, S., MWINUKA, V., NGWIRA, B., MWAUNGULU, F., WARNDORFF, D. & FINE, P. 2004. Tuberculosis and gender: exploring the patterns in a case control study in Malawi. *The International Journal of Tuberculosis and Lung Disease*, 8**,** 194-203.

DONG, B., GE, N. & ZHOU, Y. 2001. [Smoking and alcohol consumption as risk factors of pulmonary tuberculosis in Chengdu: a matched case-control study]. *Hua xi yi ke da xue xue bao= Journal of West China University of Medical Sciences= Huaxi yike daxue xuebao/[bian ji zhe, Hua xi yi ke da xue xue bao bian wei hui]*, 32**,** 104-106.

GAJALAKSHMI, V. & PETO, R. 2009. Smoking, drinking and incident tuberculosis in rural India: population-based case-control study. *International Journal of Epidemiology*, 38**,** 1018-25.

GAJALAKSHMI, V., PETO, R., KANAKA, T. S. & JHA, P. 2003. Smoking and mortality from tuberculosis and other diseases in India: retrospective study of 43 000 adult male deaths and 35 000 controls. *The Lancet*, 362**,** 507-515.

GUPTA, B., MATHUR, N., MAHENDRA, P. & SRIVASTAVA, A. 1997. A study of household environmental risk factors pertaining to respiratory diseases. *Energy Environment Monitor*, 13**,** 61-68.

GUPTA, D., AGGARWAL, A., KUMAR, S. & JINDAL, S. 2001. Smoking increases risk of pulmonary tuberculosis. *J Environ Med*, 3**,** 65-70.

JEE, S. H., GOLUB, J. E., JO, J., PARK, I. S., OHRR, H. & SAMET, J. M. 2009. Smoking and risk of tuberculosis incidence, mortality, and recurrence in South Korean men and women. *American journal of epidemiology*, 170**,** 1478-1485.

KOLAPPAN, C. & GOPI, P. 2002. Tobacco smoking and pulmonary tuberculosis. *Thorax*, 57**,** 964-966.

LEUNG, C., YEW, W., CHAN, C., TAM, C., LAM, C., CHANG, K., CHAU, C., LAU, K. & LAW, W. 2003. Smoking and tuberculosis in Hong Kong. *The International Journal of Tuberculosis and Lung Disease*, 7**,** 980-986.

LEUNG, C. C., LI, T., LAM, T. H., YEW, W. W., LAW, W. S., TAM, C. M., CHAN, W. M., CHAN, C. K., HO, K. S. & CHANG, K. C. 2004. Smoking and tuberculosis among the elderly in Hong Kong. *American journal of respiratory and critical care medicine*, 170**,** 1027-1033.

LEWIS, J. & CHAMBERLAIN, D. 1963. Alcohol consumption and smoking habits in male patients with pulmonary tuberculosis. *British journal of preventive & social medicine*, 17**,** 149-152.

LIENHARDT, C., FIELDING, K., SILLAH, J., BAH, B., GUSTAFSON, P., WARNDORFF, D., PALAYEW, M., LISSE, I., DONKOR, S. & DIALLO, S. 2005. Investigation of the risk factors for tuberculosis: a case–control study in three countries in West Africa. *International Journal of Epidemiology*, 34**,** 914-923.

LOWE, C. 1956. An association between smoking and respiratory tuberculosis. *British medical journal*, 2**,** 1081.

PEREZ-PADILLA, R., PEREZ-GUZMAN, C., BAEZ-SALDANA, R. & TORRES-CRUZ, A. 2001. Cooking with biomass stoves and tuberculosis: a case control study. *The International Journal of Tuberculosis and Lung Disease*, 5**,** 441-447.

RAO, V. G., GOPI, P. G., BHAT, J., YADAV, R., SELVAKUMAR, N. & WARES, D. F. 2012. Selected risk factors associated with pulmonary tuberculosis among Saharia tribe of Madhya Pradesh, central India. *European Journal of Public Health*, 22**,** 271-3.

SHAH, J., WARAWADEKAR, M., DESHMUKH, P. & PHUTANE, P. 1959. Institutional survey of pulmonary tuberculosis with special reference to smoking habits. *Indian journal of medical sciences*, 13**,** 381.

SHAH, S., MUJEEB, S., MIRZA, A., NABI, K. & SIDDIQUI, Q. 2003. Prevalence of pulmonary tuberculosis in Karachi juvenile jail, Pakistan.

SHETTY, N., SHEMKO, M., VAZ, M. & D'SOUZA, G. 2006. An epidemiological evaluation of risk factors for tuberculosis in South India: a matched case control study. *The International Journal of Tuberculosis and Lung Disease*, 10**,** 80-86.

TEKKEL, M., RAHU, M., LOIT, H. & BABURIN, A. 2002. Risk factors for pulmonary tuberculosis in Estonia. *The International Journal of Tuberculosis and Lung Disease*, 6**,** 887-894.

TOCQUE, K., BELLIS, M., BEECHING, N., SYED, Q., REMMINGTON, T. & DAVIES, P. 2001. A case-control study of lifestyle risk factors associated with tuberculosis in Liverpool, North-West England. *European Respiratory Journal*, 18**,** 959-964.

TOLEDO JR, A. C. D. C., GRECO, D. B. & ANTUNES, C. M. F. 2000. Risk factors for tuberculosis among human immunodeficiency virus-infected persons. A case-control study in Belo Horizonte, Minas Gerais, Brazil (1985-1996). *Memórias do Instituto Oswaldo Cruz*, 95**,** 437-443.

WANG, G., SLEIGH, A., ZHOU, G., JACKSON, S. & LIU, X. 2005. [Nonbiologic risk factors of pulmonary tuberculosis among adults in Henan: a case-control study]. *Zhonghua liu xing bing xue za zhi= Zhonghua liuxingbingxue zazhi*, 26**,** 92-96.

WANG, J. & SHEN, H. 2009. Review of cigarette smoking and tuberculosis in China: intervention is needed for smoking cessation among tuberculosis patients. *BMC public health*, 9**,** 292.

YU, G.-P., HSIEHT, C.-C. & PENG, J. 1988. Risk factors associated with the prevalence of pulmonary tuberculosis among sanitary workers in Shanghai. *Tubercle*, 69**,** 105-112.

TB Mortality

(Gajalakshmi et al., 2003, Jee et al., 2009, Gupta et al., 2005, Sitas et al., 2004, Lam et al., 2001, Liu et al., 1998, Wen et al., 2010, Pednekar and Gupta, 2007)

GAJALAKSHMI, V., PETO, R., KANAKA, T. S. & JHA, P. 2003. Smoking and mortality from tuberculosis and other diseases in India: retrospective study of 43 000 adult male deaths and 35 000 controls. *The Lancet*, 362**,** 507-515.

GUPTA, P. C., PEDNEKAR, M. S., PARKIN, D. & SANKARANARAYANAN, R. 2005. Tobacco associated mortality in Mumbai (Bombay) India. Results of the Bombay cohort study. *International journal of epidemiology*, 34**,** 1395-1402.

JEE, S. H., GOLUB, J. E., JO, J., PARK, I. S., OHRR, H. & SAMET, J. M. 2009. Smoking and risk of tuberculosis incidence, mortality, and recurrence in South Korean men and women. *American journal of epidemiology*, 170**,** 1478-1485.

LAM, T., HO, S., HEDLEY, A., MAK, K. & PETO, R. 2001. Mortality and smoking in Hong Kong: case-control study of all adult deaths in 1998. *Bmj*, 323**,** 361.

LIU, B.-Q., PETO, R., CHEN, Z.-M., BOREHAM, J., WU, Y.-P., LI, J.-Y., CAMPBELL, T. C. & CHEN, J.-S. 1998. Emerging tobacco hazards in China: 1. Retrospective proportional mortality study of one million deaths. *Bmj*, 317**,** 1411-1422.

PEDNEKAR, M. S. & GUPTA, P. C. 2007. Prospective study of smoking and tuberculosis in India. *Preventive medicine*, 44**,** 496-498.

SITAS, F., URBAN, M., BRADSHAW, D., KIELKOWSKI, D., BAH, S. & PETO, R. 2004. Tobacco attributable deaths in South Africa. *Tobacco Control*, 13**,** 396-399.

WEN, C. P., CHAN, T. C., CHAN, H. T., TSAI, M. K., CHENG, T. Y. & TSAI, S. P. 2010. The reduction of tuberculosis risks by smoking cessation. *BMC infectious diseases*, 10**,** 156.

TB Recurrence

(Leung et al., 2004, Jee et al., 2009, Thomas et al., 2005, d'Arc Lyra Batista et al., 2008)

D'ARC LYRA BATISTA, J., DE FATIMA PESSOA MILITAO DE ALBUQUERQUE, M., DE ALENCAR XIMENES, R. A. & RODRIGUES, L. C. 2008. Smoking increases the risk of relapse after successful tuberculosis treatment. *International Journal of Epidemiology*, 37**,** 841-51.

JEE, S. H., GOLUB, J. E., JO, J., PARK, I. S., OHRR, H. & SAMET, J. M. 2009. Smoking and risk of tuberculosis incidence, mortality, and recurrence in South Korean men and women. *American journal of epidemiology*, 170**,** 1478-1485.

LEUNG, C. C., LI, T., LAM, T. H., YEW, W. W., LAW, W. S., TAM, C. M., CHAN, W. M., CHAN, C. K., HO, K. S. & CHANG, K. C. 2004. Smoking and tuberculosis among the elderly in Hong Kong. *American journal of respiratory and critical care medicine*, 170**,** 1027-1033.

THOMAS, A., GOPI, P., SANTHA, T., CHANDRASEKARAN, V., SUBRAMANI, R., SELVAKUMAR, N., EUSUFF, S., SADACHARAM, K. & NARAYANAN, P. 2005. Predictors of relapse among pulmonary tuberculosis patients treated in a DOTS programme in South India. *The International Journal of Tuberculosis and Lung Disease*, 9**,** 556-561.

Appendix B. OVERVIEW OF ASSIST STUDY

STUDY BACKGROUND

Pakistan features in both the top ten high-burden TB and top ten high-burden tobacco use countries (World Health Organization, 2009). It has one of the highest TB burdens worldwide, with approximately 500,000 incident cases and 100,000 deaths annually (World Health Organisation, 2014b). Smoking is also prevalent in Pakistan.

STUDY POPULATION AND SETTING

ASSIST was conducted in 33 TB clinics/clusters (each with a catchment population ranging between 30,000- 450,000) in Pakistan between 2010 and 2011 that enrolled 1955 TB suspects who were also smokers (Siddiqi et al., 2013).

One TB paramedic was appointed per cluster (TB clinic) and was responsible for delivering behavioural support for smoking cessation to the patients attending that clinic.

STUDY DESIGN

ASSIST was an effectiveness cluster RCT trial that showed the effect of a behavioural support intervention (with and without bupropion) delivered by the TB DOTS paramedics in achieving sustained abstinence from smoking by the TB suspects at 6 months.

INTERVENTION AND CONTROL

Those who consented to participate in the trial were randomized to three groups: patients in one group received two brief behavioural support sessions (BSS group), patients in the second group received two brief BSS plus 7 weeks of bupropion therapy (BSS+ group), and patients in the control group received usual care. Given the lack of any routine advice or educational materials in Pakistan, a self-help leaflet on smoking cessation was offered to all participants (for details, refer to the Supplement 2: 'Tobacco Leaflet', available at http://annals.org/article.aspx?articleid=1684852; Siddiqi et al., 2013).

The intervention sessions consisted of two structured patient-provider interaction sessions based on the WHO's "5 As quit model": Ask- about smoking behaviour and history; Advise- about the consequences of smoking and cessation; Assess- willingness to quit smoking; Assist- in planning to quit smoking; and Arrange- follow up. The intervention was delivered by TB DOTS paramedics, following one full day's training on intervention protocol and delivery tools, using a flipbook resource developed for this purpose (for details, refer to the Supplement 1: 'Tobacco Cessation Desk Guide: Five Steps to quit', available at http://annals.org/article.aspx?articleid=1684852; Siddiqi et al., 2013).

Two sessions were delivered; one pre-quit and the other coinciding with the quit date. The aim of the first 30 minute session was to assist a smoker who was willing to set a quit date (a week after the first contact), by encouraging them to see themselves as a non-smoker, planning for their quit day and preparing them for the initial stages of the quit attempt. The second 10 minute session was scheduled for the patient's quit date and provided an opportunity for follow-up and review of progress. The content of the interactive sessions was designed to deliver information about the harms of tobacco smoking, such as effects on pregnancy and lung diseases, the social and economic benefits of stopping, identification of the social/psychological or environmental cues that trigger a smoker's desire to smoke and advice on ways to address and overcome these. Similarly, assessing the withdrawal symptoms (i.e. the strong urges/craving to smoke) was done during the session, and advice given on ways to address and overcome these.

In addition to BSS, participants in the BSS+ group also received sustained-release bupropion, 75 mg/day for the first week and 150 mg/day thereafter for six more weeks.

PRIMARY OUTCOME

The primary end-point was continuous smoking cessation at six months after the quit date, defined as an expired carbon monoxide (CO) measurement of 9 ppm or less (according to the Russell standard) at the 1- and 6-month follow-up visits.(West et al., 2005)

STUDY FINDINGS

Both treatment conditions led to 8-9 fold increase (BSS+: RR 9.3, 95%CI: 4.0-21.6; BSS: RR 8.5, 95%CI: 3.7-19.6) in continuous smoking abstinence compared to the control (Table B.1).

B.1: Continuous smoking abstinence at 6 months- ASSIST trial (2010- 2011) Pakistan

Trial arm	No. Abstinent/ N (%)	Relative Risk (95% CI)	[7]ICC
Behavioural support + bupropion (BSS+)	275/606 (45.4)	9.3 (4.0- 21.6)	0.28
Behavioural support alone (BSS)	254/620 (41.0)	8.5 (3.7- 19.6)	
Usual care control	52/615 (8.5)	1	

[7] Intra-Cluster correlation Coefficient

Appendix C. APPENDICES TO CHAPTER 6

C.1: Letter to the experts for modified Delphi procedure

Hi

Thanks for agreeing to participate in the Delphi procedure.

The fidelity criterion under development is specific to a behavioural support intervention delivered for smoking cessation to patients in primary health care settings, in Pakistan. The intervention was found to be highly effective in achieving smoking cessation in the patients. Nevertheless, there were wide differences between the cessation rates of various health clinics, where the intervention was delivered. My PhD work relates to these findings and a way forward in describing these differences.

I am attaching the fidelity criterion checklist and the scale plan (outlining the intervention dimensions) for your reference. We are at the step 4 of establishing the 'fidelity criterion'. As you will see in the scale plan, the behavioural support intervention is based on the World Health Organisation's 5 steps to quit model (5 A's: Ask, advise, assess, assist and arrange). At the time this intervention was developed the taxonomy of the behavioural change techniques (BCT) for smoking cessation (by Susan Michie and colleagues) was not established. I have linked each 'item' of the intervention with the BCTs. Once the fidelity criterion is finalized, I am hoping to record sessions on behavioural support intervention and score these using the developed fidelity criterion. The data will be used to test the tool's reliability and validity.

Now when you open the item pool for the fidelity criterion matrix, the first column describes what the item corresponds to (in the questionnaire or the flip-chart), in the original intervention. Some of the items are about the adherence to the intervention content, while others assess the competence (quality of delivery). The 2nd, 3rd and 4th columns are the BCT code, its description and the underlying domain of behaviour change. The 5th (item) and the 7th (last column on expert rating of item) are mainly concerned with you. Looking at each of the item, you have to decide on their importance in assessing fidelity of a behavioural support intervention for smoking cessation delivered in a primary health care setting; you will need to rate these on a scale of 4: (very, somewhat, not much, or not at all important). Once you and other experts return the checklist to me, I will drop those items rated by at least two experts as 'not at all important' or 'not much important' and then re-send it to you for the next round and so on. I anticipate achieving the required consensus on important items in two to three rounds.

Please let me know if you have any questions.

Many thanks

C.2: Experts rating on 'importance' of items for measuring fidelity

Behavioural Determinants	Behaviour Change Techniques (BCTs)	BCTs description	Items	Response scale	Experts rating (1= very important, 2= somewhat important, 3= not much important, 4= not at all important)			
					1	**2**	**3**	**4**
ADHERENCE								
Step 1- ASK (Status of Tobacco use)								
General aspects of the interaction (R) focusing on information gathering (I)	RC7: Information gathering and assessment	Any information gathering that provides the practitioner with the knowledge needed from the client for appropriate BCT to be delivered. * Assess current and past smoking behaviour	1. Assessing current and past smoking behaviour i. Pattern of smoking behaviour (Types of smoking? Smokers in vicinity? Children at home?) ii. Age when started smoking iii. Amount smoked	**0= not implemented** (skips all 3) **1= partially implemented** (asks about the patterns of smoking behaviour, age of starting smoking and the amount smoked without eliciting a response using categories given in the questionnaire) **2= fully implemented** (asks about the above and elicits response using categories given in the Qx; about the type of smoking by asking them to choose from cigarette, hookah, bidi, cigar or other; appropriately captures the current amount of smoking form used)	*| |	|		
Step 2- ADVISE (Risks of tobacco use and benefits of cessation). The flip book slides are designed to be used in a specific way; while the sketch on each slide faces the patient, the written material is facing the provider to help them deliver the message effectively. Each slide is designed to be delivered in three essential steps: **1- Ask the patient to describe the slide** **2- Facilitate patient with understanding the message in the slide** **3- Clarify/re-emphasise the key message in the slide**								
Specific focus on behaviour (B) and	BM1: Provide information on		2. Awareness about the various forms of tobacco smoked in the community	**0= not implemented** (skips the slide) **1= partially implemented** (delivers 2 out of 3 steps) **2= fully implemented** (all 3 essential steps)	|	| |		

addressing motivation (M)	consequences of smoking and smoking cessation	Give, or make more salient information about the harm caused by smoking and the benefits of stopping; distinguish between the harms from smoking and nicotine; debunk myths about low tar and own roll cigarettes	3. High blood pressure and heart disease	**0= not implemented** **1= partially implemented** **2= fully implemented** (same as above)	|	||		
			4. Lung diseases like chronic cough, asthma, TB and cancer	**0= not implemented** **1= partially implemented** **2= fully implemented** (same as above)	|	||		
			5. Wastage of money, staining of teeth, gum problems and bad breath	**0= not implemented** **1= partially implemented** **2= fully implemented** (same as above)		||	|	
			6. Effects on children's health: pneumonia, asthma etc.	**0= not implemented** **1= partially implemented** **2= fully implemented** (same as above)	|	||		
			7. Effects on pregnancy: complications in pregnancy, low birth weight baby	**0= not implemented** (if skipped, please mention in the comments if the patient was a male) **1= partially implemented** **2= fully implemented** (same as above)	|	||		
			8. Decide to quit, choose a quit date and utilize the money on better things	**0= not implemented** **1= partially implemented** **2= fully implemented** (same as above)	|	||		
			9. Social and economic benefits of quitting	**0= not implemented** **1= partially implemented** **2= fully implemented** (same as above)		||	|	
Step-3: ASSESS (Willingness to quit)								
General aspects of the interaction (R) focusing on information gathering (I)	RC7: Information gathering and assessment	Any information gathering that provides the practitioner with the knowledge needed from the client for appropriate BCT to be delivered. * Assess current readiness	10. Current level of motivation to stop/willingness to quit	**0= not implemented** **2= implemented**	|||			
			11. Reasons for quitting e.g. health, cost, example for others, family's health or other reason	**0= not implemented** **1= partially implemented** (asks about the reasons for quitting without eliciting a response using categories given in the questionnaire) **2= fully implemented**	|||			

		and ability to quit		(asks about the above and elicits response using categories given in the Qx)				
Specific focus on behaviour (B) and addressing motivation (M)	BM6: prompt commitment from the client there and then	Encourage the smoker to affirm or re-affirm a strong commitment to start, continue or restart the quit attempt	12. Quit from today? If 'No', Quit within next five days? If 'No', when will you be able to set a quit date?	**0= not implemented** **2= implemented**	| | |			
General aspects of the interaction (R) focusing on information gathering (I)	RC7: Information gathering and assessment	Any information gathering that provides the practitioner with the knowledge needed from the client for appropriate BCT to be delivered. * Assess past history of quit attempts	13. Attempted quit in the past, Number and duration of past quit attempts, Time since last quit attempt	**0= not implemented** **1= partially implemented** (asks 1 out of the 3 Q's) **2= fully implemented** (asks all 3 Q's)	|	| |		
			14. Factors that led back to smoking including social and physical factors (Social reasons: Family problems (tension), Smokers company, any other; Physical Symptoms: Craving, Indigestion, Insomnia, Headache, any other)	**0= not implemented** **1= partially implemented** (asks about the reasons to start again without eliciting a response using categories given in the questionnaire) **2= fully implemented** (asks about the above and elicits response using categories given in the Qx)	|	| |		
Step-4: ASSIST (in the quitting)								
General aspects of the interaction (R) focusing on information gathering (I)	RC7: Information gathering and assessment	Any information gathering that provides the practitioner with the knowledge needed from the client for appropriate BCT to be delivered. * Assess nicotine dependence	15. Fagerstrom Test for Nicotine Dependence	**0= not implemented** **1= partially implemented** (asks less than 6 Q's on the scale) **2= fully implemented** (completes the scale by asking all 6 Q's)	|	|	|	
Specific focus on behaviour(B) and maximising self-regulatory capacity/skills (S)	BS4: Facilitate goal setting	Help the smoker to set a quit date and goals that support the aim of remaining abstinent	16. Setting quit date	**0= not implemented** **2= implemented**	| | |			

Specific focus on behaviour (B) and maximising self-regulatory capacity/skills (S)	BS8: Advise on environmental restructuring	Advise on ways of changing the physical environment to minimize exposure to smoking cues (e.g. removing ashtrays from the house) or to provide cues to sustain quitting	17. Hide reminders of smoking/ways of changing the physical environment to minimise exposure to smoking cues	**0= not implemented** **2= implemented**	| |	|		
			18. Declare the house as 'smoke free home'	**0= not implemented** **2= implemented**		| | |		
Promote adjuvant activities (A)	A2: Advise on use of social support	Advise on or facilitate development of social support from friends, relatives, colleagues or "buddies"	19. Identify individual who can help support in quitting at home	**0= not implemented** **2= implemented**	|	|	|	
Specific focus on behaviour (B) and addressing motivation (M)	BM2: Boost motivation and self-efficacy	Give encouragement or bolster confidence in ability to stop. Can include telling the person that they can successfully stop smoking, arguing against self-doubts, and asserting that they can and will succeed.	20. Decide on telling people about stopping or keeping it private	**0= not implemented** **2= implemented**	|	|	|	
Specific focus on behaviour (B) and maximising self-regulatory capacity/skills	BS1: Facilitate barrier identification and problem solving BS3: Facilitate action planning /develop treatment plan	■ Help the smoker to identify general barriers (e.g. susceptibility to stress) that might make it harder to stay off cigarettes and develop general ways of addressing and overcoming these, and increasing facilitators (e.g. by generating alternative courses of action and pros and cons of each and weighing them up)	21. Trigger1: immediately after rising in the morning- Offer strategies to manage trigger 1	**0= not implemented** (skips the slide) **1= partially implemented** (2 out of 3 steps delivered) **2= fully implemented** (all 3 essential steps)	|	| |		
			22. Trigger2: defecation- Offer strategies to manage trigger 2	**0= not implemented** **1= partially implemented** **2= fully implemented** (same as above)	|	|	|	
			23. Trigger3: eating meals- Offer strategies to manage trigger 3	**0= not implemented** **1= partially implemented** **2= fully implemented** (same as above)	|	|	|	
			24. Trigger4: free at home or feeling	**0= not implemented** **1= partially implemented**	|	|	|	

(S)		■ Work with the smoker to generate a clear quit plan, including preparations for the quit attempt	bored- Offer strategies to manage trigger 4	**2= fully implemented** (same as above)				
	BS1 & 3 ; BS11: Advise on avoiding cues for smoking BS12: Facilitate restructuring of social life	In addition to BS1 & 3 ; ■ Give specific advice on how to avoid being exposed to social or other cues for smoking (e.g. staying away from places where people smoke) ■ Advise on ways of changing social interactions with family, friends, and colleagues so that they support, rather than interfere with, the goal of remaining abstinent	25. Trigger5: seeing others smoke- Offer strategies to manage trigger 5	**0= not implemented** **1= partially implemented** **2= fully implemented** (same as above)	|	| |		
			26. Trigger6: offered smoking by others- Offer strategies to manage trigger 6	**0= not implemented** **1= partially implemented** **2= fully implemented** (same as above)	|	| |		
	BS1 & 3; BS10: Advise on conserving mental resources	In addition to BS1 & 3; Advise on ways of minimizing stress and other demands on mental resources (activities that require mental effort)	27. Trigger7: intense physical/mental work- Offer strategies to manage trigger 7	**0= not implemented** **1= partially implemented** **2= fully implemented** (same as above)	|	| |		
	BS1, 3 & 10; BS14: Teach relaxation techniques	In addition to BS1, 3 & 10; Teach specific relaxation techniques and how and when to apply them	28. Trigger8: tense/anxious- Offer strategies to manage trigger 8	**0= not implemented** **1= partially implemented** **2= fully implemented** (same as above)	| |	|		
General aspects of the interaction (R) focusing on general	RC10: Provide information on withdrawal symptoms		29. Withdrawal1: Craving smoking- Offer strategies to manage withdrawal1	**0= not implemented** **1= partially implemented** **2= fully implemented** (same as above)	| | |			
			30. Withdrawal2:	**0= not implemented** **1= partially implemented**	| | |			

communication (C)		Describe to the smoker what are, and are not, nicotine withdrawal symptoms, how common they are, how long they typically last, what causes them, and what can be done to alleviate them	Restlessness/anger- Offer strategies to manage withdrawal2	**2= fully implemented** (same as above)				
			31. Withdrawal3: Headache- Offer strategies to manage Withdrawal3	**0= not implemented** **1= partially implemented** **2= fully implemented** (same as above)	|	||		
			32. Withdrawal4: Insomnia- Offer strategies to manage withdrawal4	**0= not implemented** **1= partially implemented** **2= fully implemented** (same as above)	||	|		
			33. Withdrawal5: indigestion- Offer strategies to manage withdrawal5	**0= not implemented** **1= partially implemented** **2= fully implemented** (same as above)	|	||		
			34. Withdrawal6: Anorexia and constipation- Offer strategies to manage withdrawal6	**0= not implemented** **1= partially implemented** **2= fully implemented** (same as above)	|	|	|	
			35. Withdrawal7: Cough- Offer strategies to manage withdrawal7	**0= not implemented** **1= partially implemented** **2= fully implemented** (same as above)	|	|	|	
General aspects of the interaction (R) focusing on general communication (C)& Specific focus on behaviour (B) and maximising self-regulatory capacity/skills (S)	RC10; BS13: Advise on methods of weight control	In addition to RC10; Advise on ways of minimising weight gain that do not increase motivation to smoke (e.g. take exercise, carry around "healthy snacks")	36. Withdrawal8: Weight gain- Offer strategies to manage withdrawal8	**0= not implemented** **1= partially implemented** **2= fully implemented** (same as above)	||	|		

Step-5: ARRANGE (follow-up)								
General aspects of the interaction (R) focusing on general communication (C)	RC6: Offer appropriate written material	Distinguish what are, and are not, appropriate written materials and offer/direct clients to these in ways that promote their effective use	37. Give BI information leaflet to the patient and brief about it	**0= not implemented** **2= implemented**	|	||		
QUALITY								
General aspects of the interaction (R) focusing on general communication (C)	RC1: Build general rapport	• The provider introduces him/herself to the patient and takes time to ease the patient by asking about their day (e.g. how was their day, health etc.; how was the travel to the centre; how are they doing) and appreciating their participation (e.g. by thanking them for taking time out and how useful their participation is). • Pays complete attention to the patient when s/he is talking without cutting their talk short. ■ Listens to the patient after asking a question and gives them enough time to answer back. ■ If the patient drifts from the topic the provider politely reminds the patient to focus on the topic without disregarding their talk. • Gives them feedback, prompts or probes, taking part in the discussion (e.g. encouraging the patient, verbally responding to patient's concerns at the right time, showing concern for patient's issues and negative experiences relating to smoking/cessation).	38. Build general rapport: Establish a positive, friendly and professional relationship with the smoker and foster a sense that the smoker's experiences are understood	**0= not implemented** (does not make any effort to build rapport with the patient) **1= partially implemented** **2= fully implemented** (makes all of the described examples of efforts to build rapport with the patient)	||	|		
	RC9: Explain expectations regarding treatment programme	■ Does the provider remind the patient of the information provided earlier about the intervention (before consent)? Explaining to the patient that the treatment program consists of five steps, Ask, Advise, Assess Assist, and Arrange/ the management to help you change your smoking habit. ■ Before each step does he take a moment to give a brief introduction to what is going to happen now? Taking the	39. Explain expectations regarding treatment programme: Explain to the smoker the treatment program, what it involves, the active ingredients, and what it requires of the smoker	**0= not implemented** **1= partially implemented** **2= fully implemented**	|||			

		patient along and making an effort that the key intervention messages are being understood as intended.						
	RC4: Provide reassurance	If the patient has experienced unsuccessful quit attempts or other negative experiences relating to smoking in the past, the provider: ■ Tells them that it's ok ■ Reassures them that this is normal when individuals try to quit on their own ■ Encourages them to be persistent and carry on with their attempt as this is time limited ■ Reassures them that with assistance through smoking cessation programme they are more likely to have a successful outcome ■ Provides examples from other smokers' experiences who were successful in quitting	40. Provide reassurance: Give general reassurance to the smoker that his/her experiences are normal and time limited, and provide positive expectations of success based on experience with other smokers in the same situation	**0= not implemented** (does not provide any reassurance) **1= partially implemented** (listens and answers in 'Yes' and 'No' but does not give any constructive advice) **2= fully implemented** (gives constructive advice)	|	| |		
	RC2: General practitioner communication approaches	■ Is the provider supporting his questions with clear prompts to get the right information out of the patient e.g. when he asks, "What is your monthly household income?" If the patient says 10000, the provider confirms from him if this is his cumulative household income of all working individuals at home or his alone. Then he clarifies with the patient that he needs to tell about the cumulative household income. ■ Does he prompt the patient to give views on smoking, smoking cessation and any aspects of this behavioural support programme e.g. asking the patient, "Do you feel confident about this programme that it will help you in quitting smoking".	41. General practitioner communication approaches: (Elicit and answer questions)	**0= not implemented** **1= partially implemented** **2= fully implemented**	| |	|		
		Adopt a style of interaction that involves listening carefully to the smoker and where appropriate reflecting back to the smoker key elements of what s/he is saying e.g. the patient says, "When I don't smoke I have a very heavy feeling after meals". So the provider should respond by saying, "You mean to say that when you attempt quitting smoking you get indigestion".	42. General practitioner communication approaches: (Use reflective listening)	**0= not implemented** **1= partially implemented** **2= fully implemented**	| |	|		

		Provide a summary of information exchanged and establish a clear confirmation of decisions made and commitments entered into at each step and also at the end e.g. the provider reconfirms the quit date with the patient and arranges the follow up date reaffirming the milestones agreed upon during intervention session.	43. General practitioner communication approaches: (Summarising information and confirming client decisions)	**0= not implemented 1= partially implemented 2= fully implemented**	||	|		
General aspects of the interaction (R) focusing on the delivery of the intervention (D)	RC5: Tailor interactions appropriately	Use relevant information from the client to tailor the behavioural support provided/ flexible adaptation that takes into account individual patient needs	44. Where appropriate and needed, the provider, tailors the BI according to the patients' needs e.g. in original trial some providers recommended the use of "cardamom" for chewing in place of "chewing gum", due to limited availability of chewing gums in those areas.	**0= not implemented 1= partially implemented 2= fully implemented**	||	|		
	RC3: Emphasise choice	While delivering the management strategies for dealing with various triggers and withdrawals the provider emphasises patient's choice in the type of strategy that best suits their needs rather than imposing the strategy s/he think best. If the patient is not particularly responsive then the provider emphasises each management option available and prompts the patient to choose rather than just reading out the options and deciding for the patient him/herself.	45. Emphasise choice: Emphasise client choice within the bounds of evidence based practice	**0= not implemented 1= partially implemented 2= fully implemented**	||	|		
* Tally represents each expert								

C.3: The Fidelity Index

Fidelity Index	
TB clinic ID	
Recording number	
Name of Coder ***Please insert your name here***	
Date of Coding	
Coder ID	
Total recording duration **** please enter the total amount of time (hours/mins) which it took to complete coding***	

Q#; Slide #	**Behavioural Determinants**	**BCT code: label**	**BCT description/ Operational definition**	**Items** (ingredients of BI)	**Response scale**	**Score**	**Comment**
ADHERENCE							
Step 1- ASK (Status of Tobacco use)							
Section1A, 1B, 2A and 2B	General aspects of the interaction (R) focusing on information gathering (I)	RC7: Information gathering and assessment	Any information gathering that provides the practitioner with the knowledge needed from the client for appropriate BCT to be delivered. * Assess current and past smoking behaviour	1. Assessing current and past smoking behaviour i. Pattern of smoking behaviour (Types of smoking? Smokers in vicinity? Children at home?) ii. Age when started smoking iii. Amount smoked	**0= not implemented** (skips all 3) **1= partially implemented** (asks about the patterns of smoking behaviour, age of starting smoking and the amount smoked without eliciting a response using categories given in the questionnaire) **2= fully implemented** (asks about the above and elicits response using categories given in the Qx; about the type of smoking by asking them to choose from cigarette, hookah, bidi, cigar or other; appropriately captures the current amount of smoking form used)		

Step 2- ADVISE (Risks of tobacco use and benefits of cessation). The flip book slides are designed to be used in a specific way; while the sketch on each slide faces the patient, the written material is facing the provider to help them deliver the message effectively. Each slide is designed to be delivered in three essential steps: **1- Ask the patient to describe the slide** **2- Facilitate patient with understanding the message in the slide** **3- Clarify/re-emphasise the key message in the slide**							
S 1	Specific focus on behaviour (B) and addressing motivation (M)	<u>BM1</u>: Provide information on consequences of smoking and smoking cessation	Give, or make more salient information about the harm caused by smoking and the benefits of stopping; distinguish between the harms from smoking and nicotine; debunk myths about low tar and own roll cigarettes	2. Awareness about the various forms of tobacco smoked in the community	**0= not implemented** (skips the slide) **1= partially implemented** (delivers 2 out of 3 steps) **2= fully implemented** (all 3 essential steps)		
S 2				3. High blood pressure and heart disease	**0= not implemented** **1= partially implemented** **2= fully implemented** (same as above)		
S 3				4. Lung diseases like chronic cough, asthma, TB and cancer	**0= not implemented** **1= partially implemented** **2= fully implemented** (same as above)		
S 4				5. Wastage of money, staining of teeth, gum problems and bad breath	**0= not implemented** **1= partially implemented** **2= fully implemented** (same as above)		
S 5				6. Effects on children's health: pneumonia, asthma etc.	**0= not implemented** **1= partially implemented** **2= fully implemented** (same as above)		
S 6				7. Effects on pregnancy: complications in pregnancy, low birth weight baby	**0= not implemented** (if skipped, please mention in the comments if the patient was a male) **1= partially implemented** **2= fully implemented** (same as above)		
S 7				8. Decide to quit, choose a quit date and utilize the money on better things	**0= not implemented** **1= partially implemented** **2= fully implemented** (same as above)		
S 8				9. Social and economic benefits of quitting	**0= not implemented** **1= partially implemented** **2= fully implemented** (same as above)		

Step-3: ASSESS (Willingness to quit)							
Section3A, Q-1.	General aspects of the interaction (R) focusing on information gathering (I)	RC7: Information gathering and assessment	Any information gathering that provides the practitioner with the knowledge needed from the client for appropriate BCT to be delivered. * Assess current readiness and ability to quit	10. Current level of motivation to stop/willingness to quit	**0= not implemented** **2= implemented**		
Section 3B.				11. Reasons for quitting e.g. health, cost, example for others, family's health or other reason	**0= not implemented** **1= partially implemented** (asks about the reasons for quitting without eliciting a response using categories given in the questionnaire) **2= fully implemented** (asks about the above and elicits response using categories given in the Qx)		
Section 3A	Specific focus on behaviour (B) and addressing motivation (M)	BM6: prompt commitment from the client there and then	Encourage the smoker to affirm or re-affirm a strong commitment to start, continue or restart the quit attempt	12. Quit from today? If 'No', Quit within next five days? If 'No', when will you be able to set a quit date?	**0= not implemented** **2= implemented**		
Section 3C.	General aspects of the interaction (R) focusing on information gathering (I)	RC7: Information gathering and assessment	Any information gathering that provides the practitioner with the knowledge needed from the client for appropriate BCT to be delivered. * Assess past history of quit attempts	13. Attempted quit in the past Number and duration of past quit attempts Time since last quit attempt	**0= not implemented** **1= partially implemented** (asks 1 out of the 3 Q's) **2= fully implemented** (asks all 3 Q's)		
				14. Factors that led back to smoking including social and physical factors Social reasons: Family problems (tension), Smokers company, any other Physical Symptoms: Craving, Indigestion, Insomnia, Headache, any other	**0= not implemented** **1= partially implemented** (asks about the reasons to start again without eliciting a response using categories given in the questionnaire) **2= fully implemented** (asks about the above and elicits response using categories given in the Qx)		
Step-4: ASSIST (in the quitting)							

Section 4A.	General aspects of the interaction (R) focusing on information gathering (I)	RC7: Information gathering and assessment	Any information gathering that provides the practitioner with the knowledge needed from the client for appropriate BCT to be delivered. * Assess nicotine dependence	15. Fagerstrom Test for Nicotine Dependence	**0= not implemented** **1= partially implemented** (asks less than 6 Q's on the scale) **2= fully implemented** (completes the scale by asking all 6 Q's)		
Section 4C.	Specific focus on behaviour (B) and maximising self-regulatory capacity/skills (S)	BS4: Facilitate goal setting	Help the smoker to set a quit date and goals that support the aim of remaining abstinent	16. Setting quit date	**0= not implemented** **2= implemented**		
S 8 (is text only)	Specific focus on behaviour (B) and maximising self-regulatory capacity/skills (S)	BS8: Advise on environmental restructuring	Advise on ways of changing the physical environment to minimize exposure to smoking cues (e.g. removing ashtrays from the house) or to provide cues to sustain quitting	17. Hide reminders of smoking/ways of changing the physical environment to minimise exposure to smoking cues	**0= not implemented** **2= implemented**		
				18. Declare the house as 'smoke free home'	**0= not implemented** **2= implemented**		
	Promote adjuvant activities (A)	A2: Advise on use of social support	Advise on or facilitate development of social support from friends, relatives, colleagues or "buddies"	19. Identify individual who can help support in quitting at home	**0= not implemented** **2= implemented**		
	Specific focus on behaviour (B) and addressing motivation (M)	BM2: Boost motivation and self-efficacy	Give encouragement or bolster confidence in ability to stop. Can include telling the person that they can successfully stop smoking, arguing against self-doubts, and asserting that they can and will succeed.	20. Decide on telling people about stopping or keeping it private	**0= not implemented** **2= implemented**		
S 9	Specific focus on behaviour (B) and maximising self-regulatory capacity/skills (S)	BS1: Facilitate barrier identification and problem solving	■ Help the smoker to identify general barriers (e.g. susceptibility to stress) that might make it harder to stay off cigarettes and develop general ways of addressing and overcoming these, and	21. Trigger1: immediately after rising in the morning - Offer strategies to manage trigger 1	**0= not implemented** (skips the slide) **1= partially implemented** (2 out of 3 steps delivered) **2= fully implemented** (all 3 essential steps)		

S 10		BS3: Facilitate action planning /develop treatment plan	increasing facilitators (e.g. by generating alternative courses of action and pros and cons of each and weighing them up) ■ Work with the smoker to generate a clear quit plan, including preparations for the quit attempt	22. Trigger2: defecation - Offer strategies to manage trigger 2	**0= not implemented** **1= partially implemented** **2= fully implemented** (same as above)		
S 11				23. Trigger3: eating meals - Offer strategies to manage trigger 3	**0= not implemented** **1= partially implemented** **2= fully implemented** (same as above)		
S 12				24. Trigger4: free at home or feeling bored - Offer strategies to manage trigger 4	**0= not implemented** **1= partially implemented** **2= fully implemented** (same as above)		
S 13		BS1 & 3 ; BS11: Advise on avoiding cues for smoking BS12: Facilitate restructuring of social life	In addition to BS1 & 3 ; ■ Give specific advice on how to avoid being exposed to social or other cues for smoking (e.g. staying away from places where people smoke) ■ Advise on ways of changing social interactions with family, friends, and colleagues so that they support, rather than interfere with, the goal of remaining abstinent	25. Trigger5: seeing others smoke - Offer strategies to manage trigger 5	**0= not implemented** **1= partially implemented** **2= fully implemented** (same as above)		
S 14				26. Trigger6: offered smoking by others - Offer strategies to manage trigger 6	**0= not implemented** **1= partially implemented** **2= fully implemented** (same as above)		
S 15		BS1 & 3; BS10: Advise on conserving mental resources	In addition to BS1 & 3; Advise on ways of minimizing stress and other demands on mental resources (activities that require mental effort)	27. Trigger7: intense physical/mental work - Offer strategies to manage trigger 7	**0= not implemented** **1= partially implemented** **2= fully implemented** (same as above)		
S 16		BS1, 3 & 10; BS14: Teach relaxation techniques	In addition to BS1, 3 & 10; Teach specific relaxation techniques and how and when to apply them	28. Trigger8: tense/anxious - Offer strategies to manage trigger 8	**0= not implemented** **1= partially implemented** **2= fully implemented** (same as above)		
S 18				29. Withdrawal1: Craving smoking - Offer strategies to manage withdrawal1	**0= not implemented** **1= partially implemented** **2= fully implemented** (same as above)		

S19	General aspects of the interaction (R) focusing on general communication(C)	RC10: Provide information on withdrawal symptoms	Describe to the smoker what are, and are not, nicotine withdrawal symptoms, how common they are, how long they typically last, what causes them, and what can be done to alleviate them	30. Withdrawal2: Restlessness/anger - Offer strategies to manage withdrawal2	**0= not implemented** **1= partially implemented** **2= fully implemented** (same as above)		
S 20				31. Withdrawal3: Headache - Offer strategies to manage Withdrawal3	**0= not implemented** **1= partially implemented** **2= fully implemented** (same as above)		
S 21				32. Withdrawal4: Insomnia - Offer strategies to manage withdrawal4	**0= not implemented** **1= partially implemented** **2= fully implemented** (same as above)		
S 22				33. Withdrawal5: indigestion - Offer strategies to manage withdrawal5	**0= not implemented** **1= partially implemented** **2= fully implemented** (same as above)		
Q-C4. S23				34. Withdrawal6: Anorexia and constipation - Offer strategies to manage withdrawal6	**0= not implemented** **1= partially implemented** **2= fully implemented** (same as above)		
S 24				35. Withdrawal7: Cough - Offer strategies to manage withdrawal7	**0= not implemented** **1= partially implemented** **2= fully implemented** (same as above)		
S 25	General aspects of the interaction (R) focusing on general communication(C)& Specific focus on behaviour (B) and maximising self-regulatory capacity/skills (S)	RC10; BS13: Advise on methods of weight control	In addition to RC10; Advise on ways of minimising weight gain that do not increase motivation to smoke (e.g. take exercise, carry around "healthy snacks")	36. Withdrawal8: Weight gain - Offer strategies to manage withdrawal8	**0= not implemented** **1= partially implemented** **2= fully implemented** (same as above)		
Step-5: ARRANGE (follow-up)							
Section 4D.	General aspects of the interaction (R) focusing on general communication(C)	RC6: Offer appropriate written material	Distinguish what are, and are not, appropriate written materials and offer/direct clients to these in ways that promote their effective use	37. Give BSS information leaflet to the patient and brief about it.	**0= not implemented** **2= implemented**		

Total coding score (**add the score for each item and enter the composite score for Adherence**)						0	
QUALITY							
	Behavioural Determinants	**BCT code: label**	**BCT description/ Operational definition**	**Items** (ingredients of BI)	**Anchored scale**	**Score**	**Comment**
	General aspects of the interaction (R) focusing on general communication(C)	RC1: Build general rapport	Establish a positive , friendly and professional relationship with the smoker and foster a sense that the smoker's experiences are understood	38. • The provider introduces him/herself to the patient and takes time to ease the patient by asking about their day (e.g. how was their day, health etc.; how was the travel to the centre; how are they doing) and appreciating their participation (e.g. by thanking them for taking time out and how useful their participation is). • Pays complete attention to the patient when s/he is talking without cutting their talk short. ■ Listens to the patient after asking a question and gives them enough time to answer back. ■ If the patient drifts from the topic the provider politely reminds the patient to focus on the topic without disregarding their talk. • Gives them feedback, prompts or probes, taking part in the discussion (e.g. encouraging the patient, verbally responding to patient's concerns at the right time, showing concern for patient's issues and negative experiences relating to smoking/cessation).	**0= not implemented** (does not make any effort to build rapport with the patient) **1= partially implemented** **2= fully implemented** (makes all of the described examples of efforts to build rapport with the patient)		
		RC9: Explain expectations regarding treatment programme	Explain to the smoker the treatment program, what it involves, the active ingredients, and what it requires of the smoker	39. ■ Does the provider remind the patient of the information provided earlier about the intervention (before consent)? Explaining to the patient that the treatment program consists of five steps, Ask, Advise, Assess Assist, and Arrange/ the management to help you change your smoking habit. ■ Before each step does he take a moment to give a brief introduction to what is going to happen now? Taking the patient along and making an effort that the key intervention messages are being understood as intended.	**0= not implemented** **1= partially implemented** **2= fully implemented**		

		RC4: Provide reassurance	Give general reassurance to the smoker that his/her experiences are normal and time limited, and provide positive expectations of success based on experience with other smokers in the same situation	40. If the patient has experienced unsuccessful quit attempts or other negative experiences relating to smoking in the past, the provider: ■ Tells them that it's ok ■ Reassures them that this is normal when individuals try to quit on their own ■ Encourages them to be persistent and carry on with their attempt as this is time limited ■ Reassures them that with assistance through smoking cessation programme they are more likely to have a successful outcome ■ Provides examples from other smokers' experiences who were successful in quitting	**0= not implemented** (does not provide any reassurance) **1= partially implemented** (listens and answers in 'Yes' and 'No' but does not give any constructive advice) **2= fully implemented** (gives constructive advice)		
	General aspects of the interaction (R) focusing on general communication(C)	RC2: General practitioner communication approaches	Elicit and answer questions	41. ■ Is the provider supporting his questions with clear prompts to get the right information out of the patient e.g. when he asks, "What is your monthly household income?" If the patient says 10000, the provider confirms from him if this is his cumulative household income of all working individuals at home or his alone. Then he clarifies with the patient that he needs to tell about the cumulative household income. ■ Does he prompt the patient to give views on smoking, smoking cessation and any aspects of this behavioural support programme e.g. asking the patient, "Do you feel confident about this programme that it will help you in quitting smoking".	**0= not implemented** **1= partially implemented** **2= fully implemented**		
			Use reflective listening	42. Adopt a style of interaction that involves listening carefully to the smoker and where appropriate reflecting back to the smoker key elements of what s/he is saying e.g. the patient says, "When I don't smoke I have a very heavy feeling after meals". So the provider should respond by saying, "You mean to say that when you attempt quitting smoking you get indigestion".	**0= not implemented** **1= partially implemented** **2= fully implemented**		
			Summarising information and confirm client decisions	43. Provide a summary of information exchanged and establish a clear confirmation of decisions made and commitments entered into at each step and also at the end e.g. the provider reconfirms the quit date with the patient and arranges the follow up date reaffirming the milestones agreed upon during intervention session.	**0= not implemented** **1= partially implemented** **2= fully implemented**		

	General aspects of the interaction (R) focusing on the delivery of the intervention (D)	RC5: Tailor interactions appropriately	Use relevant information from the client to tailor the behavioural support provided/ flexible adaptation that takes into account individual patient needs	44. Where appropriate and needed, the provider, tailors the BSS according to the patients' needs e.g. in original trial some DOTS facilitators recommended the use of "cardamom" for chewing in place of "chewing gum", due to limited availability of chewing gums in those areas.	**0= not implemented** **1= partially implemented** **2= fully implemented**		
		RC3: Emphasise choice	Emphasise client choice within the bounds of evidence based practice	45. While delivering the management strategies for dealing with various triggers and withdrawals the provider emphasises patient's choice in the type of strategy that best suits their needs rather than imposing the strategy s/he think best. If the patient is not particularly responsive then the provider emphasises each management option available and prompts the patient to choose rather than just reading out the options and deciding for the patient him/herself.	**0= not implemented** **1= partially implemented** **2= fully implemented**		
Total coding score (add the score for each item and enter the composite score for Quality*)**						0	

Key

1. Q#; Slide#: These are for reference to the BI flip-chart and patient questionnaire (copies attached).

2. Behavioural Determinants: presents the underlying behavioural domain for the BCTs

3. And 4. BCT code and label and description: This is the BCT's code. Each BCT has a code in the taxonomy. The BCT labels and definitions from the taxonomy have been placed in the coding framework for reference purposes.

5. Item: This is the content of the Behavioural Intervention, in the form/wording that it was delivered in. Each item in its row has been linked to the columns 1, 2, 3 and 4.

6. And 7. Scoring: Description of each scoring criteria is provided in the scale. If the given BCT has never been used, please circle '0'. If the given BCT has partially been implemented, please circle '1' and if fully implemented then circle '2'.

8. Comments: Please list any necessary comments or points of clarification

Appendix D. APPENDICES TO CHAPTER 7

D.1: Provider information sheet and consent form

Please read this document carefully.

We would like to invite you to take part in a research study. Before you decide whether to take part it is important for you to understand why the research is being done and what it will involve. Please take time to read the following information carefully. Please ask us if there is anything that is not clear. If you decide you would like to take part please return the signed **Paramedic Consent Form** to the Research officer.

You are being asked to participate in an important follow on study of the ASSIST trial conducted between 2010- 2011in Sargodha and Jhang districts. The study will explore the various aspects of DOTS paramedics' characteristics and competences that might contribute to the variation observed in smoking cessation outcomes in patients between these centres. Composite data will be used to validate a fidelity criterion for future guidance of implementation of behavioural support interventions for smoking cessation in similar settings.

Keep the information sheet for your records. After receiving the completed consent form, you will be asked to deliver the interactive flip-book intervention session (that was previously delivered by you as part of the smoking cessation intervention in the ASSIST trial) to the TB suspects presenting to the TB clinic and this session will be video recorded. The intervention session is expected to last 30 to 40 minutes in total; two to three complete sessions will be recorded. This will be followed by an interview questionnaire (lasting 15- 20 minutes) that will be filled by the research officer to record some of your characteristics and competences.

You are free to choose whether or not to participate in this study. The taping of sessions is purely to record the intervention process, you are not forced or obliged to allow video-taping of any session. You can withdraw consent at any time, including during the session and after the tape has been completed, in which case the video recording for your session will be erased by the lead researcher. The interview is also completely voluntary; you may choose not to participate or not to answer any specific question at any point during the interview.

There is no known possible risk or discomfort that you can encounter connected to the activities in this study. The research through this study will benefit the wider research evidence base by providing understanding of the factors that are important for an effective and standardised implementation of a complex smoking cessation programme in the settings of primary and secondary health care centres.

Your identity and location in this study will remain strictly confidential. The results of the study, including data, may be published for scientific purposes but will not give your name or include any identifiable references to you.

Your contact information will be kept confidential and will only be used to contact you regarding the study.

If you have any questions or would like further information, please contact:
Research student Omara Dogar, E: ofd500@york.ac.uk

Department of Health Sciences

The University of York

Thank you

Health centre _____

Study Title: Can Fidelity to a Behavioural Intervention explain outcomes in Smoking Cessation?

Provider CONSENT FORM

To be completed by the TB paramedic

Please tick (√) the boxes if you agree with the following statements.

1. I confirm that I have read and understood the Information Sheet for Paramedics for the above study and have had the opportunity to ask questions. ☐
2. I understand that my participation is voluntary and that I am free to withdraw at any time, without giving any reason. ☐
3. I understand that my delivery of the intervention session will be video/audio taped followed by a questionnaire interview and that my identity and location will be kept confidential. ☐
4. I understand that the data from the tapes and interview will be accessed by the researchers. I give permission for these individuals to have access to my records. ☐
5. I understand that my name will not be linked with the research materials, and I will not be identified or identifiable in any reports that result from the research. ☐
6. **I agree to take part in the above study.** ☐

Your Name (please print) _____

Your Signature _____ Date _____

Address _____ Telephone number _____
(Including dialing code)

_____ Mobile number _____

_____

Signature of person obtaining consent _____ Date _____

Your contact information will be kept confidential and will only be used to contact you regarding the study.

If you have any questions or would like further information, please contact:

Research student Omara Dogar, E: ofd500@york.ac.uk

Department of Health Sciences
The University of York

Thank you

D.2: Patient information sheet and consent form

Please read this document carefully.

We would like to invite you to take part in a research study. Before you decide whether to take part it is important for you to understand why the research is being done and what it will involve. Please take time to read the following information carefully. Please ask us if there is anything that is not clear. If you decide you would like to take part please return the signed **Patient Consent Form** to the TB DOTS facilitator.

You are being asked to participate in a research study designed to explore various factors related to a smoking cessation program delivered in the TB clinic of the health centres to better inform the utilisation of this program in these and similar settings in other countries.

Keep the information sheet for your records. After receiving the completed consent form, you will be asked to participate in an interactive flip-book delivered intervention session to assist you with stopping smoking and this session will be video recorded. The intervention session will be delivered by the DOTS TB paramedic at your health centre. The intervention session is expected to last 30 to 40 minutes in total.

You are free to choose whether or not to participate in this study. The taping of sessions is purely to record the intervention process, you are not forced or obliged to allow video-taping of any session. You can withdraw consent at any time, including during the session and after the tape has been completed, in which case the video recording for your session will be erased by the lead researcher.

There is no known possible risk or discomfort that you can encounter connected to the activities in this study. The research through this study will benefit the wider research evidence base by providing understanding of the factors that are important for an effective and standardised implementation of a complex smoking cessation programme in the settings of primary and secondary health care centres.

Your identity and location in this study will remain strictly confidential. The results of the study, including data, may be published for scientific purposes but will not give your name or include any identifiable references to you.

Your contact information will be kept confidential and will only be used to contact you regarding the study.

If you have any questions or would like further information, please contact:

Research student Omara Dogar, E: ofd500@york.ac.uk

Department of Health Sciences

The University of York

Thank you

Health centre _____

Study Title: Can Fidelity to a Behavioural Intervention explain outcomes in Smoking Cessation?

Patient CONSENT FORM

To be completed by the TB paramedic

Please tick (√) the boxes if you agree with the following statements.

1. I confirm that I have read and understood the Information Sheet for Participants for the above study and have had the opportunity to ask questions. ☐
2. I understand that my participation is voluntary and that I am free to withdraw at any time, without giving any reason. ☐
3. I understand that my therapy/intervention session will be video/audio taped and that my identity and location will be kept confidential. ☐
4. I understand that the tapes will be accessed by the researchers. I give permission for these individuals to have access to my records. ☐
5. I understand that my name will not be linked with the research materials, and I will not be identified or identifiable in any reports that result from the research. ☐
6. **I agree to take part in the above study.** ☐

Name (please print) _____

Participant signature/thumb impression _____ Date _____

Address _____ Telephone number _____
(Including dialing code)

_____

Mobile number _____

_____

Signature of person obtaining consent _____ Date _____

Your contact information will be kept confidential and will only be used to contact you regarding the study.

If you have any questions or would like further information, please contact:

Research student Omara Dogar, E: ofd500@york.ac.uk

Department of Health Sciences
The University of York

Thank you

D.3: Patient assessment questionnaire (ASSIST study)

Five Steps to Quit Smoking

PATIENT REGISTRATION FORM

Name of Health Facility _____ District _____

Date _____ Reg. No. _____

(Step-1: ASK)

A. Basic Information

1. Name _____

2. Age (years) ____ 3. Sex (M/F) ____ 4. Phone Line: Mobile:

5. Occupation | Office worker | labourer | businessman | Farmer | Others

6. What is your monthly household income _____ PKR/month

7. Address _____

8. Does anyone else smoke tobacco in your household?	Yes	No
9. Does anyone at your work place smoke tobacco?	Yes	No
10. Are there any children younger than 12 years of age in your household?	Yes	No
11. If yes: How many children are younger than 12 years in your household?	Count: _____	

B. Use of tobacco product

1. What form of tobacco product you usually smoke?

Cigarette	Huqqa	Pipe	Cigar	Others:

2. At what age did you start smoking tobacco? Years:

3. Duration of smoking (Age in years – age started smoking) Years: Months:

4a. How many cigarettes do you smoke in a day? Count:

5a. (If huqqa smoker), How many times in a day do you smoke huqqa? Count:

Step-2: ADVISE (Risks of tobacco use and benefits of cessation)

Use Flip-Book slides 1 to 7

Step-3: ASSESS WILLINGNESS TO QUIT

A. Willingness to Quit

1. Are you willing to quit smoking?	Yes	No
2. If yes: Can you quit from today?	Yes	No
3. If No: If not now can you quit within five days?	Yes	No
4. If no: When will you be able to set a date to quit?	Weeks:*	

*Continue only with those who set a quit date within 5 days of initial visit. Ask others to come back when they are prepared to set a quit date.

B. Why do you want to quit?

1. Health	2. Cost	3. Role Model	4. Family Health	5. Other:

C. Previous Attempt/s to quit

1. Did you ever attempt to quit?	Yes	No
2. If yes: How many times?	Number	
3. When did you attempt last time	Months:	
4. What was the longest duration in previous attempts?	Months:	
5. Why did you start again after that attempt? (Tick the relevant reason)	**Social reasons**: Family problems (tension), Smokers company, any other (specify) _____ **Physical Symptoms**: Craving, Indigestion, Insomnia, Headache, any other (specify) _____	

Step-4: ASSIST IN QUITTING

A. Assess nicotine dependence (Fagerstrom Test for Nicotine Dependence)

Sr. No	Questions	Codes	Score	Patient Score
1.	How soon after you wake do you smoke your first cigarette/Huqqa?	With in 5 minutes	3	
		5-30 minutes	2	
		31-60 minutes	1	
		After 60 minutes	0	
2.A	How many cigarettes per day do you usually smoke?	10 and Less	0	
		11-20	1	
		21-30	2	
		31 or more	3	
2.B	How many times per day you	5 times	0	

	smoke Huqqa	6-10	1	
		11-15	2	
		16 or more	3	
3.	Do you find it difficult to refrain from smoking in places where it is forbidden?	No	0	
		Yes	1	
4.	Which cigarette/Huqqa would you hate most to give up?	First one in the morning	1	
		Any other	0	
5.	Do you smoke more frequently during the first hours after awakening than during the rest of the day?	No	0	
		Yes	1	
6.	Do you smoke even if you are so ill that you are in bed most of the day?	No	0	
		Yes	1	

Total patient Score	
Patient Dependence Level*	

*Nicotine dependence level scale

Score	0 - 2	3 - 4	5	6 - 7	8 - 10
Dependence Level	Very low	Low	Medium	High	Very High

B. ~~Perform CO Test~~

~~1. Perform the test using CO Testing Chart (Guidelines)~~
~~2. Record the CO reading in Step 5-B.~~

C. Agree on a quit plan

C-1: Quit Date
Write down the identified Quit Date _____

C-2: General Guidance
Use Slide-8 on Desk-guide to provide general guidance on quitting.

C-3: Plan to manage triggers
Use desk-guide slides 9 to 16 and record in the following table:

Triggers	Strategies to manage		
Rising in the morning (immediately after)	Take some drink (Juice, green tea, lassi) after you get out of bed	Brush your teeth	Go for a walk
Defecation	Take some laxatives (Isbagol Husk) at night	Take newspaper to toilet instead of cigarette	

Eating Meals	Take a chewing gum	Go for a walk after meals	Take a nap after lunch
Relaxing/boredom at Home	Take some dry fruit/grams	Do gardening	Have tea, Watch TV
Seeing others smoking	Take a chewing gum	Advocate to the friends about the smoking harms	Avoid smoker friends in the initial period
Offered smoking	Refuse it straight away	Take chewing gum	Relax and take deep breathes
Intensive Physical/ Mental work	Take dry fruit	Relax and take deep breathes	Take small short breaks and have tea
Tension/Anxiety	Chat with colleagues	Relaxing Exercises	Offer Nimaz

E. Give the 'leaflet' to patient and brief him about it

~~**D. Dispense drug and explain how to take it (Bupropion (BSS+) Arm Only)**~~
~~Use slide-17 from desk-guide on how to dispense the drug and note down the amount given to the patient in the following table.~~

~~Date~~	~~# of Zylexx^{SR} tablets dispensed~~

~~**A. Review general guidance tasks and support**~~

~~Compliance to quit date~~	~~Y/N~~
~~Hiding of reminders of smoking e.g. tobacco, cigarettes, huqqa, ashtrays, matches etc~~	~~Y/N~~
~~Declaring of house a smoke free home~~	~~Y/N~~
~~Identification of individual who can help support in quitting at home~~	~~Y/N~~
~~Decision on telling people about stopping or keeping it private~~	~~Y/N~~

~~**C. Review of triggers management**~~
~~Use triggers table of the form to review and guide~~

D. Discuss any withdrawals and plan to manage them
Use Desk-guide slides 18 to 22 and record in the following table

Withdrawals	Strategies to manage		
Craving/Desire for cigarettes	Take a glass of juice/ cold drink	Engage in conversation with a friend	Take deep breadths and slowly breathe out

Restlessness Irritability/ frustration/ Anger	Take a glass of juice/ cold drink	Engage in conversation with a friend –	Take deep breadths and slowly breathe out
Cough	Take some throat soothers	Make a habit of morning walk	Consult doctor if cough is troublesome
Weight Gain	Take more fruits and vegetables in your diet	Exercise daily	Make a habit of morning walk
Headache	– Take a pill for headache	Engage in conversation with a friend	Offer nimaz
Insomnia	Exercise daily	Offer nimaz	Consult doctor if not getting better
Anorexia & Constipation	Take plenty of fluids/vegetables/fruits	Make a habit of morning walk	If disturbing consult your doctor
Indigestion/ Heart burns	Take more fruits and vegetables.	Make a habit of morning walk	Consult doctor if not getting better

F. Tell the patient to visit health centre for follow-up.
Next Follow-up Date _____ (Record Date)

Step-5: FOLLOW-UP

B. CO Test results

Follow-up 6 month

H. If the patient is still smoking tobacco then ask;

1. How many cigarettes do you smoke in a day?	Count:
2. (If huqqa smoker), How many times in a day do you smoke huqqa?	Count:

E. ~~Ask about compliance to taking of medicine and any side effects~~

~~1. Did you take medicine regularly?~~	~~Yes~~	~~No~~
~~2. If No: What was the reason not to take it regularly?~~		
~~3. Ask about any physical complaints and take action~~		

~~Side-effects~~	~~Action~~
~~Minor Side-effects~~ ~~– Dry mouth / unusual taste~~ ~~– Insomnia~~ ~~– Headache~~ ~~– 'Spaced out' feeling~~ ~~– Nausea~~ ~~– Constipation~~ ~~– Tremor~~	~~Tick mark on the complaint and reassure the patient, they will disappear in few days~~
~~Major Side-effects~~ ~~– Seizures (1/1000)~~ ~~– Elevated blood pressure~~ ~~– Allergic reaction - discontinue use if rash / symptoms~~ ~~– Confusion~~ ~~– Reduced appetite~~ ~~– Tinnitus, visual disturbance~~	~~Tick mark on the complaint and send the patient to the doctor after completing the form~~

~~F. Dispense remaining amount of drugs and explain how to take it~~ ~~(Bupropion Arm)~~
~~Use Slide-17 of Desk-guide and record the amount in section 'D' of Step-4.~~

~~G. Outcome~~

~~Adherence to quit~~	~~Follow up 1 (1 week)~~		~~Follow up 2 (4 weeks)~~		~~Follow up 3 (24 weeks)~~	
~~Date~~						
~~Complete Abstinence~~	~~Yes~~	~~No~~	~~Yes~~	~~No~~	~~Yes~~	~~No~~

LIST OF KEYWORDS

Behavioural interventions

Complex interventions

Smoking cessation

Variation in quit rates

Tuberculosis

Tuberculosis clinics

Tuberculosis programme

Intervention fidelity

Fidelity index

Adherence to behavioural intervention content

Quality of interaction

Patient-provider interaction sessions

Process evaluation

Interpretive evaluation

Psychometric properties

Provider competence

Provider practice behaviour

Low and Middle Income Countries

LIST OF ABBREVIATIONS

AM	Alveolar Macrophages
ASSIST	Action to Stop Smoking In Suspected TB
BECCI	Behaviour Change Counselling Index
BI	Behavioural Intervention
BSS	Behavioural Support Sessions
CI	Confidence Interval
CONSORT	Consolidated Standards of Reporting Trials
CO	Carbon Monoxide
COM-B	Capability, Opportunity, Motivation- Behaviour
DOTS	Directly Observed Treatment Short-course
GRADE	Grading of Recommendations, Assessment, Development, and Evaluation
HIC	High-Income Countries
HSRGC	Health Sciences Research Governance Committee
ICC	Intra-cluster Correlation Co-efficient
IGRA	Interferon Gamma Release Assay
INF-gamma	Interferon gamma
KALPHA	Kripendorff's alpha
LMIC	Low-and-Middle Income Countries
MRC	Medical Research Council
NRT	Nicotine Replacement Therapy
NTP	National Tuberculosis Programme
OR	Odds Ratio
PCA	Principal Components Analysis
PHC	Primary Health Care
PMRC	Pakistan Medical Research Council
RCT	Randomised Controlled Trial
RE-AIM	Reach Effectiveness Adoption Implementation Maintenance
SHS	Second-Hand Smoking
SPIRIT	Standard Protocol Items: Recommendations for Interventional Trials
TB	Tuberculosis
TDF	Theoretical Domains Framework
TIDieR	Template for Intervention Description and Replication
TNF	Tumour Necrosis Factor
TST	Tuberculin Skin Test

REFERENCES

ABRAHAM, C. & MICHIE, S. 2008. A taxonomy of behavior change techniques used in interventions. *Health psychology,* 27**,** 379.

ADAMS, G., GULLIFORD, M. C., UKOUMUNNE, O. C., ELDRIDGE, S., CHINN, S. & CAMPBELL, M. J. 2004. Patterns of intra-cluster correlation from primary care research to inform study design and analysis. *Journal of Clinical Epidemiology,* 57**,** 785-794.

ADELSTEIN, A. & RIMINGTON, J. 1967. Smoking and pulmonary tuberculosis: an analysis based on a study of volunteers for mass miniature radiography. *Tubercle,* 48**,** 219-226.

ADIB, S. M., AL-TAKASH, H. & AL-HAJJ, C. 1999. Tuberculosis in Lebanese jails: prevalence and risk factors. *European journal of epidemiology,* 15**,** 253-260.

AJZEN, I. 1991. The theory of planned behavior. *Organizational behavior and human decision processes,* 50**,** 179-211.

AL-DARRAJI, H. A. A., TAN, C., KAMARULZAMAN, A. & ALTICE, F. L. 2015. Prevalence and correlates of latent tuberculosis infection among employees of a high security prison in Malaysia. *Occupational and environmental medicine,* 72**,** 442-447.

ALBRECHT, L., ARCHIBALD, M., ARSENEAU, D. & SCOTT, S. D. 2013. Development of a checklist to assess the quality of reporting of knowledge translation interventions using the Workgroup for Intervention Development and Evaluation Research (WIDER) recommendations. *Implement Sci,* 8**,** 52.

ALCAIDE, J., ALTET, M., PLANS, P., PARRON, I., FOLGUERA, L., SALTO, E., DOMINGUEZ, A., PARDELL, H. & SALLERAS, L. 1996. Cigarette smoking as a risk factor for tuberculosis in young adults: a casecontrol study. *Tubercle and Lung Disease,* 77**,** 112-116.

ALTET-GOMEZ, M., ALCAIDE, J., GODOY, P., ROMERO, M. & HERNANDEZ, R. I. 2005. Clinical and epidemiological aspects of smoking and tuberculosis: a study of 13038 cases. *The International Journal of Tuberculosis and Lung Disease,* 9**,** 430-436.

ANDERSON, R. H., SY, F. S., THOMPSON, S. & ADDY, C. 1997. Cigarette smoking and tuberculin skin test conversion among incarcerated adults. *American journal of preventive medicine,* 13**,** 175.

ARCAVI, L. & BENOWITZ, N. L. 2004. Cigarette smoking and infection. *Archives of Internal Medicine,* 164**,** 2206.

ARIYOTHAI, N., PODHIPAK, A., AKARASEWI, P., TORNEE, S., SMITHTIKARN, S. & THONGPRATHUM, P. 2004. Cigarette smoking and its relation to pulmonary tuberculosis in adults.

AWAISU, A., NIK MOHAMED, M. H., MOHAMAD NOORDIN, N., ABD AZIZ, N., SYED SULAIMAN, S. A., MUTTALIF, A. R. & AHMAD MAHAYIDDIN, A. 2011. The SCIDOTS Project: evidence of benefits of an integrated tobacco cessation intervention in tuberculosis care on treatment outcomes. *Substance abuse treatment, prevention, and policy* [Online]. Available: http://onlinelibrary.wiley.com/o/cochrane/clcentral/articles/668/CN-00805668/frame.html.

BAKER, T. B., COLLINS, L. M., MERMELSTEIN, R., PIPER, M. E., SCHLAM, T. R., COOK, J. W., BOLT, D. M., SMITH, S. S., JORENBY, D. E. & FRASER, D. 2016. Enhancing the effectiveness of smoking treatment research: conceptual bases and progress. *Addiction,* 111**,** 107-116.

BANDURA, A. 1997. *Self-efficacy: The exercise of control*, New York: Freeman.

BARANOWSKI, T. & STABLES, G. 2000. Process evaluations of the 5-a-day projects. *Health Education & Behavior,* 27**,** 157-166.

BASU, S., STUCKLER, D., BITTON, A. & GLANTZ, S. A. 2011. Projected effects of tobacco smoking on worldwide tuberculosis control: mathematical modelling analysis. *British Medical Journal,* 343.

BATES, M. N., KHALAKDINA, A., PAI, M., CHANG, L., LESSA, F. & SMITH, K. R. 2007. Risk of tuberculosis from exposure to tobacco smoke: a systematic review and meta-analysis. *Archives of Internal Medicine,* 167**,** 335.

BAUMEISTER, R. F., HEATHERTON, T. F. & TICE, D. M. 1994. *Losing control: How and why people fail at self-regulation*, Academic press.

BELL, R. A. & KRAVITZ, R. L. 2014. DIRECT OBSERVATION AND CODING OF PHYSICIAN–PATIENT INTERACTIONS. *Research Methods in Health Communication: Principles and Application* [Online].

BELLG, A. J., BORRELLI, B., RESNICK, B., HECHT, J., MINICUCCI, D. S., ORY, M., OGEDEGBE, G., ORWIG, D., ERNST, D. & CZAJKOWSKI, S. 2004. Enhancing treatment fidelity in health behavior change studies: best practices and recommendations from the NIH Behavior Change Consortium. *Health Psychology-Hillsdale-,* 23**,** 443-451.

BERGSTRÖM, A., SKEEN, S., DUC, D. M., BLANDON, E. Z., ESTABROOKS, C., GUSTAVSSON, P., HOA, D. T. P., KÄLLESTÅL, C., MÅLQVIST, M. & NGA, N. T. 2015. Health system context and implementation of evidence-based practices—development and validation of the Context Assessment for Community Health (COACH) tool for low-and middle-income settings. *Implementation Science,* 10**,** 1.

BLACK, N., MURPHY, M., LAMPING, D., MCKEE, M., SANDERSON, C., ASKHAM, J. & MARTEAU, T. 1999. Consensus development methods: a review of best practice in creating clinical guidelines. *Journal of health services research & policy,* 4**,** 236-248.

BLAND, J. M. & ALTMAN, D. 1986. Statistical methods for assessing agreement between two methods of clinical measurement. *The lancet,* 327**,** 307-310.

BLAND, J. M. & ALTMAN, D. G. 1999. Measuring agreement in method comparison studies. *Statistical methods in medical research,* 8**,** 135-160.

BOELAERT, J. R., GOMES, M. S. & GORDEUK, V. R. 2003. Smoking, iron, and tuberculosis. *The Lancet,* 362**,** 1243-1244.

BOND, G., BECKER, D., DRAKE, R. & VOGLER, K. 1997a. A fidelity scale for the Individual Placement and Support model of supported employment. *Rehabilitation Counseling Bulletin,* 40**,** 265-284.

BOND, G., PICONE, J., MAUER, E., FISHBEIN, S. & STOUT, R. 2000a. The Quality of Supported Employment Implementation Scale. *Journal of Vocational Rehabilitation,* 14**,** 201-212.

BOND, G. R., BECKER, D. R., DRAKE, R. E. & VOGLER, K. M. 1997b. A fidelity scale for the Individual Placement and Support model of supported employment. *Rehabilitation Counseling Bulletin*.

BOND, G. R., WILLIAMS, J., EVANS, L., SALYERS, M., KIM, H. & SHARPE, H. 2000b. PN-44-psychiatric rehabilitation fidelity toolkit. Cambridge, MA: Human Services Research Institute

BONELL, C., OAKLEY, A., HARGREAVES, J., STRANGE, V. & REES, R. 2006. Research methodology: Assessment of generalisability in trials of health interventions: suggested framework and systematic review. *BMJ: British Medical Journal,* 333**,** 346.

BORRELLI, B. 2011. The assessment, monitoring, and enhancement of treatment fidelity in public health clinical trials. *Journal of public health dentistry,* 71**,** S52-S63.

BORRELLI, B., SEPINWALL, D., ERNST, D., BELLG, A. J., CZAJKOWSKI, S., BREGER, R., DEFRANCESCO, C., LEVESQUE, C., SHARP, D. L. & OGEDEGBE, G. 2005. A new tool to assess treatment fidelity and evaluation of treatment fidelity across 10 years of health behavior research. *Journal of consulting and clinical psychology,* 73**,** 852.

BOSLAUGH, S. 2007. *Secondary data sources for public health: A practical guide*, Cambridge University Press.

BOUTRON, I., MOHER, D., ALTMAN, D. G., SCHULZ, K. F. & RAVAUD, P. 2008. Extending the CONSORT statement to randomized trials of nonpharmacologic treatment: explanation and elaboration. *Annals of internal medicine,* 148**,** 295-309.

BREHM, J. W. & SELF, E. A. 1989. The intensity of motivation.

BREHM, J. W., WRIGHT, R. A., SOLOMON, S., SILKA, L. & GREENBERG, J. 1983. Perceived difficulty, energization, and the magnitude of goal valence. *Journal of Experimental Social Psychology,* 19**,** 21-48.

BRENNAN, R. L. 2001. *Generalizability Theory*, Springer.

BROSE, L. S., MCEWEN, A. & WEST, R. 2012. Does it matter who you see to help you stop smoking? Short-term quit rates across specialist stop smoking practitioners in England. *Addiction,* 107**,** 2029-2036.

BROWN, K. & CAMPBELL, A. 1961. Tobacco, alcohol and tuberculosis. *British Journal of Diseases of the Chest,* 55**,** 150-158.

BROWNSON, R. C., COLDITZ, G. A. & PROCTOR, E. K. 2012. *Dissemination and implementation research in health: translating science to practice*, Oxford University Press.

BRUNET, L., PAI, M., DAVIDS, V., LING, D., PARADIS, G., LENDERS, L., MELDAU, R., VAN ZYL SMIT, R., CALLIGARO, G., ALLWOOD, B., DAWSON, R. & DHEDA, K. 2011. High prevalence of smoking among patients with suspected tuberculosis in South Africa. *European Respiratory Journal,* 38**,** 139-46.

BUMBARGER, B. & PERKINS, D. 2008. After randomised trials: issues related to dissemination of evidence-based interventions. *Journal of Children's Services,* 3**,** 55-64.

BUSKIN, S. E., GALE, J. L., WEISS, N. S. & NOLAN, C. M. 1994. Tuberculosis risk factors in adults in King County, Washington, 1988 through 1990. *American Journal of Public Health,* 84**,** 1750-1756.

BYGBJERG, I. 2012. Double Burden of Noncommunicable and Infectious Diseases in Developing Countries. *Science,* 337**,** 1499-1501.

CAMPBELL, M., DONNER, A. & KLAR, N. 2007a. Developments in cluster randomized trials and Statistics in Medicine. *Statistics in medicine,* 26**,** 2-19.

CAMPBELL, M., FITZPATRICK, R., HAINES, A., KINMONTH, A. L., SANDERCOCK, P., SPIEGELHALTER, D. & TYRER, P. 2000. Framework for design and evaluation of complex interventions to improve health. *Bmj,* 321**,** 694-696.

CAMPBELL, M. K., FAYERS, P. M. & GRIMSHAW, J. M. 2005. Determinants of the intracluster correlation coefficient in cluster randomized trials: the case of implementation research. *Clinical Trials,* 2**,** 99-107.

CAMPBELL, N. C., MURRAY, E., DARBYSHIRE, J., EMERY, J., FARMER, A., GRIFFITHS, F., GUTHRIE, B., LESTER, H., WILSON, P. & KINMONTH, A. L. 2007b. Designing and evaluating complex interventions to improve health care. *Bmj,* 334**,** 455-459.

CANE, J., O'CONNOR, D. & MICHIE, S. 2012. Validation of the theoretical domains framework for use in behaviour change and implementation research. *Implement Sci,* 7**,** 37.

CARROLL, C., PATTERSON, M., WOOD, S., BOOTH, A., RICK, J. & BALAIN, S. 2007. A conceptual framework for implementation fidelity. *Implement Sci,* 2**,** 40.

CARROLL, K. M., NICH, C., SIFRY, R. L., NURO, K. F., FRANKFORTER, T. L., BALL, S. A., FENTON, L. & ROUNSAVILLE, B. J. 2000. A general system for evaluating therapist adherence and competence in psychotherapy research in the addictions. *Drug and alcohol dependence,* 57**,** 225-238.

CELLA, D., GERSHON, R., LAI, J.-S. & CHOI, S. 2007. The future of outcomes measurement: item banking, tailored short-forms, and computerized adaptive assessment. *Quality of Life Research,* 16**,** 133-141.

CENTERS FOR DISEASE CONTROL. 2014. *Fact sheets: The Difference Between Latent TB Infection and TB Disease* [Online]. Available: http://www.cdc.gov/tb/publications/factsheets/general/ltbiandactivetb.htm [Accessed August 3, 2015].

CHENG, H. G. & PHILLIPS, M. R. 2014. Secondary analysis of existing data: opportunities and implementation. *Shanghai Arch Psychiatry,* 26**,** 371-375.

CLARKE, G. 1998. Intervention fidelity in the psychosocial prevention and treatment of adolescent depression. *Journal of Prevention & Intervention in the Community,* 17**,** 19-33.

COOK, T. D., CAMPBELL, D. T. & DAY, A. 1979. *Quasi-experimentation: Design & analysis issues for field settings*, Houghton Mifflin Boston.

CRAIG, P., DIEPPE, P., MACINTYRE, S., MICHIE, S., NAZARETH, I. & PETTICREW, M. 2008. Developing and evaluating complex interventions: the new Medical Research Council guidance. *BMJ: British Medical Journal,* 337.

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**,** 587-592.

CRAMPIN, A., GLYNN, J., FLOYD, S., MALEMA, S., MWINUKA, V., NGWIRA, B., MWAUNGULU, F., WARNDORFF, D. & FINE, P. 2004. Tuberculosis and gender: exploring the patterns in a case control study in Malawi. *The International Journal of Tuberculosis and Lung Disease,* 8**,** 194-203.

CRAUN, G. F. & CALDERON, R. L. *How to interpret epidemiological associations* [Online]. U.S. Environmental Protection Agency, National Health and Environmental Effects Research Laboratory, Research Triangle Park, NC, USA Available: http://www.who.int/water_sanitation_health/dwq/nutrientschap9.pdf [Accessed September 9, 2015].

CURRAN-EVERETT, D. & MILGROM, H. 2013. Post-hoc data analysis: benefits and limitations. *Current opinion in allergy and clinical immunology,* 13**,** 223-224.

CURRAN, G. M., BAUER, M., MITTMAN, B., PYNE, J. M. & STETLER, C. 2012. Effectiveness-implementation hybrid designs: combining elements of clinical effectiveness and implementation research to enhance public health impact. *Medical care,* 50**,** 217.

D'ARC LYRA BATISTA, J., DE FATIMA PESSOA MILITAO DE ALBUQUERQUE, M., DE ALENCAR XIMENES, R. A. & RODRIGUES, L. C. 2008. Smoking increases the risk of relapse after successful tuberculosis treatment. *International Journal of Epidemiology,* 37**,** 841-51.

D O'HALLORAN, P., BLACKSTOCK, F., SHIELDS, N., HOLLAND, A., ILES, R., KINGSLEY, M., BERNHARDT, J., LANNIN, N., MORRIS, M. E. & TAYLOR, N. F. 2014. Motivational interviewing to increase physical activity in people with chronic health conditions: a systematic review and meta-analysis. *Clinical Rehabilitation***,** 0269215514536210.

DANE, A. V. & SCHNEIDER, B. H. 1998. Program integrity in primary and early secondary prevention: Are implementation effects out of control? *Clinical psychology review,* 18**,** 23-45.

DE WINTER, J., DODOU, D. & WIERINGA, P. 2009. Exploratory Factor Analysis With Small Sample Sizes. *Multivariate Behavioral Research,* 44**,** 147-181.

DEN BOON, S., VAN LILL, S., BORGDORFF, M., VERVER, S., BATEMAN, E., LOMBARD, C., ENARSON, D. & BEYERS, N. 2005. Association between smoking and tuberculosis infection: a population survey in a high tuberculosis incidence area. *Thorax,* 60**,** 555-557.

DEVELLIS, R. F. 2003. *Scale Development: Theory and Applications*, Sage.

DI STEFANO, A., VALLESE, D., PITRUZZELLA, A., ZANINI, A., SPANEVELLO, A. & BALBI, B. 2012. Airway inflammation in healthy smokers. *Monaldi Arch Chest Dis,* 77**,** 53-56.

DOGAR, O., BULLEN, C., NOVOTNY, T. E. & SIDDIQI, K. 2013. Smoking cessation and respiratory disease in low-income and middle-income countries. *The Lancet Respiratory Medicine,* 1**,** e23-e24.

DOGAR, O., ELSEY, H., KHANAL, S. & SIDDIQI, K. 2016. Challenges of integrating tobacco cessation interventions in TB programs: Case studies from Nepal and Pakistan. *Journal of Smoking Cessation*.

DOGAR, O., PILLAI, N., SAFDAR, N., SHAH, S., ZAHID, R. & SIDDIQI, K. 2015. Second-hand smoke and the risk of tuberculosis: a systematic review and a meta-analysis. *Epidemiology and infection***,** 1-15.

DOGAR, O. F. & SIDDIQI, K. 2013. An evidence-based guide to smoking cessation therapies. *Nurse Prescribing,* 11**,** 543-548.

DOLL, R. 1999. Risk from tobacco and potentials for health gain. *The International Journal of Tuberculosis and Lung Disease,* 3**,** 90-99.

DOLL, R., PETO, R., BOREHAM, J. & SUTHERLAND, I. 2004. Mortality in relation to smoking: 50 years' observations on male British doctors. *Bmj,* 328**,** 1519.

DONG, B., GE, N. & ZHOU, Y. 2001. [Smoking and alcohol consumption as risk factors of pulmonary tuberculosis in Chengdu: a matched case-control study]. *Hua xi yi ke da xue xue bao= Journal of West China University of Medical Sciences= Huaxi yike daxue xuebao/[bian ji zhe, Hua xi yi ke da xue xue bao bian wei hui],* 32**,** 104-106.

DONNER, A. & KLAR, N. 2000. *Design and Analysis of Cluster Randomization Trials in Health Research,* Arnold, London

DRAKE, R. E., GOLDMAN, H. H., LEFF, H. S., LEHMAN, A. F., DIXON, L., MUESER, K. T. & TORREY, W. C. 2001. Implementing evidence-based practices in routine mental health service settings. *Psychiatric services*.

DURLAK, J. A. & DUPRE, E. P. 2008. Implementation matters: A review of research on the influence of implementation on program outcomes and the factors affecting implementation. *American journal of community psychology,* 41**,** 327-350.

DUSENBURY, L., BRANNIGAN, R., FALCO, M. & HANSEN, W. B. 2003. A review of research on fidelity of implementation: implications for drug abuse prevention in school settings. *Health education research,* 18**,** 237-256.

DYE, C. 2006. Global epidemiology of tuberculosis. *Lancet,* 367**,** 938-40.

EAMES, C., DALEY, D., HUTCHINGS, J., HUGHES, J., JONES, K., MARTIN, P. & BYWATER, T. 2008. The Leader Observation Tool: A process skills treatment fidelity measure for the Incredible Years parenting programme. *Child: care, health and development,* 34**,** 391-400.

ECCLES, M., GRIMSHAW, J., CAMPBELL, M. & RAMSAY, C. 2003. Research designs for studies evaluating the effectiveness of change and improvement strategies. *Quality and Safety in Health Care,* 12**,** 47-52.

ENGEL, R. J. & SCHUTT, R. K. 2012. *The practice of research in social work*, Sage.

ESSON, M. & LEEDER, R. 2004. The Millennium development goals and tobacco control: an opportunity for global partnership. *World Health Organization*.

FARMER, A., HARDEMAN, W., HUGHES, D., PREVOST, A. T., KIM, Y., CRAVEN, A., OKE, J., BOASE, S., SELWOOD, M. & KELLAR, I. 2012. An explanatory randomised controlled trial of a nurse-led, consultation-based intervention to support patients with adherence to taking glucose lowering medication for type 2 diabetes. *BMC family practice,* 13**,** 1.

FIORE, M. C. 2000. US public health service clinical practice guideline: treating tobacco use and dependence. *Respiratory Care,* 45**,** 1200-1262.

FIRPO-TRIPLETT, R. & FULLER, T. R. 2012. General Adaptation Guidance; A Guide to Adapting Evidence-Based Sexual Health Curricula. Scotts Valley, CA: ETR Associates/CDC.

FISHBEIN, M. A theory of reasoned action: some applications and implications. Nebraska Symposium on Motivation, 1979. 65-116.

FISHBEIN, M., TRIANDIS, H. C., KANFER, F. H., BECKER, M. & MIDDLESTADT, S. E. 2001. Factors influencing behavior and behavior change.

FOWLER, F. J. 1995. *Improving survey questions: Design and Evaluation,* Thousand Oaks: Sage.

GAJALAKSHMI, V. & PETO, R. 2009. Smoking, drinking and incident tuberculosis in rural India: population-based case-control study. *International Journal of Epidemiology,* 38**,** 1018-25.

GAJALAKSHMI, V., PETO, R., KANAKA, T. S. & JHA, P. 2003. Smoking and mortality from tuberculosis and other diseases in India: retrospective study of 43 000 adult male deaths and 35 000 controls. *The Lancet,* 362**,** 507-515.

GEARING, R. E., EL-BASSEL, N., GHESQUIERE, A., BALDWIN, S., GILLIES, J. & NGEOW, E. 2011. Major ingredients of fidelity: A review and scientific guide to improving quality of intervention research implementation. *Clinical psychology review,* 31**,** 79-88.

GLASGOW, R. E., VOGT, T. M. & BOLES, S. M. 1999. Evaluating the public health impact of health promotion interventions: the RE-AIM framework. *American journal of public health,* 89**,** 1322-1327.

GRADE WORKING GROUP 2004. Grading quality of evidence and strength of recommendations. *BMJ: British Medical Journal,* 328**,** 1490.

GRANT, A., TREWEEK, S., DREISCHULTE, T., FOY, R. & GUTHRIE, B. 2013. Process evaluations for cluster-randomised trials of complex interventions: a proposed framework for design and reporting. *Trials,* 14**,** 15.

GREEN, E. C. & MURPHY, E. 2014. Health Belief Model. *The Wiley Blackwell Encyclopedia of Health, Illness, Behavior, and Society.* John Wiley & Sons, Ltd.

GUPTA, D., AGGARWAL, A., KUMAR, S. & JINDAL, S. 2001. Smoking increases risk of pulmonary tuberculosis. *J Environ Med,* 3**,** 65-70.

HABICHT, J.-P., VICTORA, C. & VAUGHAN, J. P. 1999. Evaluation designs for adequacy, plausibility and probability of public health programme performance and impact. *International journal of epidemiology,* 28**,** 10-18.

HAGGER, M. S. & HARDCASTLE, S. J. 2014. Interpersonal style should be included in taxonomies of behavior change techniques. *Frontiers in psychology,* 5**,** 254.

HAHN, E. A., DEVELLIS, R. F., BODE, R. K., GARCIA, S. F., CASTEL, L. D., EISEN, S. V., BOSWORTH, H. B., HEINEMANN, A. W., ROTHROCK, N. & CELLA, D. 2010. Measuring social health in the patient-reported outcomes measurement information system (PROMIS): item bank development and testing. *Quality of Life Research,* 19**,** 1035-1044.

HALL, D., JAMES, P. & ROBERTS, S. 1997. Evaluation of training in behaviour change counselling skills: the application of clinical-audit methodology. *Health Education Journal,* 56**,** 393-403.

HARDCASTLE, S. J., FORTIER, M., BLAKE, N. & HAGGER, M. S. 2016. Identifying content-based and relational techniques to change behavior in Motivational Interviewing. *Health psychology review***,** 1-46.

HARDEMAN, W., GRIFFIN, S., JOHNSTON, M., KINMONTH, A. & WAREHAM, N. 2000. Interventions to prevent weight gain: a systematic review of psychological models and behaviour change methods. *International journal of obesity,* 24**,** 131-143.

HARDEMAN, W., MICHIE, S., FANSHAWE, T., PREVOST, A. T., MCLOUGHLIN, K. & KINMONTH, A. L. 2008. Fidelity of delivery of a physical activity intervention: Predictors and consequences. *Psychology and Health,* 23**,** 11-24.

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**,** 676-687.

HASSON, H. 2010. Systematic evaluation of implementation fidelity of complex interventions in health and social care. *Implement Sci,* 5**,** 67.

HAWE, P., SHIELL, A. & RILEY, T. 2004. Complex interventions: how "out of control" can a randomised controlled trial be? *BMJ: British Medical Journal,* 328**,** 1561.

HAYES, A. F. & KRIPPENDORFF, K. 2007. Answering the call for a standard reliability measure for coding data. *Communication Methods and Measures,* 1**,** 77-89.

HECKMAN, C. J., EGLESTON, B. L. & HOFMANN, M. T. 2010. Efficacy of motivational interviewing for smoking cessation: a systematic review and meta-analysis. *Tobacco control,* 19**,** 410-416.

HECKMAN, J. J. 1981. Heterogeneity and state dependence. *Studies in labor markets.* University of Chicago Press.

HINKIN, T. 1995. A review of scale development practices in the study of organizations. *Journal of Management,* 21**,** 967-988.

HOFFMANN, T. C., GLASZIOU, P. P., BOUTRON, I., MILNE, R., PERERA, R., MOHER, D., ALTMAN, D. G., BARBOUR, V., MACDONALD, H. & JOHNSTON, M. 2014. Better reporting of interventions: template for intervention description and replication (TIDieR) checklist and guide. *BMJ: British Medical Journal,* 348.

HOLTER, M. C., MOWBRAY, C. T., BELLAMY, C. D., MACFARLANE, P. & DUKARSKI, J. 2004. Critical ingredients of consumer run services: Results of a national survey. *Community Mental Health Journal,* 40**,** 47-63.

HORNE, D. J., CAMPO, M., ORTIZ, J. R., OREN, E., ARENTZ, M., CROTHERS, K. & NARITA, M. 2012. Association between smoking and latent tuberculosis in the U.S. population: an analysis of the National Health and Nutrition Examination Survey. *PloS one,* 7**,** e49050.

HULSCHER, M., LAURANT, M. & GROL, R. 2003. Process evaluation on quality improvement interventions. *Quality and Safety in Health Care,* 12**,** 40-46.

HUSSAIN, H., AKHTAR, S. & NANAN, D. 2003. Prevalence of and risk factors associated with Mycobacterium tuberculosis infection in prisoners, North West Frontier Province, Pakistan. *International journal of epidemiology,* 32**,** 794-799.

INSTITUTE FOR HEALTH METRICS AND EVALUATION. 2010. *Global Burden of Disease (GBD) Visualizations.* [Online]. Available: http://www.healthmetricsandevaluation.org/gbd/visualizations/country [Accessed September 1, 2014].

JANSEN, Y. J., FOETS, M. M. & DE BONT, A. A. 2010. The contribution of qualitative research to the development of tailor-made community-based interventions in primary care: a review. *The European Journal of Public Health,* 20**,** 220-226.

JEE, S. H., GOLUB, J. E., JO, J., PARK, I. S., OHRR, H. & SAMET, J. M. 2009. Smoking and risk of tuberculosis incidence, mortality, and recurrence in South Korean men and women. *American journal of epidemiology,* 170**,** 1478-1485.

JOHNSTON, M. 2014. Improving the reporting of behaviour change interventions. *European Health Psychologist,* 16**,** 181-189.

KENDALL, M. G. & SMITH, B. B. 1939. The problem of m rankings. *The annals of mathematical statistics,* 10**,** 275-287.

KINMONTH, A.-L., WAREHAM, N. J., HARDEMAN, W., SUTTON, S., PREVOST, A. T., FANSHAWE, T., WILLIAMS, K. M., EKELUND, U., SPIEGELHALTER, D. & GRIFFIN, S. J. 2008. Efficacy of a theory-based behavioural intervention to increase physical activity in an at-risk group in primary care (ProActive UK): a randomised trial. *The Lancet,* 371**,** 41-48.

KIRKWOOD, B. R. & STERNE, J. A. 2003. *Essential medical statistics*, Blackwell Scientific Publications.

KNITTLE, K. 2014. Fidelity in intervention delivery: A rough field guide. *European Health Psychologist,* 16**,** 190-195.

KOLAPPAN, C. & GOPI, P. 2002. Tobacco smoking and pulmonary tuberculosis. *Thorax,* 57**,** 964-966.

KOZIOL, N. & ARTHUR, A. 2011. An Introduction to Secondary Data Analysis. *Research Methodology Series*.

KRIPPENDORFF, K. 2004. *Content analysis: An introduction to its methodology,* Thousand Oaks, California: Sage.

KUMARI, P. & MEENA, L. S. 2014. Factors Affecting Susceptibility to Mycobacterium tuberculosis: a Close View of Immunological Defence Mechanism. *Applied biochemistry and biotechnology,* 174**,** 2663-2673.

LAI, D., CAHILL, K., QIN, Y. & TANG, J. L. 2010. Motivational interviewing for smoking cessation. *Cochrane Database Syst Rev,* 1.

LAKES, K. D. & HOYT, W. T. 2009. Applications of Generalizability Theory to Clinical Child and Adolescent Psychology Research∗. *Journal of Clinical Child & Adolescent Psychology,* 38**,** 144-165.

LAM, T., HO, S., HEDLEY, A., MAK, K. & PETO, R. 2001. Mortality and smoking in Hong Kong: case-control study of all adult deaths in 1998. *Bmj,* 323**,** 361.

LANCASTER, T. & STEAD, L. 2005. Individual behavioural counselling for smoking cessation. *Cochrane Database of Systematic Reviews,* 2.

LANE, C., HUWS-THOMAS, M., HOOD, K., ROLLNICK, S., EDWARDS, K. & ROBLING, M. 2005. Measuring adaptations of motivational interviewing: the development and validation

of the behavior change counseling index (BECCI). *Patient education and counseling,* 56**,** 166-173.

LANE, D. 2013. *HyperStat Online Statistics Textbook- Chapter 6: Sampling distributions* [Online]. Available: http://davidmlane.com/hyperstat/A50760.html [Accessed March 11, 2015].

LARRY, H. & EDWARD, J. S. 1999. *A Step-by-Step Approach to Using SAS System for Univariate and Multivariate Statistics,* Cary, NC: SAS Institute Inc., 1994. 552 pp.

LAVIGNE, M., ROCHER, I., STEENSMA, C. & BRASSARD, P. 2006. The impact of smoking on adherence to treatment for latent tuberculosis infection. *BMC Public Health,* 6**,** 66.

LAWSON, P. J. & FLOCKE, S. A. 2009. Teachable moments for health behavior change: a concept analysis. *Patient education and counseling,* 76**,** 25-30.

LAWTON, R., MCEACHAN, R., JACKSON, C., WEST, R. & CONNER, M. 2015. Intervention fidelity and effectiveness of a UK worksite physical activity intervention funded by the Bupa Foundation, UK. *Health promotion international,* 30**,** 38-49.

LEECH, N. L. & ONWUEGBUZIE, A. J. 2010. Guidelines for conducting and reporting mixed research in the field of counseling and beyond. *Journal of Counseling & Development,* 88**,** 61-69.

LENNOX, A. S., BAIN, N., TAYLOR, R. J., MCKIE, L., DONNAN, P. T. & GROVES, J. 1998. Stages of change training for opportunistic smoking intervention by the primary health care team. Part I: randomised controlled trial of the effect of training on patient smoking outcomes and health professional behaviour as recalled by patients. *Health Education Journal,* 57**,** 140-149.

LEUNG, C. C., LI, T., LAM, T. H., YEW, W. W., LAW, W. S., TAM, C. M., CHAN, W. M., CHAN, C. K., HO, K. S. & CHANG, K. C. 2004. Smoking and tuberculosis among the elderly in Hong Kong. *American journal of respiratory and critical care medicine,* 170**,** 1027-1033.

LIENHARDT, C., FIELDING, K., SILLAH, J., BAH, B., GUSTAFSON, P., WARNDORFF, D., PALAYEW, M., LISSE, I., DONKOR, S. & DIALLO, S. 2005. Investigation of the risk factors for tuberculosis: a case–control study in three countries in West Africa. *International Journal of Epidemiology,* 34**,** 914-923.

LIM, S. S., VOS, T., FLAXMAN, A. D., DANAEI, G., SHIBUYA, K., ADAIR-ROHANI, H., AMANN, M., ANDERSON, H. R., ANDREWS, K. G. & ARYEE, M. 2013. A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010. *The Lancet,* 380**,** 2224-2260.

LIN, H. H., EZZATI, M. & MURRAY, M. 2007. Tobacco smoke, indoor air pollution and tuberculosis: a systematic review and meta-analysis. *PLoS medicine,* 4**,** e20.

LINDSAY, R. P., SHIN, S. S., GARFEIN, R. S., RUSCH, M. L. & NOVOTNY, T. E. 2014. The Association between active and passive smoking and latent tuberculosis infection in adults and children in the united states: results from NHANES. *PLoS ONE [Electronic Resource],* 9**,** e93137.

LINNAN, L. & STECKLER, A. 2002. *Process evaluation for public health interventions and research*, Jossey-Bass San Francisco, California.

LITTLETON, J. & PARK, J. 2009. Tuberculosis and syndemics: Implications for Pacific health in New Zealand. *Social Science & Medicine,* 69**,** 1674-1680.

LIU, B.-Q., PETO, R., CHEN, Z.-M., BOREHAM, J., WU, Y.-P., LI, J.-Y., CAMPBELL, T. C. & CHEN, J.-S. 1998. Emerging tobacco hazards in China: 1. Retrospective proportional mortality study of one million deaths. *Bmj,* 317**,** 1411-1422.

LONNROTH, K., JARAMILLO, E., WILLIAMS, B., DYE, C. & RAVIGLIONE, M. 2009. Drivers of tuberculosis epidemics: the role of risk factors and social determinants. *Soc Sci Med,* 68**,** 2240-6.

LORD, F. M. 1959. Statistical inferences about true scores. *Psychometrika,* 24**,** 1-17.

LORENCATTO, F., WEST, R., BRUGUERA, C., BROSE, L. S. & MICHIE, S. 2016. Assessing the Quality of Goal Setting in Behavioural Support for Smoking Cessation and its Association with Outcomes. *Annals of Behavioral Medicine,* 50**,** 310-318.

LORENCATTO, F., WEST, R., CHRISTOPHERSON, C. & MICHIE, S. 2013a. Assessing fidelity of delivery of smoking cessation behavioural support in practice. *Implementation Science,* 8**,** 40.

LORENCATTO, F., WEST, R., SEYMOUR, N. & MICHIE, S. 2013b. Developing a method for specifying the components of behavior change interventions in practice: the example of smoking cessation. *J Consult Clin Psychol,* 81**,** 528-44.

LORENCATTO, F., WEST, R., STAVRI, Z. & MICHIE, S. 2013c. How well is intervention content described in published reports of smoking cessation interventions? *nicotine & tobacco research,* 15**,** 1273-1282.

LOWE, C. 1956. An association between smoking and respiratory tuberculosis. *British medical journal,* 2**,** 1081.

LUCCA, A. 2000. A clubhouse fidelity index: Preliminary reliability and validity results. *Mental Health Services Research,* 2**,** 89-94.

MADSON, M. B. & CAMPBELL, T. C. 2006. Measures of fidelity in motivational enhancement: A systematic review. *Journal of substance abuse treatment,* 31**,** 67-73.

MARION, K. C., JILL, M. & JEREMY, M. G. 2001. Cluster trials in implementation research: estimation of intracluster correlation coefficients and sample size. *Statistics in Medicine,* 20**,** 391-399.

MASON, J., WOOD, J. & FREEMANTLE, N. 1999. Designing evaluations of interventions to change professional practice. *Journal of health services research & policy,* 4**,** 106.

MATEOS, F., BROCK, J. H. & PÉREZ-ARELLANO, J. L. 1998. Iron metabolism in the lower respiratory tract. *Thorax,* 53**,** 594-600.

MATHERS, C. D. & LONCAR, D. 2006. Projections of global mortality and burden of disease from 2002 to 2030. *PLoS medicine,* 3**,** e442.

MCBRIDE, C., EMMONS, K. & LIPKUS, I. 2003. Understanding the potential of teachable moments: the case of smoking cessation. *Health education research,* 18**,** 156-170.

MCCURDY, S. A., ARRETZ, D. S. & BATES, R. O. 1997. Tuberculin reactivity among California Hispanic migrant farm workers. *American journal of industrial medicine,* 32**,** 600-605.

MCEWEN, A. 2008. Call content manual for pro-active telephone smoking cessation intervention (Pro-active of reactive telephone support for stopping smoking [PORTSSS] trial.ed). London: Department of Health

MCEWEN, A., HAJEK, P., MCROBBIE, H. & WEST, R. 2006. *Manual of Smoking Cessation: A Guide for Counsellors and Practitioners,* Oxford: Wiley-Blackwell

MCGRAW, S. A., MCKINLAY, S. M., MCCLEMENTS, L., LASATER, T. M., ASSAF, A. & CARLETON, R. A. 1989. Methods in Program Evaluation The Process Evaluation System of the Pawtucket Heart Health Program. *Evaluation Review,* 13**,** 459-483.

MCGREW, J. H. & BOND, G. R. 1995. Critical ingredients of assertive community treatment: Judgments of the experts. *The Journal of Mental Health Administration,* 22**,** 113-125.

MEADER, N., KING, K., LLEWELLYN, A., NORMAN, G., BROWN, J., RODGERS, M., MOE-BYRNE, T., HIGGINS, J. P., SOWDEN, A. & STEWART, G. 2014. A checklist designed to aid consistency and reproducibility of GRADE assessments: development and pilot validation. *Systematic reviews,* 3**,** 1-9.

MEDICAL RESEARCH COUNCIL 2000. A framework for the development and evaluation of RCTs for complex interventions to improve health. *London: MRC, 2000*.

MESSICK, S. 1980. Test validity and the ethics of assessment. *American psychologist,* 35**,** 1012.

MICHIE, S. & ABRAHAM, C. 2004. Interventions to change health behaviours: Evidence-based or evidence-inspired? *Psychology & Health,* 19**,** 29-49.

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. *Implement Sci,* 6**,** 10.

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**,** 690.

MICHIE, S., ABRAHAM, C., WHITTINGTON, C., MCATEER, J. & GUPTA, S. 2009b. Identifying effective techniques in interventions: A meta-analysis and meta-regression. *Health Psychology,* 28**,** 690-671.

MICHIE, S., CHURCHILL, S. & WEST, R. 2011b. Identifying evidence-based competences required to deliver behavioural support for smoking cessation. *Annals of Behavioral Medicine,* 41**,** 59-70.

MICHIE, S., FIXSEN, D., GRIMSHAW, J. M. & ECCLES, M. P. 2009c. Specifying and reporting complex behaviour change interventions: the need for a scientific method. *Implementation Science,* 4**,** 1-6.

MICHIE, S., HYDER, N., WALIA, A. & WEST, R. 2011c. Development of a taxonomy of behaviour change techniques used in individual behavioural support for smoking cessation. *Addictive behaviors,* 36**,** 315-319.

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

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**,** 26-33.

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**,** 660-680.

MICHIE, S., JOHNSTON, M., ROTHMAN, A., KELLY, M. & DE BRUIN, M. 2014-2017. *Developing methodology for designing and evaluating theory-based complex interventions: an ontology for linking behaviour change techniques to theory* [Online]. University College London: Medical Research Council. Available: http://www.ucl.ac.uk/health-psychology/research/theories-techniques [Accessed].

MICHIE, S., RICHARDSON, M., JOHNSTON, M., ABRAHAM, C., FRANCIS, J., HARDEMAN, W., ECCLES, M. P., CANE, J. & 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**,** 81-95.

MICHIE, S., VAN STRALEN, M. & WEST, R. 2011d. The behaviour change wheel: a new method for characterising and designing behaviour change interventions. *Implementation Science,* 6**,** 42.

MIHALIC, S. 2004. The importance of implementation fidelity. *Emotional & Behavioral Disorders In Youth,* 4**,** 83-105.

MILLER, S. J. & BINDER, J. L. 2002. The effects of manual-based training on treatment fidelity and outcome: A review of the literature on adult individual psychotherapy. *Psychotherapy: Theory, Research, Practice, Training,* 39**,** 184.

MILLER, W. & ROLLNICK, S. 2013. *Motivational interviewing: Preparing people for change,* New York: Guildford Press.

MONCHER, F. J. & PRINZ, R. J. 1991. Treatment fidelity in outcome studies. *Clinical psychology review,* 11**,** 247-266.

MONTGOMERY, P., UNDERHILL, K., GARDNER, F., OPERARIO, D. & MAYO-WILSON, E. 2013. The Oxford Implementation Index: a new tool for incorporating implementation data into systematic reviews and meta-analyses. *Journal of clinical epidemiology,* 66**,** 874-882.

MOOK, D. 1996. *Motivation: the Organization of Action. ,* New York, W. W. Norton.

MOORE, G., AUDREY, S., BARKER, M., BOND, L., BONELL, C., HARDEMAN, W., MOORE, L., O'CATHAIN, A., TINATI, T. & WIGHT, D. 2015. Process evaluation of complex interventions. *Complex Interventions in Health: An Overview of Research Methods***,** 222.

MORGAN, D. L. 2007. Paradigms Lost and Pragmatism Regained: Methodological Implications of Combining Qualitative and Quantitative Methods. *Journal of Mixed Methods Research* 1**,** 48-76.

MOTALLEBZADEH, R., BLAND, J. M., MARKUS, H. S., KASKI, J. C. & JAHANGIRI, M. 2007. Neurocognitive function and cerebral emboli: randomized study of on-pump versus off-pump coronary artery bypass surgery. *The Annals of Thoracic Surgery,* 83**,** 475-482.

MOWBRAY, C. T., HOLTER, M. C., TEAGUE, G. B. & BYBEE, D. 2003. Fidelity criteria: Development, measurement, and validation. *American Journal of Evaluation,* 24**,** 315-340.

MOYERS, T., MARTIN, T., MANUEL, J., MILLER, W. & ERNST, D. 2010. Revised global scales: Motivational interviewing treatment integrity 3.1. 1 (MITI 3.1. 1). *Unpublished manuscript, University of New Mexico, Albuquerque, NM*.

MOYERS, T. B., MARTIN, T., MANUEL, J. K., HENDRICKSON, S. M. & MILLER, W. R. 2005. Assessing competence in the use of motivational interviewing. *Journal of substance abuse treatment,* 28**,** 19-26.

MURAVEN, M. & SLESSAREVA, E. 2003. Mechanisms of self-control failure: Motivation and limited resources. *Personality and Social Psychology Bulletin,* 29**,** 894-906.

NELSON, M. C., CORDRAY, D. S., HULLEMAN, C. S., DARROW, C. L. & SOMMER, E. C. 2012. A procedure for assessing intervention fidelity in experiments testing educational and behavioral interventions. *The journal of behavioral health services & research,* 39**,** 374-396.

NEUTENS, J. J. & RUBINSON, L. 2001. *Research techniques for the health sciences*, Benjamin-Cummings Publishing Company.

NG, M., FREEMAN, M. K., FLEMING, T. D., ROBINSON, M., DWYER-LINDGREN, L., THOMSON, B., WOLLUM, A., SANMAN, E., WULF, S. & LOPEZ, A. D. 2014. Smoking prevalence and cigarette consumption in 187 countries, 1980-2012. *Jama,* 311**,** 183-192.

NICE 2008. Smoking cessation services in primary care, pharmacies, local authorities and workplaces, particularly for manual working groups, pregnant women and hard to reach communities. Public Health Guidance: No. 10, February. NICE, London.

NIGG, C. R., ALLEGRANTE, J. P., & ORY, M. 2002. *Behavior Change Consortium,* Health Education Research.

NIGG, C. R. & LONG, C. R. 2012. A systematic review of single health behavior change interventions vs. multiple health behavior change interventions among older adults. *Translational behavioral medicine,* 2**,** 163-179.

NOVICK, M. R. 1966. The axioms and principle results of classical test theory. *Journal of Mathematical Psychology,* 3**,** 1-18.

NUNNALLY, J. C. 1978. Psychometric theory. *NY: McGraw-Hill*.

NUNNALLY, J. C. & BERNSTEIN, I. H. 1994. Psychometric theory. New York: McGraw-Hill

NUTBEAM, D. 1998. Evaluating health promotion—progress, problems and solutions. *Health promotion international,* 13**,** 27-44.

O'DONNELL, C. L. 2008. Defining, conceptualizing, and measuring fidelity of implementation and its relationship to outcomes in K–12 curriculum intervention research. *Review of Educational Research,* 78**,** 33-84.

OAKLEY, A., STRANGE, V., BONELL, C., ALLEN, E. & STEPHENSON, J. 2006a. Ripple Study Team, process evaluation in randomised controlled trials of complex interventions. *British Medical Journal,* 332**,** 413-416.

OAKLEY, A., STRANGE, V., BONELL, C., ALLEN, E., STEPHENSON, J. & TEAM, R. S. 2006b. Health services research: process evaluation in randomised controlled trials of complex interventions. *BMJ: British Medical Journal,* 332**,** 413.

OMRAN, A. R. 1971. EPIDEMIOLOGIC TRANSITION - THEORY OF EPIDEMIOLOGY OF POPULATION CHANGE. *Milbank Memorial Fund Quarterly,* 49**,** 509-&.

ONWUEGBUZIE, A. J. & COLLINS, K. M. 2007. A typology of mixed methods sampling designs in social science research. *The qualitative report,* 12**,** 281-316.

ONYEBUJOH, P., LEVIN, J., FOURIE, F., GARHIRAM, V., TEMBE, L., PHILI, N., MTHIYANE, T., MONIWA, T., BAYER, G. & RAMAJOE, I. 1999. Immunotherapy with Mycobacterium vaccae in patients with newly diagnosed pulmonary tuberculosis: a randomised controlled trial. *Lancet,* 354**,** 116-119.

ORNSTEIN, S. M., MUSHAM, C., REID, A., JENKINS, R. G., ZEMP, L. D. & GARR, D. R. 1993. Barriers to adherence to preventive services reminder letters: the patient's perspective. *Journal of Family Practice,* 36**,** 195-201.

OSBORNE, J. W. & COSTELLO, A. B. 2004. Sample size and subject to item ratio in principal components analysis. *Practical assessment, research & evaluation,* 9**,** 8.

OXFORD DICTIONARIES. 2015. *"motivation". Oxford Dictionaries* [Online]. Oxford University Press. Available: http://www.oxforddictionaries.com/definition/english/motivation [Accessed September 29, 2015].

PAI, M., MOHAN, A., DHEDA, K., LEUNG, C. C., YEW, W. W., CHRISTOPHER, D. J. & SHARMA, S. K. 2007. Lethal interaction: the colliding epidemics of tobacco and tuberculosis. *Expert review of anti-infective therapy,* 5**,** 385-391.

PARKER, D., EVANGELOU, E. & EATON, C. 2005. Intraclass correlation coefficients for cluster randomized trials in primary care: the cholesterol education and research trial (CEART). *Contemp Clin Trials,* 26.

PATRA, J., BHATIA, M., SURAWEERA, W., MORRIS, S. K., PATRA, C., GUPTA, P. C. & JHA, P. 2015. Exposure to Second-Hand Smoke and the Risk of Tuberculosis in Children and Adults: A Systematic Review and Meta-Analysis of 18 Observational Studies. *PLOS Med,* 12**,** e1001835.

PAWSON, R. & TILLEY, N. 1997. *Realistic evaluation*, Sage.

PEDNEKAR, M. S. & GUPTA, P. C. 2007. Prospective study of smoking and tuberculosis in India. *Preventive medicine,* 44**,** 496-498.

PEREPLETCHIKOVA, F. & KAZDIN, A. E. 2005. Treatment integrity and therapeutic change: Issues and research recommendations. *Clinical Psychology: Science and Practice,* 12**,** 365-383.

PETERS, D. H. 2009. *Improving health service delivery in developing countries: from evidence to action*, World Bank Publications.

PETERS, D. H., TRAN, N. T. & ADAM, T. 2013. Implementation research in health: a practical gudie. Alliance for Health Policy and Systems Research, World Health Organization.

PETO, R., CHEN, Z.-M. & BOREHAM, J. 1999. Tobacco- the growing epidemic. *Nature medicine,* 5**,** 15-17.

PETTICREW, M., REHFUESS, E., NOYES, J., HIGGINS, J. P., MAYHEW, A., PANTOJA, T., SHEMILT, I. & SOWDEN, A. 2013. Synthesizing evidence on complex interventions: how meta-analytical, qualitative, and mixed-method approaches can contribute. *Journal of clinical epidemiology,* 66**,** 1230-1243.

PIGOTT, T. & SHEPPERD, S. 2013. Identifying, documenting, and examining heterogeneity in systematic reviews of complex interventions. *Journal of clinical epidemiology,* 66**,** 1244-1250.

PINÉ-ABATA, H., MCNEILL, A., MURRAY, R., BITTON, A., RIGOTTI, N. & RAW, M. 2013. A survey of tobacco dependence treatment services in 121 countries. *Addiction,* 108**,** 1476-1484.

PIPER, M. E., FIORE, M. C., SMITH, S. S., FRASER, D., BOLT, D. M., COLLINS, L. M., MERMELSTEIN, R., SCHLAM, T. R., COOK, J. W. & JORENBY, D. E. 2016. Identifying effective intervention components for smoking cessation: a factorial screening experiment. *Addiction,* 111**,** 129-141.

PIRIE, P. L., STONE, E. J., ASSAF, A. R., FLORA, J. A. & MASCHEWSKY-SCHNEIDER, U. 1994. Program evaluation strategies for community-based health promotion programs: perspectives from the cardiovascular disease community research and demonstration studies. *Health Education Research,* 9**,** 23-36.

PREUSS, R. A. 2013. Using Generalizability Theory to Develop Clinical Assessment Protocols. *Physical therapy,* 93**,** 562-569.

PROCHASKA, J. & DICLEMENTE, C. 1984. Self change processes, self efficacy and decisional balance across five stages of smoking cessation. *Progress in clinical and biological research,* 156**,** 131.

PROCTOR, E., SILMERE, H., RAGHAVAN, R., HOVMAND, P., AARONS, G., BUNGER, A., GRIFFEY, R. & HENSLEY, M. 2011. Outcomes for implementation research: conceptual distinctions, measurement challenges, and research agenda. *Administration and Policy in Mental Health and Mental Health Services Research,* 38**,** 65-76.

PUTKA, D. J. & MCCLOY, R. A. 2008. Estimating Variance Components in SPSS and SAS: An Annotated Reference Guide.

RAO, V. G., GOPI, P. G., BHAT, J., YADAV, R., SELVAKUMAR, N. & WARES, D. F. 2012. Selected risk factors associated with pulmonary tuberculosis among Saharia tribe of Madhya Pradesh, central India. *European Journal of Public Health,* 22**,** 271-3.

RESNICK, B., INGUITO, P., ORWIG, D., YAHIRO, J. Y., HAWKES, W., WERNER, M., ZIMMERMAN, S. & MAGAZINER, J. 2005. Treatment fidelity in behavior change research: A case example. *Nursing research,* 54**,** 139-143.

RESNICOW, K. & VAUGHAN, R. 2006. A chaotic view of behavior change: a quantum leap for health promotion. *International Journal of Behavioral Nutrition and Physical Activity,* 3**,** 25.

RICE, V. & STEAD, L. 2008. Nursing interventions for smoking cessation. *Cochrane Database Syst Rev,* 1.

ROSSI, P. H., LIPSEY, M. W. & FREEMAN, H. E. 2003. *Evaluation: A systematic approach*, Sage publications.

ROTH, A. D. & PILLING, S. 2008. Using an evidence-based methodology to identify the competences required to deliver effective cognitive and behavioural therapy for depression and anxiety disorders. *Behavioural and Cognitive Psychotherapy,* 36**,** 129-147.

SAPNAS, K. G. & ZELLER, R. A. 2002. Minimizing sample size when using exploratory factor analysis for measurement. *Journal of nursing measurement,* 10**,** 135-154.

SCHNEIDER, N. K. & NOVOTNY, T. E. 2007. Addressing smoking cessation in tuberculosis control. *Bulletin of the World Health Organization,* 85**,** 820-821.

SCHOENWALD, S. K., GARLAND, A. F., CHAPMAN, J. E., FRAZIER, S. L., SHEIDOW, A. J. & SOUTHAM-GEROW, M. A. 2011. Toward the effective and efficient measurement of implementation fidelity. *Administration and Policy in Mental Health and Mental Health Services Research,* 38**,** 32-43.

SCHÜNEMANN, H., HILL, S., GUYATT, G., AKL, E. A. & AHMED, F. 2011. The GRADE approach and Bradford Hill's criteria for causation. *Journal of epidemiology and community health,* 65**,** 392-395.

SCOTT, A. G. & SECHREST, L. 1989. Strength of theory and theory of strength. *Evaluation and Program Planning,* 12**,** 329-336.

SERENO, A. B., SOARES, E. C. C., LAPA E SILVA, J. R., NÁPOLES, A. M., BIALOUS, S. A., DA COSTA E SILVA, V. L. & NOVOTONY, E. T. 2012. Feasibility study of a smoking cessation intervention in Directly Observed Therapy Short-Course tuberculosis treatment clinics in Rio de Janeiro, Brazil. *Rev Panam Salud Publica,* 32**,** 451-6.

SHACHAR, R. & ECKSTEIN, Z. 2007. Correcting for bias in retrospective data. *Journal of applied econometrics,* 22**,** 657-675.

SHELDON, T. A. 2001. It ain't what you do but the way that you do it. *Journal of health services research & policy,* 6**,** 3-5.

SHETTY, N., SHEMKO, M., VAZ, M. & D'SOUZA, G. 2006. An epidemiological evaluation of risk factors for tuberculosis in South India: a matched case control study. *The International Journal of Tuberculosis and Lung Disease,* 10**,** 80-86.

SHIELL, A., HAWE, P. & GOLD, L. 2008. Complex interventions or complex systems? Implications for health economic evaluation. *BMJ: British Medical Journal,* 336**,** 1281.

SHOUKRI, M. M. & CIHON, C. 1998. *Statistical methods for health sciences*, CRC Press.

SIDDIQI, K. 2014. Supporting economically disadvantaged tobacco users to help them stop: time to review priorities? *Addiction,* 109**,** 1221-1222.

SIDDIQI, K., KHAN, A., AHMAD, M., DOGAR, O., KANAAN, M., NEWELL, J. & THOMSON, H. 2013. Action to Stop Smoking In Suspected Tuberculosis (ASSIST) in Pakistan: a Cluster Randomized, Controlled Trial. *Annals of internal medicine,* 158**,** 667-675.

SIDDIQUI, O., HEDEKER, D., FLAY, B. R. & HU, F. B. 1996. Intraclass Correlation Estimates in a School-based Smoking Prevention Study Outcome and Mediating Variables, by Sex and Ethnicity. *American Journal of Epidemiology,* 144**,** 425-433.

SLAMA, K., CHIANG, C. Y., ENARSON, D. A., HASSMILLER, K., FANNING, A., GUPTA, P. & RAY, C. 2007. Tobacco and tuberculosis: a qualitative systematic review and meta-analysis *The International Journal of Tuberculosis and Lung Disease,* 11**,** 1049-1061.

SMALL, R., YELLAND, J., LUMLEY, J., RICE, P. L., COTRONEI, V. & WARREN, R. 1999. Cross-cultural research: trying to do it better 2. Enhancing data quality. *Australian and New Zealand Journal of Public Health,* 23**,** 390-395.

SMITH, A. K., AYANIAN, J. Z., COVINSKY, K. E., LANDON, B. E., MCCARTHY, E. P., WEE, C. C. & STEINMAN, M. A. 2011. Conducting high-value secondary dataset analysis: an introductory guide and resources. *Journal of general internal medicine,* 26**,** 920-929.

STEAD, L. F. & LANCASTER, T. 2012. Combined pharmacotherapy and behavioural interventions for smoking cessation. *The Cochrane Library,* 10.

STETLER, C. B., LEGRO, M. W., WALLACE, C. M., BOWMAN, C., GUIHAN, M., HAGEDORN, H., KIMMEL, B., SHARP, N. D. & SMITH, J. L. 2006. The role of formative evaluation in implementation research and the QUERI experience. *Journal of General Internal Medicine,* 21**,** S1-S8.

STREINER, D. L. & NORMAN, G. R. 2008. *Health measurement scales: A practical guide to their development and use,* Oxford University Press

STUCKLER, D. 2008. Population causes and consequences of leading chronic diseases: A comparative analysis of prevailing explanations. *Milbank Quarterly,* 86**,** 273-326.

STUCKLER, D., BASU, S. & MCKEE, M. 2010. Drivers of inequality in Millennium Development Goal progress: a statistical analysis. *PLoS medicine,* 7**,** e1000241.

SZKLO, M. & NIETO, J. 2007. *Epidemiology: Beyond the Basics*, Jones & Bartlett Publishers.

TALJAARD, M., DONNER, A., VILLAR, J., WOJDYLA, D., VELAZCO, A., BATAGLIA, V., FAUNDES, A., LANGER, A., NARVÁEZ, A. & VALLADARES, E. 2008. Intracluster correlation coefficients from the 2005 WHO Global Survey on Maternal and Perinatal Health: implications for implementation research. *Paediatric and perinatal epidemiology,* 22**,** 117-125.

TEAGUE, G., DRAKE, R. E. & ACKERSON, T. 1995. Evaluating use of continuous treatment teams for persons with mental illness and substance abuse. *Psychiatric Services,* 46**,** 689-695.

THOMAS, A., GOPI, P., SANTHA, T., CHANDRASEKARAN, V., SUBRAMANI, R., SELVAKUMAR, N., EUSUFF, S., SADACHARAM, K. & NARAYANAN, P. 2005. Predictors of relapse among pulmonary tuberculosis patients treated in a DOTS programme in South India. *The International Journal of Tuberculosis and Lung Disease,* 9**,** 556-561.

TOBER, G., CLYNE, W., FINNEGAN, O., FARRIN, A., TEAM, I. R. & TEAM, U. R. 2008. Validation of a scale for rating the delivery of psycho-social treatments for alcohol dependence and misuse: the UKATT process rating scale (PRS). *Alcohol and alcoholism,* 43**,** 675-682.

TONES, K. 2000. Evaluating health promotion: a tale of three errors. *Patient education and counseling,* 39**,** 227-236.

UNESCO INSTITUTE FOR STATISTICS. *United Nations Educational, Scientific and Cultural Organisation (UNESCO), country profile-Pakistan* [Online]. Available: http://en.unesco.org/countries/pakistan [Accessed Sep 20, 2016].

VAN ZYL-SMIT, R. N., BRUNET, L., PAI, M. & YEW, W. W. 2010. The convergence of the global smoking, COPD, tuberculosis, HIV, and respiratory infection epidemics. *Infectious disease clinics of North America,* 24**,** 693.

VARTANIAN, T. P. 2010. *Secondary data analysis*, Oxford University Press.

WALTZ, J., ADDIS, M. E., KOERNER, K. & JACOBSON, N. S. 1993. Testing the integrity of a psychotherapy protocol: assessment of adherence and competence. *Journal of consulting and clinical psychology,* 61**,** 620.

WANG, G., SLEIGH, A., ZHOU, G., JACKSON, S. & LIU, X. 2005. [Nonbiologic risk factors of pulmonary tuberculosis among adults in Henan: a case-control study]. *Zhonghua liu xing bing xue za zhi= Zhonghua liuxingbingxue zazhi,* 26**,** 92-96.

WANG, J. & SHEN, H. 2009. Review of cigarette smoking and tuberculosis in China: intervention is needed for smoking cessation among tuberculosis patients. *BMC public health,* 9**,** 292.

WEBB, G. 1918. The effect of the inhalation of cigarette smoke on the lungs. A clinical study. *Am Rev Tuberc,* 2**,** 25-27.

WEISMAN, A., TOMPSON, M. C., OKAZAKI, S., GREGORY, J., GOLDSTEIN, M. J., REA, M. & MIKLOWITZ, D. J. 2002. Clinicians' Fidelity to a Manual-Based Family Treatment as a Predictor of the One-Year Course of Bipolar Disorder. *Family process,* 41**,** 123-131.

WEISS, C. H. 1997. Theory-based evaluation: Past, present, and future. *New directions for evaluation,* 1997**,** 41-55.

WEN, C. P., CHAN, T. C., CHAN, H. T., TSAI, M. K., CHENG, T. Y. & TSAI, S. P. 2010. The reduction of tuberculosis risks by smoking cessation. *BMC infectious diseases,* 10**,** 156.

WEST, R. 2005. Time for a change: putting the Transtheoretical (Stages of Change) Model to rest. *Addiction,* 100**,** 1036-1039.

WEST, R. 2009. The multiple facets of cigarette addiction and what they mean for encouraging and helping smokers to stop. *COPD: Journal of Chronic Obstructive Pulmonary Disease,* 6**,** 277-283.

WEST, R. 2013. *The SmokeFree Formula: A Revolutionary Way to Stop Smoking Now*, Hachette UK.

WEST, R., HAJEK, P., STEAD, L. & STAPLETON, J. 2005. Outcome criteria in smoking cessation trials: proposal for a common standard. *Addiction,* 100**,** 299-303.

WEST, R., WALIA, A., HYDER, N., SHAHAB, L. & MICHIE, S. 2010. Behavior change techniques used by the English Stop Smoking Services and their associations with short-term quit outcomes. *Nicotine & Tobacco Research,* 12**,** 742-747.

WEST, R., ZATONSKI, W., PRZEWOZNIAK, K. & JARVIS, M. J. 2007. Can we trust national smoking prevalence figures? Discrepancies between biochemically assessed and self-reported smoking rates in three countries. *Cancer Epidemiology Biomarkers & Prevention,* 16**,** 820-822.

WHYTE, F. & KEARNEY, N. 2003. Enhancing the nurse's role in tobacco control. *Tobacco Control Factsheets*.

WILSON, K., SENAY, I., DURANTINI, M., SÁNCHEZ, F., HENNESSY, M., SPRING, B. & ALBARRACÍN, D. 2015. When it comes to lifestyle recommendations, more is sometimes less: A meta-analysis of theoretical assumptions underlying the effectiveness of interventions promoting multiple behavior domain change. *Psychological bulletin,* 141**,** 474.

WORLD HEALTH ORGANISATION. 2014a. *Definitions and reporting framework for tuberculosis- 2013 revision* [Online]. WHO/HTM/TB/2013.2. Available: http://www.who.int/tb/publications/definitions/en/ [Accessed August 3, 2015].

WORLD HEALTH ORGANISATION. 2014b. *Global Tuberculosis Report* [Online]. Available: http://www.who.int/tb/publications/global_report/en/ [Accessed May 21, 2015].

WORLD HEALTH ORGANISATION. 2014c. *Toolkit for delivering the 5A's and 5R's brief tobacco interventions in primary care* [Online]. Available: http://www.who.int/tobacco/publications/smoking_cessation/9789241506953/en/ [Accessed March 04, 2015].

WORLD HEALTH ORGANIZATION 2008. WHO Report on the Global Tobacco Epidemic, 2008: the MPOWER package. Geneva: World Health Organization.

WORLD HEALTH ORGANIZATION. 2009. *Tuberculosis & Tobacco- a strong association* [Online]. Available: http://www.who.int/tobacco/resources/publications/factsheet_tb_tobacco_sep09.pdf [Accessed May 28, 2015].

YOUNG, J. & BECK, A. T. 1980. Cognitive therapy scale: Rating manual. *Unpublished manuscript*.

YU, G.-P., HSIEHT, C.-C. & PENG, J. 1988. Risk factors associated with the prevalence of pulmonary tuberculosis among sanitary workers in Shanghai. *Tubercle,* 69**,** 105-112.

ZELLER, R. A. & MARTSOLF, D. 2002. Statistical tools in applied research. *Unpublished manuscript*.

ZHAO, L., CHEN, Y. & SCHAFFNER, D. W. 2001. Comparison of logistic regression and linear regression in modeling percentage data. *Applied and environmental microbiology,* 67**,** 2129-2135.

ZIGNOL, M., HOSSEINI, M. S., WRIGHT, A., LAMBREGTS–VAN WEEZENBEEK, C., NUNN, P., WATT, C. J., WILLIAMS, B. G. & DYE, C. 2006. Global incidence of multidrug-resistant tuberculosis. *Journal of Infectious Diseases,* 194**,** 479-485.