

Measuring fidelity to behavioural support delivery for smoking cessation and its association with outcomes

Omara Dogar^{1,2} , Jan R. Boehnke^{1,3}, Fabiana Lorencatto⁴, Trevor A. Sheldon^{1,5} & Kamran Siddiqi^{1,5} 

Department of Health Sciences, University of York, York, UK,¹ The University of Edinburgh, Usher Institute, Edinburgh, UK,² School of Nursing and Health Sciences, University of Dundee, Dundee, UK,³ Centre for Behaviour Change, University College London, London, UK⁴ and Hull York Medical School, University of York, York, UK⁵

ABSTRACT

Background and aims Behavioural support increases smoking cessation in clinical settings, but effect sizes differ among providers, due possibly to variations in delivery. This study evaluates a measure ('fidelity index') intended to capture fidelity to delivery of content- and interaction-based items of a behavioural support (BS) for smoking cessation and the association of fidelity with quit rates. **Methods** A fidelity index for scoring the adherence and quality domains of a specific BS intervention, '5As for quit', was developed by classifying the intervention components using the taxonomy of behaviour change techniques. The index was applied to code 154 BS sessions audiotaped among 18 chest clinics in Pakistan to assess their fidelity and explore reliability of coding. The association between intervention fidelity and successful quit achieved by the same providers in a previous study was explored using regression analysis. **Results** The index represented two domains: adherence to delivery of content-based activities of 5As (37 items) and quality of interaction-based activities (eight items). The intercoder reliability was good for content-based (average Krippendorff's $\alpha = 0.80$) and moderate for interaction-based (average Krippendorff's $\alpha = 0.66$) items. Approximately 70% (intraclass correlation coefficient: adherence scores = 0.72, quality scores = 0.71) of variation in BS delivery was contributed by providers, which increased to 97% (g -coefficient: adherence scores = 0.973, quality scores = 0.974) after accounting for other sources of variation. Higher quit rates were positively associated with average quality scores [risk ratio = 2.15; 95% confidence interval (CI) = 1.43–3.24], but negatively associated with average adherence scores (risk ratio = 0.55; 95% CI = 0.40–0.77) within services. **Conclusions** The fidelity index is a reliable measure for quantifying intervention fidelity of delivering smoking cessation behavioural support. Recommended revisions of the fidelity index include incorporation of additional interaction-based items, such as the relational techniques used in motivational interviewing.

Keywords Behaviour change techniques, behavioural support, fidelity index, fidelity scores, intervention fidelity, smoking cessation.

Correspondence to: Omara Dogar, University of York, Department of Health Sciences, Seeborn Rowntree Building, Alcuin way, YO10 5DD, Heslington, York, UK. E-mail: omara.dogar@york.ac.uk

Submitted 17 May 2019; initial review completed 16 July 2019; final version accepted 2 September 2019

INTRODUCTION

Behaviour change interventions are complex, consisting of multiple, often interacting features [1]. Behavioural support, a promising intervention for smoking cessation in low resource settings, consists of advice, discussion, encouragement and activities designed to change smoking behaviours [2,3]. Behavioural support, if delivered as planned, on average produces expected cessation outcomes [4,5]. For inferences to be drawn from effectiveness studies, participant characteristics, intervention delivery, competence of the provider and the interaction between provider

and patient need to be measured and their influence estimated [6].

In practice, behavioural interventions are often delivered inconsistently [7,8], resulting in variations in outcomes [9–11]. For example, a behavioural support (BS) intervention was offered to smokers attending chest clinics in Pakistan: Action to Stop Smoking In Suspected Tuberculosis (ASSIST), and was found effective [relative risk 8.5, 95% confidence interval (CI) = 3.7–19.6] [12]. However, the wide variation in quit rates among chest clinics (7–70%) and a high intraclass correlation coefficient (ICC = 0.28), indicated the strong influence of clinics in

determining the success of the intervention. Context, factors extrinsic to the intervention that tend to hinder or bolster its effect [13] including training and technical support to the providers, is likely to influence outcomes [14]. The competences of the providers, their motivation and intention to provide support, the efficiency of the system and organizational structure stretched practice time for cessation provision and the providers' own smoking status; all supplemented the complexity of the BS delivered in ASSIST [15].

Intervention fidelity helps to attribute effects more accurately by explaining variation and increases confidence in interpretation of outcomes [7]. For example, where there are no fidelity assessments, effects could be wrongly attributed to the intervention itself rather than a difference in provider competence [6]. Equally, the potential emergence of an intervention effect may be masked by variations in the extent and quality of its delivery [6]. Conceptual models of fidelity differentiate between fidelity of design, training, delivery, receipt and enactment [13,16]. Fidelity of delivery is the degree to which an intervention is actually implemented as intended to the patient by the provider [17] and encompasses two main concepts: what is delivered and how it is delivered [18]. Fidelity to delivery of BS for smoking cessation was the mainstay of this study, and thus the term 'fidelity' in this paper refers broadly to the element of delivery. Measuring fidelity could document differences in delivery practices [19] and compare the implemented intervention components to the theory base by linking with outcomes [20,21].

It is feasible to reliably assess the fidelity of behaviour change interventions in clinical practice [22]. However, there is limited evidence regarding methods for quantifying fidelity within complex behaviour change interventions [23]. A method to measure fidelity to BS delivery requires capturing what is delivered: both its content- and interaction-based activities [13,16] and quantifying how these are delivered [24,25]. These can be conceptualized as compositional/structural features forming items of the index and functional/process features forming the anchors to quantify each item. While studies have previously developed methods to characterize the content of BS using the behaviour change techniques (BCTs) taxonomies [26–29] and BCT associations with outcomes [30], these limit ratings to the delivery process and their predictive value for outcomes. The extent to which fidelity of delivery of a BS intervention influences smoking cessation outcomes remains unclear.

The purpose of this study was to develop a BCT-linked, reliable method of rating intervention delivery ('fidelity index') and determine fidelity's association with smoking cessation outcomes. We report how we developed this index using BCT taxonomy, evaluated its psychometric properties and explored associations with quit rates.

METHODS

The study was conducted in routine tuberculosis (TB) services in Pakistan, involving the same sites as the ASSIST trial [12] and the health workers trained to deliver BS for smoking cessation. Ethics approval was obtained from the Health Sciences Research Governance Committee (HSRGC) at the University of York and the National Bioethics Committee at Pakistan Medical Research Council (PMRC).

STUDY I: DEVELOPMENT OF THE FIDELITY INDEX

Quantifying fidelity was conceptualized using measurement methods in the psychotherapeutic [31], health and education literature [21,32,33]. First, items were generated, then response scale options for each item were constructed, and finally experts carried out the content validation of the fidelity index.

The logic model for measuring fidelity and the conceptualization of its compositional and functional features are shown in the Supporting information, Figs S1 and S2. BS in ASSIST study was structured using the 5As (Ask, Advise, Assess, Assist, Arrange) to quit model [34] (Supporting information, Table S1). A confirmatory approach [31] to critical components technique was used to generate index items (Supporting information, Table S2) by identifying intervention activities that were measurable [32]. Activities were mapped using the smoking cessation BCTs [35], which provided the coding framework for the intended BS delivery. BCTs with a focus on specific behaviour (address motivation, maximize self-regulation) were mapped to the content-based items, which were expected to constitute a score representing the adherence to the BS (A score). Those with a focus on general aspects of interaction (delivery of intervention and general communication) specified the interaction-based items, expected to represent the quality of interaction (Q score) (Supporting information, Table S3).

To detect meaningful variation in delivery rather than the mere presence or absence of delivered items [36], a three-point ordinal response scale—fully, partially or not implemented—was chosen per item, such that the sum across item scores would provide a fidelity score reflecting the degree of implementation [37,38]. The response scale of each item used behavioural anchors that were operationally defined using the interaction style specified in the BS training manual (Supporting information, Table S4).

The items of the fidelity index were evaluated in a small Delphi study, which did not lead to any changes to the index (details in Supporting information, Appendix S1).

Psychometric validation of the fidelity index

Design and settings

The index was validated using an observational study in chest clinics in the Jhang and Sargodha districts of Pakistan in 2014, the sites of the ASSIST trial.

Participants

Providers were health workers responsible for dispensing medicines and recording patient data at the chest clinics. All those who delivered smoking cessation in the ASSIST trial were eligible to participate. Those who consented were offered a refresher training on BS for smoking cessation; each chest clinic had one provider.

Adult TB patients attending the chest clinics and smoking on a daily basis, except those who required hospitalization, were eligible to participate upon consent.

The number of providers ($n = 18$) was set by the number of participating chest clinics from ASSIST and was considered sufficiently diverse for the main analysis focusing on inter-coder reliability of the individual items. The sample size for individual items was evaluated via Krippendorff's α [39]. As the distribution of this statistic is not known [39], determining a minimal sample size for a specific level of accuracy is not directly possible. We evaluated the impact of different sample sizes ($n = 50$ – 250) on the breadth of its bootstrapped confidence interval for scenarios with $\alpha \approx 0.70$ and low as well as average fidelity ratings across three coders. The results indicated that the breadth of the index's confidence interval shrank substantially from $n = 50$ (~ 0.26) to $n = 100$ (~ 0.15), but that the gains were less substantial beyond that (e.g. $n = 200$, interval breadth ~ 0.13). We therefore aimed for at least $n = 100$ observations.

Measures

Three bilingual coders with a graduate degree in social or health sciences, trained on reliably specifying BS [40], coded sessions independently by listening to audiotapes. Disagreements on item ratings were resolved by discussion and a fourth (consensus) index filled for each item on agreed scores.

Analyses

The scores were summarized both for each item using mean, median and standard deviations and for each domain of the index, i.e. adherence (Ascore) and quality (Qscore), using the mean and 95% CIs supplied by the provider.

Psychometric properties of the fidelity index were assessed using three approaches: inter-coder reliability, principal components analysis (PCA) and via a generalizability study (G-study) [41]. Inter-coder reliability and PCA assessed the item scores, while the G-study assessed

the domain scores. All analyses were conducted in SAS version 9.4 (Cary, NC, USA; syntax in Supporting information, Table S5).

Inter-coder reliability of the item scores was computed using Krippendorff's alpha [42], and interpreted as poor for values of $\alpha < 0.67$, moderate between 0.67 and 0.80 and good for $\alpha > 0.80$ [43]. To identify items that reliably loaded across all coders, PCA with promax rotation was employed [44]. Adherence was further analysed to discriminate dimensions of the content-based items. The G-study assessed the reliability for measuring provider differences in intervention fidelity under generalizability theory. It involved three sequential steps: identification of important sources of variation, variance partitioning and computation of coefficient of generalizability (g), detailed in the Supporting information, Appendix S2.

Sensitivity analyses included: (i) estimation of the crude ICC and (ii) modelling district as a fixed effect to test for systematic differences in fidelity (Supporting information, Appendix S2).

STUDY 2: ASSOCIATION BETWEEN INTERVENTION FIDELITY AND OUTCOMES

Design

The fidelity scores (study conducted in 2014) for both adherence and quality were linked with quit rates achieved by the same providers in the ASSIST trial, conducted in 2010 [12]. Patient data were aggregated at service-level for both studies.

Measures

Outcome involved self-reported 6-month continuous abstinence, verified by carbon monoxide measurement of 9 parts per million (p.p.m.) or less [45] for 1299 patients from the ASSIST trial [12]. The proportion of patients who quit by providers ($n = 22$) were used in analysis.

The predictor variable involved standardized adherence and quality scores aggregated by provider.

Contextual variables (from both studies) used provider practice scores based on the self-record checklist of selected intervention components. The purpose was to establish consistency in providers' practice despite the time interval between ASSIST trial and this study (detailed in Supporting information, Appendix S3).

Analysis

Intervention fidelity was described by the percentage of sessions that fully, partially or did not implement BS activities. Patient characteristics for both studies are provided in Supporting information, Table S6. Kendall's W statistic

measured the pairwise concordance in providers' practice over time.

Binomial regression was used to estimate the association between average fidelity scores and quit rates, within a service (SAS PROC GENMOD). Risk ratios (RR) and 95% CIs were reported. Both fidelity scores were centred for use in the interaction term.

RESULTS

The fidelity index (Supporting information, Table S7) consisted of 45 items overall, 37 for measuring adherence: content-based items and eight for measuring quality: interaction-based items.

Study 1. Psychometric validation of the fidelity index

Between February and May 2014, 180 patients were enrolled into the study and received BS for smoking cessation at 19 chest clinics (Fig. 1). Three coders scored 154 audiotaped BS sessions for fidelity providing 462 data-points.

Description of item and domain scores and inter-coder reliability

Items 10, 12, 39, 40 and 44 showed very little variance, exhibiting floor and ceiling effects (Supporting information, Table S8). For items 10 and 12 the average score approached the maximum possible value of 2 (ceiling effect) as the majority of the providers fully implemented them. For items 39, 40 and 44 a floor effect was observed, as the majority of the providers did not implement them.

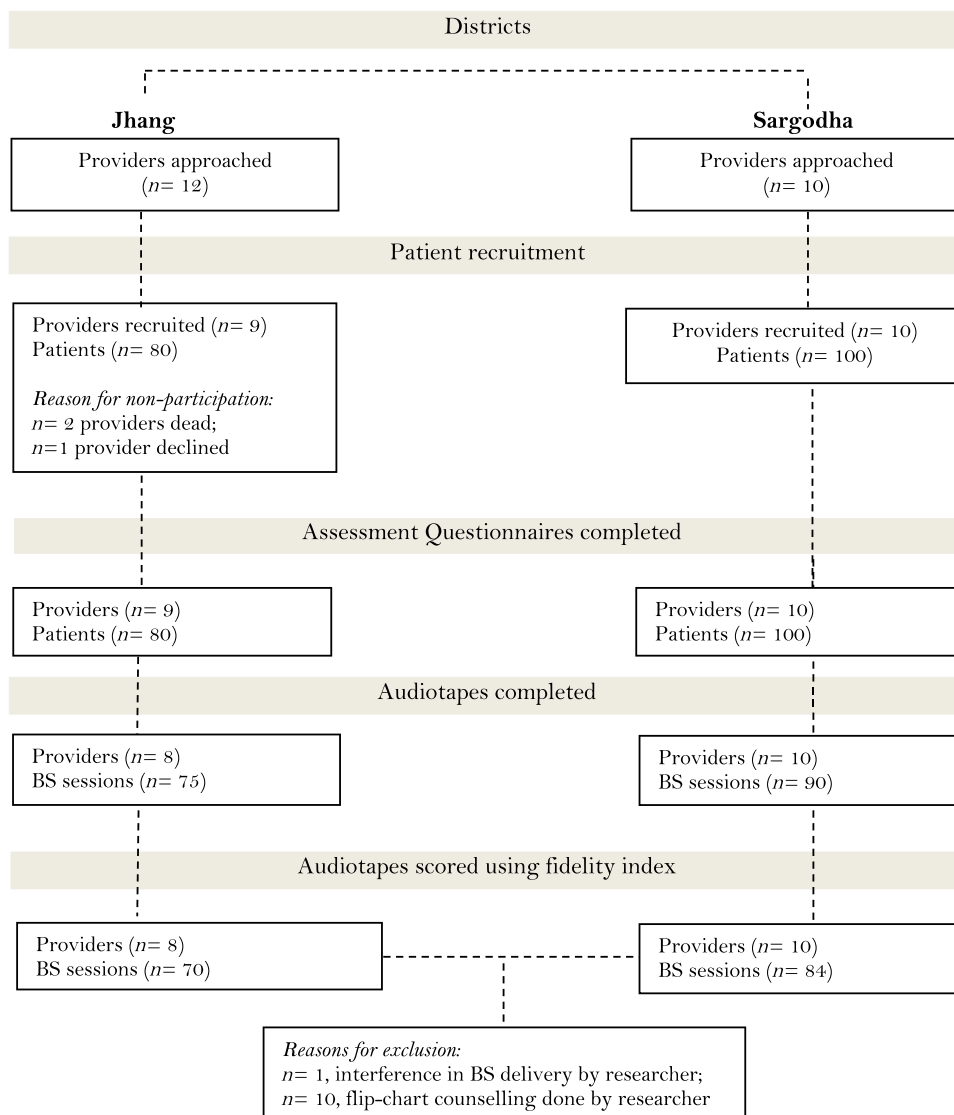


Figure 1 Study flowchart [Colour figure can be viewed at wileyonlinelibrary.com]

Items 12, 22 and 24 to 28, 38, 39 and 40 showed lower agreement (Krippendorff's $\alpha < 0.67$) between the three coders (Table 1). The overall mean Krippendorff's α was 0.80 for content-based items and 0.66 for interaction-based items.

The fidelity scores ranged from 11.17 (95% CI = 8.62–13.71) to 62.94 (95% CI = 59.51–66.36) for adherence, with a maximum possible score of 74, and ranged from 0 to 11.06 (95% CI = 10.00–12.13) for quality, with a maximum possible score of 16 (Table 2). The mean scores and

Table 1 Item description, average coder scores and inter-coder reliability estimates.

Item number	Item content	Coder 1 score mean	Coder 2 score mean	Coder 3 score mean	Krippendorff's α (95% CI)
Adherence (content-based items)					
	Assess smoking behaviour	1.29	1.27	1.28	0.868 (0.80–0.93)
Item 2	Provide awareness about forms of tobacco	1.35	1.23	1.24	0.755 (0.69–0.82)
Item 3	Provide information on harms of tobacco use: blood pressure and heart	1.34	1.20	1.11	0.715 (0.64–0.79)
Item 4	Provide information on harms of tobacco use: lung	1.37	1.27	1.19	0.738 (0.66–0.81)
Item 5	Provide information on harms of tobacco use: \$, oral health	1.33	1.16	1.11	0.707 (0.63–0.78)
Item 6	Provide information on harms of tobacco use: children health	1.39	1.30	1.20	0.755 (0.68–0.82)
Item 7	Provide information on harms of tobacco use: pregnancy	1.24	1.18	1.16	0.719 (0.64–0.79)
Item 8	Decide to quit	0.99	0.97	0.95	0.759 (0.68–0.83)
Item 9	Provide information on benefits of quitting	0.97	0.92	0.88	0.861 (0.81–0.90)
Item 10	Assess readiness to quit	1.92	1.88	1.86	0.675 (0.34–0.91)
Item 11	Elicit reason to quit	1.41	1.37	1.37	0.782 (0.71–0.85)
Item 12	Prompt commitment to quit	1.89	1.82	1.82	0.519 (0.18–0.79)
Item 13	Assess quit in past	1.74	1.77	1.79	0.734 (0.61–0.85)
Item 14	Assess past history of quit attempts	1.35	1.37	1.44	0.729 (0.63–0.83)
Item 15	Assess nicotine dependence	1.74	1.77	1.79	0.885 (0.78–0.98)
Item 16	Set quit date	0.75	0.75	0.67	0.774 (0.63–0.89)
Item 17	Advise on hiding reminders	0.83	0.83	0.81	0.856 (0.73–0.96)
Item 18	Advise on declaring home smoke-free	0.77	0.70	0.69	0.858 (0.74–0.96)
Item 19	Advise on social support	0.77	0.79	0.77	0.890 (0.79–0.98)
Item 20	Advise on telling others about the quit attempt	0.76	0.81	0.75	0.862 (0.75–0.96)
Item 21	Identify trigger, prepare and plan to manage: after rising	1.12	1.01	0.94	0.678 (0.58–0.77)
Item 22	Trigger management: defecation	0.97	0.88	0.70	0.640 (0.54–0.73)
Item 23	Trigger management: eating meals	0.94	0.85	0.75	0.681 (0.59–0.76)
Item 24	Trigger management: free or bored	0.99	0.91	0.70	0.636 (0.54–0.72)
Item 25	Trigger management: seeing others smoke	0.92	0.88	0.71	0.645 (0.54–0.73)
Item 26	Trigger management: offered smoke	0.84	0.81	0.64	0.655 (0.56–0.74)
Item 27	Trigger management: intense work	0.89	0.73	0.66	0.638 (0.53–0.74)
Item 28	Trigger management: anxious	0.83	0.73	0.56	0.610 (0.50–0.71)
Item 29	Provide information on withdrawal symptoms: craving	0.63	0.55	0.51	0.812 (0.75–0.87)
Item 30	Provide information on withdrawal symptoms: restlessness	0.58	0.51	0.42	0.776 (0.70–0.85)
Item 31	Provide information on withdrawal symptoms: headache	0.61	0.57	0.47	0.810 (0.74–0.87)
Item 32	Provide information on withdrawal symptoms: insomnia	0.60	0.51	0.41	0.760 (0.67–0.84)
Item 33	Provide information on withdrawal symptoms: indigestion	0.62	0.62	0.45	0.740 (0.66–0.82)
Item 34	Provide information on withdrawal symptoms: anorexia and constipation	0.60	0.56	0.49	0.749 (0.67–0.83)
Item 35	Provide information on withdrawal symptoms: cough	0.55	0.51	0.45	0.720 (0.64–0.80)
Item 36	Provide information on withdrawal symptoms: weight gain	0.50	0.44	0.40	0.802 (0.73–0.87)
Item 37	Offer information leaflet	0.44	0.43	0.39	0.879 (0.76–0.97)
Quality (interaction-based items)					

(Continues)

Table 1. (Continued)

Item number	Item content	Coder 1 score mean	Coder 2 score mean	Coder 3 score mean	Krippendorff's α (95% CI)
Item 38	Build general rapport	0.50	0.37	0.40	0.632 (0.52–0.74)
Item 39	Explain expectations regarding treatment programme	0.05	0.02	0.03	0.052 (–0.60–0.64)
Item 40	Provide reassurance	0.25	0.18	0.16	0.481 (0.31–0.66)
Item 41	Communication approaches: elicit and answer questions	0.36	0.43	0.39	0.741 (0.65–0.82)
Item 42	Communication approaches: use reflective listening	0.34	0.23	0.23	0.737 (0.63–0.84)
Item 43	Communication approaches: summarize information and confirm client decisions	0.48	0.34	0.38	0.679 (0.57–0.78)
Item 44	Tailor interactions appropriately	0.18	0.21	0.12	0.729 (0.58–0.85)
Item 45	Emphasize choice	0.47	0.51	0.42	0.692 (0.60–0.78)

Mean Krippendorff's α for adherence to content was 0.80 and for quality of interaction was 0.66. Items in bold type show low agreement ($\alpha < 0.67$). The response option anchors for each item were 0 = not implemented, 1 = partially implemented and 2 = fully implemented except for items 10, 12, 16–20 and 37, where the response option anchors were 0 = not implemented and 2 = implemented. CI = confidence interval.

Table 2 Average fidelity scores and proportion quit (from ASSIST study) by providers.

Provider ID	Number of patients	Fidelity scores		Standardized fidelity scores		Number of patients	Proportion quit
		Adherence: content-based items Mean (95% CI)	Quality: interaction-based items Mean (95% CI)	Adherence: content-based items Mean (SD)	Quality: interaction-based items Mean (SD)		
24	6	11.17 (8.62–13.71)	0.04 (–0.04–0.13)	0.31 (0.19)	0 (0)	61	0.26
26	5	13.05 (11.18–14.92)	0.10 (–0.04–0.24)	0.36 (0.07)	0 (0)	57	0.72
4	9	14.83 (12.87–16.79)	0	0.39 (0.16)	0 (0)	64	0.41
27	10	21.33 (17.96–24.69)	0.28 (0.08–0.47)	0.41 (0.11)	0 (0)	54	0.35
7	9	17.06 (15.26–18.85)	0	0.43 (0.14)	0 (0)	60	0.20
22	10	24.85 (20.72–28.98)	0.75 (–0.05–1.55)	0.62 (0.22)	0.03 (0.08)	55	0.40
25	10	31.85 (29.80–33.90)	1.53 (0.93–2.12)	0.82 (0.17)	0.03 (0.05)	55	0.62
11	10	34.13 (29.20–39.05)	2.13 (1.36–2.89)	0.93 (0.43)	0.26 (0.27)	60	0.57
28	9	34.19 (31.88–36.51)	0.67 (0.45–0.88)	0.94 (0.19)	0.08 (0.06)	55	0.67
21	8	35.16 (30.12–40.20)	0.81 (0.48–1.15)	0.96 (0.38)	0.09 (0.11)	60	0.55
5	9	37.86 (33.74–41.98)	2.75 (2.38–3.12)	1.04 (0.36)	0.33 (0.11)	60	0.52
19	8	43.06 (40.36–45.77)	11.06 (10.00–12.13)	1.14 (0.18)	1.38 (0.42)	62	0.55
3	10	43.85 (41.40–46.30)	7.98 (7.39–8.56)	1.19 (0.22)	1.00 (0.22)	60	0.55
23	8	48.56 (42.64–54.48)	3.19 (2.22–4.15)	1.20 (0.41)	0.23 (0.10)	57	0.54
2	9	47.50 (43.79–51.21)	0.72 (0.14–1.31)	1.30 (0.28)	0.07 (0.17)	60	0.32
20	10	54.02 (52.11–56.29)	2.50 (1.57–3.43)	1.48 (0.15)	0.31 (0.36)	60	0.07
9	10	59.23 (56.91–61.54)	2.33 (1.53–3.12)	1.61 (0.18)	0.29 (0.29)	59	0.46
8	4	62.94 (59.51–66.36)	4.13 (2.55–5.70)	1.72 (0.16)	0.50 (0.35)	59	0.19

ASSIST = Action to Stop Smoking In Suspected Tuberculosis; SD = standard deviation; 95% CI = 95% confidence interval.

standardized mean scores for each fidelity domain by provider are shown in Table 2.

PCA

Fifty-two per cent of the variance for the full index was explained by a single component. Items 10, 12, 15, 16, 37, 39 and 40 did not load well (threshold < 0.4) across the

coders. When analysing the items from the two fidelity domains separately, a single component explained 60 and 91% of the variance, respectively. Content-based items (10, 12, 15, 16 and 37) showed loadings below the threshold of 0.4; the interaction-based item (39) also did not load well (Supporting information, Tables S9, S10 and S11).

For adherence domain (Table 3), four components showed eigenvalues > 1 (Supporting information, Table

S12) in addition to the first component that explained 60% of variance. These components roughly map to the five aspects of the BS intervention: assessments of dependency, providing information about the consequences of smoking and cessation, preparing for the quit attempt, managing triggers to smoke and managing nicotine withdrawals. Items 1 and 37 did not load well. Overall, three items (10, 12 and 15) did not load well in spite of being delivered.

G-study

The *g*-coefficients were 0.973 for adherence and 0.974 for quality, showing that the providers, after accounting for other sources of variation (Supporting information, Appendix S2), contributed approximately 97% of the variation in BS delivery. The crude ICC for provider differences in adherence scores was 0.72 and in quality scores was 0.71, indicating that providers contributed approximately 70% of the variation in intervention delivery when other sources were not adjusted. The inclusion of district in the model as a fixed-effect neither correlated significantly with the adherence scores ($\beta = -7.21$, $F_{(df = 1, 272)} = 0.94$, $P = 0.33$) nor the quality scores ($\beta = -0.29$, $F_{(df = 1, 272)} = 0.04$, $P = 0.84$).

The findings of the psychometric analysis do not suggest removing items from the fidelity index, as the core problem was largely lack of variance in the sample. Overall, fidelity and item scores showed variability across providers, indicating the applicability of the coding system (i.e. few ceiling or floor effects; services not bunched together clumping in one limited score range). Krippendorff's alpha and *g*-coefficients estimates confirmed that the coding structure of the index items reliably captured variation in providers intervention delivery.

Study 2. Association of fidelity with outcomes

Relative to content-based items, interaction-based items of the BS intervention were poorly implemented (fully implemented in < 10% and partially in 10–20% of the sessions) (Fig. 2).

The degree of similarity in the providers' practice between the current study and ASSIST was found to be moderate for management of patient triggers (Kendall's *W*: 0.69, $P = 0.06$) and high for withdrawal symptoms (Kendall's *W*: 0.75, $P = 0.03$).

The binomial regression (model 1, Table 4) indicates that an increase in the quality scores of the provider led to a 50% increase in the quit rates ($P < 0.05$). An increase in the adherence scores of the provider led to a 24% decrease in the quit rates ($P = 0.055$). When including both variables in the same model (model 2, Table 4), the association of adherence scores with quit rates became statistically significant, although remaining negative, and the association of quality scores with quit rates became

stronger (model 2, Table 4). A negative interaction was found between adherence and quality scores ($P < 0.0001$) (model 3, Table 4).

DISCUSSION

The BCT taxonomy for smoking cessation was used to code the activities of the ASSIST 5As intervention into content- and interaction-based items. The derived index was applied to BS sessions recorded in chest clinics in Pakistan to measure fidelity and explore reliability. Comprising 45 items, 37 on adherence (content-based) and eight on quality (interaction-based), the fidelity index was found reliable for coding BS despite varying coder skills. Items on setting a quit date, offering an information leaflet, explaining treatment expectations and providing reassurance seemed valid, but were not fully assessed due to low variability across the sample. Linking average fidelity scores for providers with successful quit provided useful insights into the delivery of BS for smoking cessation. Quantifying BCTs for the quality domain that required tailoring to patient needs, using an anchored scale, might be an over-simplification of a somewhat complex process. Most frequently delivered BCTs in the study included information-gathering and addressing motivation (e.g. providing information about the harms of smoking and assessments of nicotine dependence). Least implemented were maximizing self-regulatory capacity, communication and delivery of intervention (e.g. management of triggers and withdrawals, building rapport, providing reassurance), as reported by the fidelity study in England's Stop Smoking Services [7]. Some of the least implemented BCTs in our study, such as emphasizing choice, confirming client decisions, advising on changing routine and facilitating relapse prevention, might have been difficult to capture, being more subjective than the content-based activities, or it might reflect provider competence, motivation or patient preferences. These phenomena need to be explored in future research.

Given the interpersonal nature of counselling, it is unsurprising that BS effectiveness was strongly influenced by interaction with providers [24]. A recent meta-analysis of intervention fidelity for motivational interviewing [46] found that interventions with higher implementation of interpersonal BCTs produced larger effects on behaviour change. The association of successful quit with better quality scores could also reflect a non-specific effect due to higher patient satisfaction. The finding that adherence was negatively associated with successful quit, even after adjusting for the quality, appears counter-intuitive [5]. However, complete adherence to intervention ingredients as a desirable goal is contested [7,16]. High and moderate quit rates showed similar levels of adherence across a sample of English Stop Smoking Services [7]. Furthermore, the

Table 3 Identifying the principal components for adherence domain.

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
Item 1	-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
Item 2	-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	0.02	0.09	0.1	0.02	0.05	-0.06	0.07
Item 3	-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	-0.07	-0.08	-0.07	0.11	0.17	0.07	0.16
Item 4	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
Item 5	-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	-0.04	-0.02	-0.01	-0.04	-0.01	0	-0.06
Item 6	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
Item 7	-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
Item 8	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
Item 9	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
Item 10	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
Item 11	-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
Item 12	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
Item 13	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
Item 14	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
Item 15	-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
Item 16	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
Item 17	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
Item 18	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
Item 19	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
Item 20	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
Item 21	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
Item 22	-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
Item 23	-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
Item 24	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
Item 25	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
Item 26	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
Item 27	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
Item 28	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
Item 29	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
Item 30	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
Item 31	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
Item 32	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
Item 33	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

(Continues)

Table 3. (Continued)

Items	Loading 1				Loading 2				Loading 3				Loading 4				Loading 5			
	C1	C2	C3	C4	C1	C2	C3	C4	C1	C2	C3	C4	C1	C2	C3	C4	C1	C2	C3	C4
Item 34	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
Item 35	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
Item 36	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
Item 37	-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; Shade codes: lightest grey for component 1 to darkest grey for component 5. Loading values ≥ 0.4 were considered meaningful.

service with lower adherence (56%) was potentially more effective in achieving higher patient quit compared to services with higher adherence (69%). Delivering complex interventions possibly requires providers to manage too many dimensions at the same time, making the task longer and tedious, and so less effective if fully implemented [47,48]. More simultaneously delivered behaviour change elements could also potentially make it complex for the patients to understand and engage in. We should consider simplifying the content of smoking cessation interventions and shift focus to the interpersonal aspects of the patient-provider interaction as 'more is sometimes less' in health behaviour change [49].

A fidelity index offers numerous applications for smoking cessation practice. Compared to practice manuals, it can make the content of time-limited interventions more focused if used as a self-recording checklist by the providers [7]. It could help to standardize approaches to reduce wide variations in implementation of active ingredients and reduce failure to achieve the original effectiveness of the interventions [50]. In addition to the content-based items, using the quality domain on interaction-based items, the index can facilitate identification of specific skills and competence, provider training needs. Local practices could link implemented ingredients with quit outcomes and recommend adaptations to the services for their respective context by studying trends over time. The development of this fidelity index focused on provider practice and attitude in delivering BS, relying heavily on BCT taxonomy. This may have missed the wider psychological processes of behaviour change not covered by BCT taxonomies. Incorporation of additional interaction-based items in the index, such as the relational techniques used in motivational interviewing [24], might improve the operational value of the fidelity index. There is potential for developing shorter versions of the fidelity index and for optimized intervention delivery if the active ingredients can independently predict quit outcomes. Studies using multi-phase optimization strategy and sequential multiple-assignment randomized trials could move the field forward in providing such mechanistic evidence [51–53].

The study had several limitations. The use of audiotaping meant that important aspects of non-verbal interaction will have been missed. Capturing both verbal and non-verbal interaction using conversation analysis [54] of videotaped BS sessions should be considered. Nevertheless, in this study, direct observation of patient-provider sessions via audio recordings was considered less intrusive than video for capturing the key elements of the intervention delivery process [7,40].

The retrospective linking of provider fidelity scores with patient quit rates from the past raises questions about causality. The practice scores, used to test the assumption that providers' BS delivery practice was similar

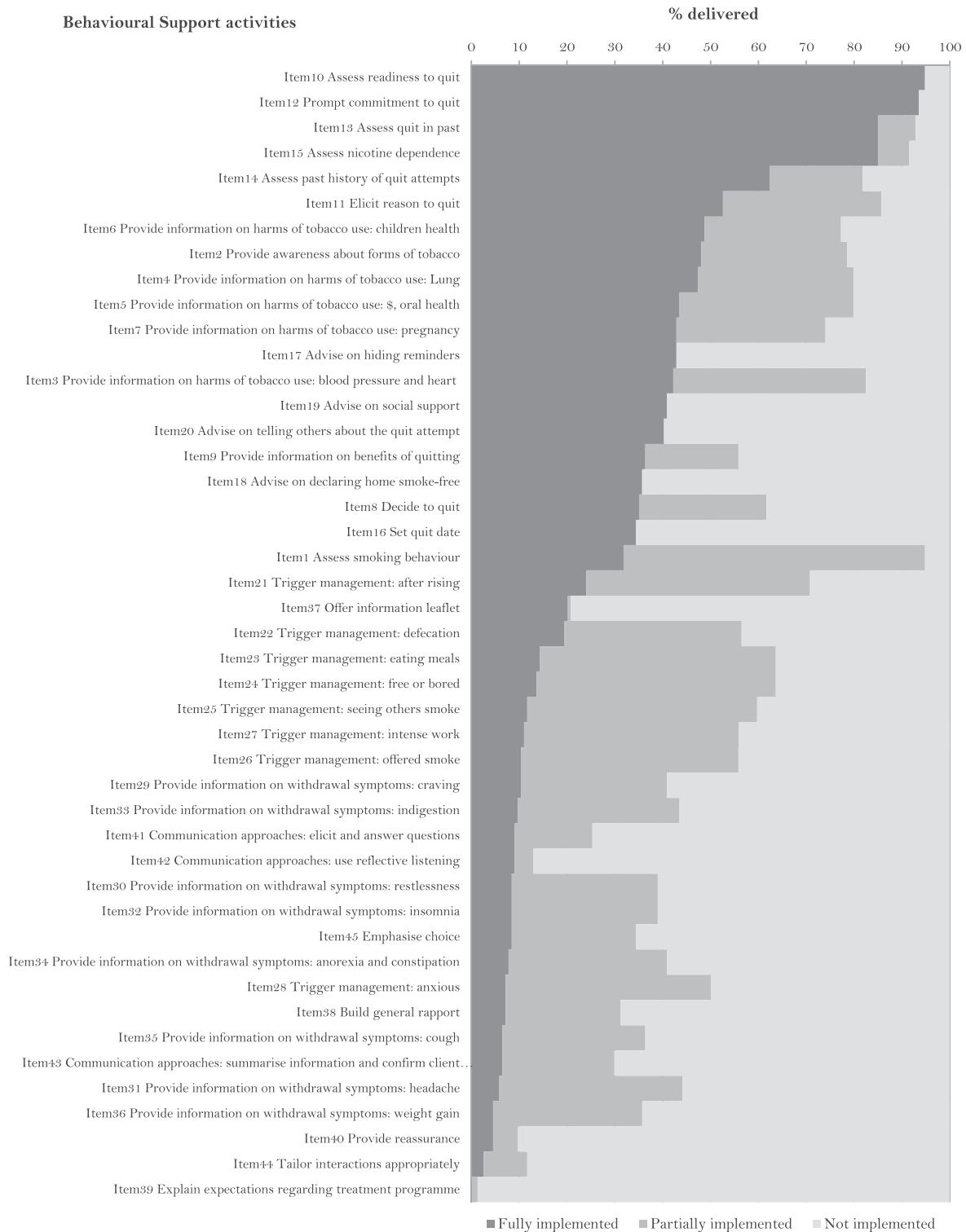


Figure 2 Smoking cessation behavioural support intervention fidelity

in the gap between the two studies, were self-recorded and might be subject to reporting bias. Such retrospective linkage can introduce bias in the study estimates, especially when there is state dependence in respondent's choices [55]. True state dependence, where behaviour relevant to future choices is altered because of

experiencing an event, is unlikely in the case of providers completing the self-record checklist in this study, as they did not receive any behaviour change training between the two studies.

The study had to consider data at the provider- rather than patient-level, thus reducing the statistical power of

Table 4 Association of fidelity scores with patient quit rates by providers ($n = 18$).

Independent variable	Successful quit at 6 months risk ratio (95% CI)	P-value (Wald's χ^2)
Model 1: univariable effects		
Adherence to content	0.76 (0.57–1.01)	0.0552
Quality of interaction	1.50 (1.06–2.12)	0.0226
Model 2: multivariable (adjusted effects—when both adherence and interaction are in the model)		
Adherence to content	0.55 (0.40–0.77)	0.0005
Quality of interaction	2.15 (1.43–3.24)	0.0002
Model 3: model 2 + interaction term		
Adherence to content	0.20 (0.12–0.34)	<0.0001
Quality of interaction	8.15 (4.39–15.15)	<0.0001
Adherence \times interaction	0.0007 (0.0001–0.0083)	<0.0001

the analysis. Although the implementation pattern of BCTs indicated that the delivery of smoking cessation BS in this study was similar to that of stop-smoking advisers in the United Kingdom, generalizability of the findings from this study is still limited. The fidelity index represented a specific smoking cessation intervention delivered in chest clinics of Pakistan. Replication of the methods used in this study and application of different versions of the fidelity index in diverse settings is required.

CONCLUSION

The fidelity index is a reliable measure of the delivery of behaviour change features of a smoking cessation intervention in clinical practice. It provides a method to score intervention fidelity for individual ingredients, enabling its use as an intermediate variable to explain the intervention–outcome relationship. Recommended revisions of the fidelity index include incorporation of additional interaction-based items, such as the relational techniques used in motivational interviewing.

Author contributions

O.D. conceptualized the research idea, conducted the study, performed statistical analyses and drafted the manuscript [56]. K.S. and T.A.S. contributed to conceptualizing and refining the research idea, interpretation of statistical analysis and discussion of findings. J.R.B. contributed to research design of the psychometric elements of the study, selection of statistical tests, interpretation and discussion of findings. E.L. provided advice for fidelity coding and support for training of the coders. All authors provided critical revisions and approved the final manuscript.

Declaration of interests

K.S. received a research grant from Pfizer (2015–17) to study the effect of varenicline on waterpipe smoking cessation.

Acknowledgements

The authors thank Dr Riaz Ahmad and Iqbal Mahmood for facilitating data collection in the clinics in Sargodha and Jhang and health workers that participated in the study, Dr Amina Khan and Qadeer Ashraf for fidelity coding using the audiotapes and Professor Robert West for providing critical feedback on the thesis as the examiner. Bupa foundation prize money for the ASSIST trial funded the field activities of the study.

References

- Craig P., Dieppe P., Macintyre S., Michie S., Nazareth I., Petticrew M. Developing and evaluating complex interventions: the new Medical Research Council guidance. *Int J Nurs Stud* 2013; **50**: 587–92.
- Dogar O., Bullen C., Novotny T. E., Siddiqi K. Smoking cessation and respiratory disease in low-income and middle-income countries. *Lancet Respir Med* 2013; **1**: e23–e24.
- Siddiqi K. Supporting economically disadvantaged tobacco users to help them stop: time to review priorities? *Addiction* 2014; **109**: 1221–2.
- Grimshaw J. M., Shirran L., Thomas R., Mowatt G., Fraser C., Bero L. *et al.* Changing provider behavior: an overview of systematic reviews of interventions. *Med Care* 2001; **39**: II2–II45.
- Durlak J. A., DuPre E. P. Implementation matters: a review of research on the influence of implementation on program outcomes and the factors affecting implementation. *Am J Community Psychol* 2008; **41**: 327–50.
- Tober G., Clyne W., Finnegan O., Farrin A., Team I. R., Team U. R. Validation of a scale for rating the delivery of psycho-social treatments for alcohol dependence and misuse: the UKATT process rating scale (PRS). *Alcohol Alcohol* 2008; **43**: 675–82.
- Lorenatto F., West R., Christopherson C., Michie S. Assessing fidelity of delivery of smoking cessation behavioural support in practice. *Implement Sci* 2013; **8**: 40.
- Lorenatto F., West R., Bruguera C., Michie S. A method for assessing fidelity of delivery of telephone behavioral support for smoking cessation. *J Consult Clin Psychol* 2014; **82**: 482–91.
- Lennox A. S., Bain N., Taylor R. J., McKie L., Donnan P. T., Groves J. 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 Educ J* 1998; **57**: 140–9.
10. Hawe P., Shiell A., Riley T. Complex interventions: how 'out of control' can a randomised controlled trial be? *BMJ* 2004; **328**: 1561–3.
 11. Papadakis S., Cole A. G., Reid R. D., Coja M., Aitken D., Mullen K.-A. *et al.* Increasing rates of tobacco treatment delivery in primary care practice: evaluation of the Ottawa model for smoking cessation. *Ann Fam Med* 2016; **14**: 235–43.
 12. Siddiqi K., Khan A., Ahmad M., Dogar O., Kanaan M., Newell J. *et al.* Action to stop smoking in suspected tuberculosis (ASSIST) in Pakistan: a cluster randomized, controlled trial. *Ann Intern Med* 2013; **158**: 667–75.
 13. Moore G., Audrey S., Barker M., Bond L., Bonell C., Hardeman W. *et al.* Process evaluation of complex interventions. *Complex Intervent Health* 2015; <https://doi.org/10.1136/bmj.h1258>.
 14. Dane A. V., Schneider B. H. Program integrity in primary and early secondary prevention: are implementation effects out of control? *Clin Psychol Rev* 1998; **18**: 23–45.
 15. Dogar O., Elsej H., Khanal S., Siddiqi K. Challenges of integrating tobacco cessation interventions in TB programs: case studies from Nepal and Pakistan. *J Smok Cessat* 2016; **11**: 108–15.
 16. Borrelli B. The assessment, monitoring, and enhancement of treatment fidelity in public health clinical trials. *J Public Health Dent* 2011; **71**: S52–S63.
 17. Peters D. H., Tran N. T., Adam T. Implementation Research in Health: a Practical Guide. Contract no.: 11 September 2015. World Health Organization Alliance for Health Policy and Systems Research; 2013.
 18. Carroll C., Patterson M., Wood S., Booth A., Rick J., Balain S. A conceptual framework for implementation fidelity. *Implement Sci* 2007; **2**: 40.
 19. Bond G. R., Becker D. R., Drake R. E., Vogler K. M. A fidelity scale for the individual placement and support model of supported employment. *Rehab Counsel Bull* 1997; **40**: 265–84.
 20. Drake R. E., Goldman H. H., Leff H. S., Lehman A. F., Dixon L., Mueser K. T. *et al.* Implementing evidence-based practices in routine mental health service settings. *Psychiatr Serv* 2001; **52**: 179–82.
 21. Mowbray C. T., Holter M. C., Teague G. B., Bybee D. Fidelity criteria: development, measurement, and validation. *Am J Eval* 2003; **24**: 315–40.
 22. Hardeman W., Lamming L., Kellar I., De Simoni A., Gaffry J., Boase S. *et al.* Implementation of a nurse-led behaviour change intervention to support medication taking in type 2 diabetes: beyond hypothesised active ingredients (SAMS consultation study). *Implement Sci* 2014; **9**: 70.
 23. Toomey E., Matthews J., Hurley D. A. Using mixed methods to assess fidelity of delivery and its influencing factors in a complex self-management intervention for people with osteoarthritis and low back pain. *BMJ Open* 2017; **7**: e015452.
 24. Hardcastle S. J., Fortier M., Blake N., Hagger M. S. Identifying content-based and relational techniques to change behavior in motivational interviewing. *Health Psychol Rev* 2011; **11**: 1–16.
 25. Hagger M. S., Hardcastle S. J. Interpersonal style should be included in taxonomies of behavior change techniques. *Front Psychol* 2014; **5**: 254.
 26. Taylor A. H., Thompson T. P., Greaves C. J., Taylor R. S., Green C., Warren F. C. *et al.* A pilot randomised trial to assess the methods and procedures for evaluating the clinical effectiveness and cost-effectiveness of exercise assisted reduction then stop (EARS) among disadvantaged smokers. *Health Technol Assess* 2014; **18**: 1–324.
 27. Lorencatto E., West R., Bruguera C., Brose L. S., Michie S. Assessing the quality of goal setting in behavioural support for smoking cessation and its association with outcomes. *Ann Behav Med* 2016; **50**: 310–8.
 28. Gainforth H. L., Lorencatto E., Erickson K., West R., Michie S. Characterizing clients' verbal statements in behavioural support interventions: the case of smoking cessation. *Br J Health Psychol* 2016; **21**: 600–12.
 29. Keogh A., Matthews J., Hurley D. A. An assessment of physiotherapist's delivery of behaviour change techniques within the SOLAS feasibility trial. *Br J Health Psychol* 2018; **23**: 908–32.
 30. West R., Walia A., Hyder N., Shahab L., Michie S. Behavior change techniques used by the English stop smoking services and their associations with short-term quit outcomes. *Nicotine Tob Res* 2010; **12**: 742–7.
 31. Bond G. R., Williams J., Evans L., Salyers M., Kim H., Sharpe H. *PN-44-Psychiatric Rehabilitation Fidelity Toolkit*. Cambridge, MA: Human Services Research Institute; 2000.
 32. Streiner D. L., Norman G. R. *Health Measurement Scales: A Practical Guide to Their Development and Use, 4th edn*. Oxford: Oxford University Press; 2008.
 33. Bellg A. J., Borrelli B., Resnick B., Hecht J., Minicucci D. S., Ory M. *et al.* Enhancing treatment fidelity in health behavior change studies: best practices and recommendations from the NIH behavior change consortium. *Health Psychol Hillsdale* 2004; **23**: 443–51.
 34. World Health Organization. Toolkit for delivering the 5A's and 5R's brief tobacco interventions in primary care 2014. Available at: http://www.who.int/tobacco/publications/smoking_cessation/9789241506953/en/ (Archived at <http://www.webcitation.org/78Gg9PRca> on 10 May 2019) (accessed on March 04, 2015)
 35. Michie S., Hyder N., Walia A., West R. Development of a taxonomy of behaviour change techniques used in individual behavioural support for smoking cessation. *Addict Behav* 2011; **36**: 315–9.
 36. Nelson M. C., Cordray D. S., Hulleman C. S., Darrow C. L., Sommer E. C. A procedure for assessing intervention fidelity in experiments testing educational and behavioral interventions. *J Behav Health Serv Res* 2012; **39**: 374–96.
 37. Gearing R. E., El-Bassel N., Ghesquiere A., Baldwin S., Gillies J., Ngeow E. Major ingredients of fidelity: a review and scientific guide to improving quality of intervention research implementation. *Clin Psychol Rev* 2011; **31**: 79–88.
 38. Clarke G. Intervention fidelity in the psychosocial prevention and treatment of adolescent depression. *J Prev Interv Community* 1998; **17**: 19–33.
 39. Zapf A., Castell S., Morawietz L., Karch A. Measuring inter-rater reliability for nominal data—which coefficients and confidence intervals are appropriate? *BMC Med Res Methodol* 2016; **16**: 93.
 40. Lorencatto E., West R., Seymour N., Michie S. Developing a method for specifying the components of behavior change interventions in practice: the example of smoking cessation. *J Consult Clin Psychol* 2013; **81**: 528–44.
 41. Brennan R. L. *Generalizability Theory*. New York: Springer; 2001.
 42. Hayes A. F., Krippendorff K. Answering the call for a standard reliability measure for coding data. *Commun Methods Meas* 2007; **1**: 77–89.

43. Krippendorff K. *Content Analysis: An Introduction to its Methodology*. Thousand Oaks, CA: Sage; 2004.
44. Motallebzadeh R., Bland J. M., Markus H. S., Kaski J. C., Jahangiri M. Neurocognitive function and cerebral emboli: randomized study of on-pump versus off-pump coronary artery bypass surgery. *Ann Thorac Surg* 2007; **83**: 475–82.
45. West R., Hajek P., Stead L., Stapleton J. Outcome criteria in smoking cessation trials; proposal for a common standard. *Addiction* 2005; **100**: 299–303.
46. O'Halloran P. D., Blackstock F., Shields N., Holland A., Iles R., Kingsley M. *et al*. Motivational interviewing to increase physical activity in people with chronic health conditions: a systematic review and meta-analysis. *Clin Rehabil* 2014; **28**: 1159–71.
47. Boutron I., Moher D., Altman D. G., Schulz K. F., Ravaud P. Extending the CONSORT statement to randomized trials of nonpharmacologic treatment: explanation and elaboration. *Ann Intern Med* 2008; **148**: 295–309.
48. Sereno A. B., Soares E. C. C., Lapa e Silva J. R., Nápoles A. M., Bialous S. A., da Costa e Silva V. L. *et al*. 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* 2012; **32**: 451–6.
49. Wilson K., Senay I., Durantini M., Sánchez F., Hennessy M., Spring B. *et al*. 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. *Psychol Bull* 2015; **141**: 474–509.
50. Bumbarger B., Perkins D. After randomised trials: issues related to dissemination of evidence-based interventions. *J Child Serv* 2008; **3**: 55–64.
51. Piper M. E., Fiore M. C., Smith S. S., Fraser D., Bolt D. M., Collins L. M. *et al*. Identifying effective intervention components for smoking cessation: a factorial screening experiment. *Addiction* 2016; **111**: 129–41.
52. Collins L. M., Murphy S. A., Strehler V. The multiphase optimization strategy (MOST) and the sequential multiple assignment randomized trial (SMART): new methods for more potent eHealth interventions. *Am J Prev Med* 2007; **32**: S112–S118.
53. Howard M. C., Jacobs R. R. The multiphase optimization strategy (MOST) and the sequential multiple assignment randomized trial (SMART): two novel evaluation methods for developing optimal training programs. *J Organ Behav* 2016; **37**: 1246–70.
54. Hutchby I., Wooffitt R. *Conversation analysis*. Polity; 2008.
55. Shachar R., Eckstein Z. Correcting for bias in retrospective data. *J Appl Economet* 2007; **22**: 657–75.
56. Dogar O. Understanding variations in the outcome of a smoking cessation programme in tuberculosis patients in Pakistan—the role of fidelity? White Rose eThesis. University of York; 2017. Available at: <http://etheses.whiterose.ac.uk/id/eprint/15900>. (Archived at <http://www.webcitation.org/78GD9vuk> on 10 May 2019) (accessed on May 10, 2019).

Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Figure S1 Logic model for measuring intervention fidelity.

Figure S2 Concept of compositional features versus functional features to form items and response options of the fidelity index.

Table S1 Behavioural Support intervention provided in ASSIST study.

Table S2 List of content-based ingredients identified from Behavioural Support for smoking cessation.

Table S3 Specifying Behavioural Support components using Behaviour Change Techniques taxonomy.

Table S4 Behavioural anchors for the response scale.

Table S5 SAS syntax.

Table S6 Participant characteristics in Fidelity study and ASSIST study.

Table S7 Fidelity Index.

Table S8 Descriptive statistics for item scores.

Table S9 Item loadings for full fidelity index (all 45 items).

Table S10 Item loadings of the 37 items for Adherence.

Table S11 Item loadings of the 8 items for Quality.

Table S12 Variance explained by each Principal Component for Adherence.

Table A1.1 Experts rating on 'importance' of items for measuring fidelity.

Table A2.1 Identifying important sources of variation.

Table A2.2 Variance partitioning.

Table A3.1 Self-record checklist of the providers implemented in both studies.