A method for assessing fidelity of delivery of telephone behavioural support for smoking cessation

Fabiana Lorencatto, [1,4] Robert West,[3,4] Carla Bruguera[31] & Susan Michie. [2,4]

[1] Division of Health Services Research & Management, School of Health Sciences, City University London.

[2] Centre for Outcomes Research and Effectiveness, Dept. Clinical Educational & Health Psychology, University College London.

[3] CRUK Health Behaviour Research Centre, Dept. Epidemiology & Public Health, University College London.

[4] National Centre for Smoking Cessation & Training, University College London.

Author note: Correspondence concerning this article should be addressed to Fabiana Lorencatto Division of Health Services Research & Management, School of Health Sciences, City University London, EC1V 0HB. Email: Fabiana.Lorencatto.2@city.ac.uk .

Disclosure of competing interests: Robert West has undertaken research and consultancy for companies that develop and manufacture smoking cessation medications. Robert West has a share of a patent in a novel nicotine delivery device. Fabiana Lorencatto, Susan Michie, and Carla Bruguera declare no competing interests.

ABSTRACT

Objectives: Behavioural support for smoking cessation is delivered through different modalities, often guided by treatment manuals. Recently developed methods for assessing fidelity of delivery have shown that face-to-face behavioural support is often not delivered as specified in the service treatment manual. This study aimed to extend this method to evaluate fidelity of telephone-delivered behavioural support.

Methods: A treatment manual and transcripts of 75 audio-recorded behavioural support sessions were obtained from the UK's national quitline service and coded into component behaviour change techniques (BCTs) using a taxonomy of 45 smoking cessation BCTs. Inter-rater reliability was assessed using percentage agreement. Fidelity was assessed by comparing the number of BCTs identified in the manual with those delivered in telephone sessions by four counselors. Fidelity was assessed according to session type, duration, counselor, and BCT. Differences between self-reported and actual BCT use were examined.

Results: Average coding reliability was high (81%). On average, 41.8% of manual-specified BCTs were delivered per session (SD16.2), with fidelity varying by counselor from 32-49%. Fidelity was highest in pre-quit sessions (46%), and for BCT 'give options for additional support' (95%). Fidelity was lowest for quit-day sessions (35%) and BCT 'set graded tasks' (0%). Session duration was positively correlated with fidelity (r=0.585; p<0.01). Significantly fewer BCTs were used than were reported as being used [t=-5.52 (15), p<0.001].

Conclusions: The content of telephone-delivered behavioural support can be reliably coded in terms of BCTs. This can be used assess fidelity to treatment manuals and in turn identify training

needs. The observed low fidelity underlines the need to establish routine procedures for monitoring delivery of behavioural support.

Key words: Behaviour change interventions, smoking cessation, delivery, fidelity

INTRODUCTION

Approximately 70% of adult smokers would like to quit smoking (Orleans, 2007; Thyrian, West, John, 2007). There are numerous aids available to assist smokers in quitting, including behavioural support, which involves advice, discussion and targeted activities aimed at maximising a smoker's motivation to quit and facilitating relapse prevention and coping (West & Stapleton, 2008). Behavioural support has been delivered through various modalities with demonstrated effectiveness, including: face-to-face individual and group support sessions, internet- and telephone-based support (Lancaster & Stead, 2005; Stead & Lancaster, 2005; Shahab & McEwen, 2009; Stead, Perera, & Lancaster 2009).

There has recently been an increased investment in delivering behavioural support interventions as part of wider clinical practice. For example, in England, medications and free, weekly behavioural support sessions are available via a network of 152 National Health Service (NHS) Stop Smoking Services. However, the outcomes across these services are extremely heterogeneous: between April 2011 and March 2012 four-week carbon monoxide (CO) validated quit outcomes ranged from 2% to 78% (NHS Information Centre, 2012). One potential service-level factor contributing to this variability in outcome may be differences in the content of support delivered across services.

However, behavioural support interventions are complex, comprising multiple, potentially interacting behaviour change techniques (BCTs) (Michie, Fixsen, Eccles, et al. 2009). Such complexity therefore often renders the process of clearly identifying, characterising and comparing the content of behavioural support interventions difficult. Consequently, the translation of complex interventions from trial to service settings is rarely uniform; meaning that

interventions with demonstrated effectiveness can sometimes fail when implemented in practice on a wide scale or in new settings (Borrelli, Sepinwall, Bellg et al. 2005). In the UK, national guidelines outlining the recommended content and format of behavioural support interventions have been published as part of an effort to standardise the content of support delivered across services and to promote evidence-based practice (Croghan, 2011; West, McNeil, & Raw, 2000). Methods are needed to support the implementation of these recommendations into routine practice (Eccles, Armstrong, Baker et al. 2009).

Treatment manuals represent one potential vehicle by which clinical findings and guidelines may be translated into practice, and a tool for standardising the content of support provided within services. Manuals are widely used in the delivery of smoking cessation behavioural support both in research (e.g. Zhu, 1996; Curry, 1995; Lawrence, 2006; Zhu et al. 2002) and in clinical practice (West, Walia, Hyder, Shahab, & Michie, 2010). However, there is evidence in the English NHS Stop Smoking Services that counselors delivering support according to the same treatment manual have widely differing success rates (Brose, McEwen, & West, *2013*). This raises the question as to whether counselors are adhering to manual specifications when delivering support.

Fidelity of delivery refers to the extent to which core intervention components are delivered as intended (Borrelli, 2011). This is distinguished from *how* components are delivered, such as quality of delivery (Borrelli, 2011). Knowledge of fidelity of delivery can contribute to the accurate interpretation of intervention outcomes, enhance theoretical understanding of how interventions work, and aid identification of training needs or aspects of intervention delivery requiring improvement (Borrelli, Sepinwall, Bellg et al. 2005). Although the need to assess fidelity has been emphasised in the CONSORT statement (Boutron, Moher, Altman, et al. 2008),

systematic reviews indicate that fidelity is rarely assessed,reported or accounted for in research (Moncher & Prinz, 1991; Dane & Schneider, 1998; Dusenbury et al. 2003). Where fidelity has been assessed for behaviour change interventions in other domains, including physical activity (Hardeman, Michie, Fanshawe, et al. 2008), excessive alcohol use (Tober, Clyne, Finnegan et al. 2009), and medication adherence (Dewing et al., 2013), it has often been shown that fidelity is inconsistent and poor, with typically less than 55% of manual-specified content delivered during intervention sessions. However to date, such fidelity assessments have typically been conducted as part of a process evaluation for research trials (Grant, Treweek, Dreischulte, Foy, & Guthrie, 2013). There are few examples of established, routine, procedures for monitoring fidelity of behaviour change interventions as delivered in the context of actual clinical practice (Moncher & Prinz, 1991; Dane & Schneider, 1998; Dusenbury et al. 2003).

The recommended 'gold standard' strategy for assessing fidelity of delivery involves objectively verifying delivery by comparing the content of recorded intervention sessions to a pre-specified criterion such as a treatment manual (Borrelli, 2011). A pre-requisite to applying such a method is the ability to clearly specify the content of both intervention sessions and treatment manuals. A taxonomy-based method for objectively verifying fidelity of delivery of smoking cessation behavioural support interventions has recently been developed (Lorencatto, West, Christopherson, Michie, 2013). This involves applying a reliable taxonomy of 45 smoking cessation BCTs to first specify and subsequently compare content of service treatment manuals and audio-recorded one-to-onebehavioural support sessions delivered (Michie, Hyder, Walia, & West, 2010; Lorencatto, West, Seymour, & Michie, *In press*). This method has been applied to assess fidelity in two English NHS Stop Smoking Services, with the finding that fidelity of delivery was moderate, with on average only 66% of manual-specified BCTs delivered routinely

in practice (Lorencatto, West, Christopherson, Michie, 2013). Fidelity was also shown to vary substantially across different types of behavioural support sessions (i.e. pre-quit, quit-day, and post-quit), according to individual stop smoking counselors and BCTs. It was also found that on average, the majority (65%) of BCTs delivered within a session were not originally included within the service treatment manual. These findings provide initial insight into variation in the content of behavioural support delivered across NHS Stop Smoking Services. However, these findings were based on a limited number of audio-recorded behavioural support sessions delivered in a single context; the extent to which such findings may be generalised to a larger sample or to behavioural support delivered through different modalities is unknown.

For example, telephone-delivered smoking cessation behavioural support has been increasingly used as both a primary intervention and as a supplement to face-to-face support and/or pharmacotherapy (Lichtenstein, 1996). National quitlines have been established in the UK, Australia, and USA (Zhu et al. 2002; Stead, Perera, & Lancaster, 2009). There is evidence from evaluative trials for the effectiveness of telephone-delivered support (Stead, Perera, & Lancaster, 2006; Zhu et al., 2002). However, the outcomes of telephone-delivered behavioural support have also been shown to vary substantially in practice and research NHS Information Centre, 2012; Stead, Perera, & Lancaster, 2009). This variability remains despite the widespread use of treatment manuals in the delivery of telephone-based behavioural support interventions (Stead, Perera, & Lancaster, 2009; Zhu et al.2 002). There are also inherent differences in delivering behavioural support via telephone rather than in person. For example, compared to sessions in person, it may be easier for counselors delivering telephone support to complete assessment forms or refer to a manual throughout the session without disrupting the clinical interaction.

This principal aims of this study were therefore to (i) evaluate the reliability of the BCT taxonomy method for specifying the content of telephone-based behavioural support and (ii) assess fidelity of delivery of a UK national quitline, as reported and as measured by the BCT analysis of sessions. A secondary aim was to examine discrepancies between stop smoking counselors' self-reported and actual practice, that is, what practitioners '*say they do*' vs. '*what they actually do.*' This is important to examine given the well-established differences in the wider medical literature between health care professional's reported oractuce abd ibserved practice (Cabana et al., 1999; Jones, Gerrity, & Earp, 1990). The extent to which this is applicable to stop smoking counsellors is unknown.

The specific objectives were to:

1. Assess the reliability of an established fidelity assessment method when applied to telephone-based smoking cessation behavioural support

2. Assess the fidelity of delivery of behavioural support in a UK national smoking cessation quitline

3. Investigate variation of fidelity according to: i) session type, ii) counselor, iii) session duration, and iv) specific BCTs

4. Examine the extent of use of additional BCTs not included in the service treatment manual

5. Examine discrepancies between self-reported and actual delivery of 16-evidence based BCTs.

METHODS

Ethical Approval

This study received ethical approval by the Clinical, Educational, and Health Psychology Research Department Ethics Committee (UCL) [Reference: CEHP/2011/038].

Design

This cross-sectional study objectively verified fidelity of delivery by comparing the content, in terms of component BCTs, of treatment manuals to that of transcripts of audio-recorded, telephone-delivered behavioural support sessions.

Study Sample & Materials

Data were obtained from a national UK quitline, which employs four trained stop smoking counselors to provide dedicated telephone-based smoking cessation behavioural support. Counselors had on average 13.5 years experience working as a dedicated stop smoking counselor (range: 13-15) and three had passed the National Centre for Smoking Cessation and Training's skills and knowledge training for delivery of smoking cessation behavioural support (see: www.ncsct.co.uk) (Brose, West, Michie, Kenyon, & McEwen, 2012). The behavioural support provided by the quitline service is typically delivered over four sessions reflecting the different stages of the quit attempt: a pre-quit session, a quit-day session, and two post-quit sessions. The service also offers pharmacological support in the form of nicotine replacement therapy vouchers that are mailed to clients. In 2011, the service had an average self-reported four-week successful quit rate of 51.8%. All counselors are aware of the service treatment manual and have been observed in practice and received feedback as part of their training.

From the quitline, two sets of data were obtained. First, the quitline service treatment manual, which is based on the UK national treatment guidance and training standards for delivering smoking cessation behavioural support (Croghan, E., 2011; West, Lorencatto, Michie, McEwen, et al. 2010). The treatment manual clearly outlines the format and content of sessions to be delivered to *all* clients in either a pre-quit, quit-day, or post-quit behavioural support session. This is accompanied by illustrative dialogues demonstrating how to deliver the recommended content.

Secondly, a set of seventy five sessions consecutively delivered to smokers consenting to the study were audio-recorded by the counselor using a discrete device during a six month data collection period. This minimised the risk of couneslors selecting which sessions to audio-record. Informed consent was obtained from counselors in writing and from clients by audio-recorded telephone. Eleven were excluded as they were incomplete, resulting in a final sample of 64 audio-recordings of three different types of sessions : pre-quit (n=27), quit-day (n=16), and post-quit (n=21). Recordings were anonymised and transcribed verbatim.

Procedure

The procedure and analyses followed those developed for assessing fidelity of face-to-face smoking cessation behavioural support (Lorencatto et al. , 2013). Two researchers (Coder initials: FL, CB) independently coded the treatment manual and session transcripts into component BCTs using an established taxonomy of 45 smoking cessation BCTs (Michie, Hyder, Walia, & West, 2010). This taxonomy has demonstrated reliability as a framework for identifying and characterising component BCTs in the content of treatment manuals and session transcripts (Michie, Hyder, Walia & West, 2010; Lorencatto, West, Seymour, & Michie, 2013). Data were extracted on the number of BCTs identified within each section of the manual (pre-

quit, quit-day or post-quit), as well as within each transcript. If coders identified the same BCT within a section of text, agreement was registered. Where one coder identified a BCT and the other did not, or a different BCT was identified, disagreement was registered. If an intervention component could not be coded by a BCT label from the taxonomy, this was identified as a potential new BCT. Discrepancies were resolved through discussion or consultation with a behaviour change expert (SM).

All practitioners were asked to complete the NCSCT's annual counselor's survey, from which the demographic and professional characteristics of the counselors were drawn (McDermott, Thomson, West, et al. 2012). The survey also assessed self-reported use of the 16 BCTs that have been shown to be significantly associated with improved four-week quit outcomes, and therefore represent those BCTs with the currently best established evidence-base (West, Walia, Hyder, Shahab, & Michie, 2010). Counselor's self-reported use of evidence-based BCTs was assessed for each BCT with the item: 'Thinking about all of the sessions you have delivered over the last 3 months, with what proportion of your clients do you think that you performed the following activities?' (0- *none of them* to 4: *all of them*).

Analyses

A sub-sample of 25% transcripts was double-coded to assess inter-rater coding reliability. Percentage agreement was used to assess reliability rather than Cohen's Kappa. Given the high number of BCTs in the taxonomy (i.e. 45), the probability of chance selecting a particular code is low, and, as Kappa corrects for chance agreement amongst multiple coders, use of Kappa is likely to underestimate reliability (Steinjijans, Diletti, Bomches, Greis, & Sikkeder, 1997). Moreover, the items being coded (i.e. sentences in transcripts) are not mutually exclusive, and multiple BCTs may occur within a single item and at multiple points within the transcript with

coders potentially agreeing on one instance of the identification of a particular BCT but not the other; thus precluding a global present/absent rating that is required to calculate Kappa (Cohen, 1968).

Fidelity was examined by assessing the proportion of BCTs specified in the service treatment manual that was delivered in practice. This was first done according to session type rather than overall as the service treatment manual had individual sections pertaining to the three different stages of the quit attempt and BCTs did not feature uniformly across these three sections of the manual. For example, fidelity of delivery for pre-quit sessions was established by examining the proportion of BCTs specified in the pre-quit section of the manual that was delivered in each pre-quit behavioural support session. This was in turn repeated for quit-day and post-quit sessions then compared across session types to examine variability in fidelity. To obtain an overall estimate of fidelity, the percentage of manual-specified BCTs delivered across the three types of sessions was averaged.

Variation in the extent of fidelity was also examined according to numerous factors. First, variation in fidelity according to individual counselors was examined by comparing the average proportion of manual-specified BCTs delivered by each counselor within their sessions. Secondly, the association between session duration and variation in the proportion of manual-specified BCTs delivered was examined using Pearson correlation. Subsequently, to assess variation in fidelity across BCTs, the proportion of sessions in which each BCT was delivered according to the manual was calculated. This was first done according to session type, and then averaged across session types as not all BCTs featured consistently across all three sections of the manual. Lastly, to ascertain what proportion of session content was not manual-specified, the

number of BCTs delivered that were not included in the manual was calculated as a percentage of the total number of BCTs delivered within a session.

To establish a percentage of self-reported use of evidence-based BCTs across counselors, the total scores for *self-reported* use of each of the 16-evidence based BCTs was established by summing response ratings for each BCT across the four counselors. For each BCT, this total score was then presented as a percentage of the maximum possible total score of 16. The resulting percentages represent the percentage of sessions in which counselors report using each of the evidence-based BCTs in. The percentage *actual* use for each evidence-based BCT was assessed by calculating the total number of sessions across counselors each BCT was actually delivered. This was then presented as a percentage of the maximum possible 64 sessions. Differences between percentage self-reported use and percentage actual use were then examined for each evidence-based BCT using paired sample t-test.

RESULTS

1. *Reliability of fidelity assessment method*

Average inter-rater reliability for coding was 81.9% across transcripts (range 75.4%-89.9%), which is high (i.e. >75%) (Cohen, 1968). Any discrepancies were easily resolved through discussion, and no new additional BCTs were identified during coding.

2. *Fidelity of delivery*

A full list of BCTs identified in each section of the service treatment manual is available in Supplementary Table 1. Across all transcripts (n=64), on average 41.8% (range: 8-82%) of

manual-specified content was identified in session transcripts, and therefore delivered in practice (Table 1; Figure 1).

3. Variation of fidelity

(i) *According to type of session*

The pre-quit section of the manual contained 22 BCTs (Supplementary Table 1), of which on average 10 (46%) were delivered in pre-quit sessions (SD 16.9; range: 14-82%) (Table 1). The manual-specified content relating to quit-day support contained 25 BCTs (Supplementary Table 1) of which on average 9 (35%) were delivered in quit-day sessions (SD 14.8; range: 8-60%). The post-quit support section of the manual featured 28 BCTs (Supplementary table 1), of which on average 12 (42%) were delivered in post-quit sessions (SD 16.3; range: 8-82%) (Table 1).

(ii). According to individual counselor

Of the 64 sessions examined, the four counselors delivered on average 16 sessions each (range: 4-27). The average proportion of manual-specified BCTs delivered by each counselor was 41.8%, ranging from 32% to 49% across the four counselors (Table 1).

(iii) *As a function of session duration*

On average, sessions lasted 12.40 minutes (SD 6.55). There was a positive correlation between the duration of a session and the proportion of manual-specified BCTs delivered in the session ($r=0.452$, $p<0.01$) (Table 1).

(iv) *By specific BCT*

Each manual-specified BCT was delivered on average in 40% of the appropriate sessions as specified by the service treatment manual (range: 0-95%)(see Supplementary Table 2). The

BCTs for which fidelity was highest included: 'Giving options for additional and/or later support (delivered appropriately in 95% of sessions),' 'Information gathering and assessment (88%),' and 'Providing feedback on current behaviour and progress (85%).' The BCTs for which fidelity was lowest included: 'set graded tasks (0%),' 'Measure CO and explain purposes of CO monitoring (2%),' and 'prompt commitment from the client there and then (3%)' (Table 2).

4. *Delivery of BCTs not included in the manual (i.e. additional content)*

Sessions contained on average a total of 14 BCTs per session (SD 5.3) (Table 1). Of these, on average 3 (23%) were non-manual specified (range: 0-53%). A full list of BCTs most frequently delivered as additional content in each type of session is available in Supplementary Table 3.

5. *Self-reported vs. Actual use of 16 evidence-based BCTs*

The average score for self-reported use of each of the 16 evidence-based BCTs was 12 (SD 3.03); indicating that counselors reported using a BCT on average in 75% of sessions (range: 31-100%). In contrast, the average number of sessions each BCT was actually delivered in was 22, corresponding to 34% actual use of each evidence-based BCT (range: 2-89%). Differences between average percentage reported and average percentage actual use across each of the 16 evidence-based BCTs were significant[$t=-5.52$ (15), $p<0.001$].

DISCUSSION

The content of telephone-delivered behavioural support for smoking cessation can be reliably coded into component BCTs using an established taxonomy (Michie, Hyder, Walia, & West, 2011; Lorencatto, West, Seymour, & Michie, *2013*). Inter-rater coding reliability was consistently high (average percentage agreement 81.9%), similar to levels achieved for coding

face-to-face behavioural support (Lorencatto, West, Seymour, & Michie, *2013;* Lorencatto, West, Christopherson, & Michie, *2013*). Using this method to examine delivery of support in a national quitline service found that on average less than half (42%) of manual-specified content was routinely delivered in practice. This is considered to be 'low' fidelity as guidelines currently suggest that if 80-100% of pre-specified content in a manual is delivered, this is classifiable as 'high' fidelity, whereas less than 50% delivery is deemed to be 'low' (Borrelli, 2011). Although these levels of fidelity are lower than those observed for behavioural support delivered in person (66%) (Lorencatto, West, Michie et al., *Under Review)*, they are consistent with those found in systematic reviews and trials of behaviour change interventions in other domains (range: 40-50%) (Dane & Schneider, 1998; Dusenbury et al. 2003; Tober et al. 2009; Hardeman et al. 2008). The current findings therefore add to an increasing body of evidence demonstrating the poor and variable delivery of behaviour change interventions.

The few published fidelity assessment methods have typically been conducted as part of trial evaluations in research settings (Dane & Schneider, 1998; Dusenbury et al. 2003; Tober et al. 2009; Hardeman et al. 2008.) The present study provides an example of a reliable fidelity assessment method applied in the context of clinical practice. The observed variation in fidelity of delivery has implications for improving practice within services and designing more effective interventions. For instance, it was possible to examine variation in fidelity according to session type, counselor, and individual BCT. Fidelity was shown to be lowest in pre-quit sessions (35%), for a particular counselor (35%), and BCTs 'set graded tasks' (0%) and 'measure and explain purpose of CO monitoring (2%),' These are aspects of service delivery requiring improvement and represent specific training needs to be addressed in future training programs in order to

improve consistency in service delivery and fidelity to manual specifications. This allows for more beneficial and targeted use of training and professional development resources.

The finding that longer session duration was associated with higher fidelity is consistent with reviews of fidelity of delivery for other complex interventions (Moncher & Prinz, 1991), but not with findings for face-to-face behavioural support for smoking cessations (Lorencatto, West, Christopherson, & Michie, *2013*). This lack of association may reflect the narrower range of session duration in face-to-face than telephone-delivered support (5 – 36 mins and 3-62 mins respectively).

Furthermore, telephone-delivered sessions contained less additional content than face-to-face sessions (23% vs 65%) (Lorencatto, West, Michie et al, *2013)*. It is not clear whether additional content is beneficial or detrimental to the effect of the intervention: it may dilute the effect of manual-specified content and introduce substantial variability into the content of sessions delivered in practice, or it may enhance the effect by delivering additional effective or complementary BCTs. Another area requiring further research is the issue of quality of BCT delivery, that is, how well BCTs are delivered, as it is not enough to simply state whether a BCT is delivered or not.

The observed over-reporting of BCT delivery is an important finding from both research and clinical perspectives, demonstrating that clinician self-reported practice in questionnaires, interviews and assessments cannot be assumed to accurately reflect actual practice. This finding is consistent with studies demonstrating differences between what physicians say they do and what they actually do (Jones, Gerrity, & Earp, 1990; Cabana, Rand, & Powe, 1999; Moy, Grant, Weiss et al. 1999). Such differences underline the need to increase observation of counselors in

practice, both in order to monitor practice in audit and research studies and also to provide accurate feedback to clinicians in improving their practice.

More importantly, these findings also raise the debate surrounding the rationale for expecting 100% fidelity and the issue of the extent to which treatment manuals are fit for purpose. One explanation of the observed low fidelity in the quitline service may be that manuals reflect unrealistic expectations of what can or should be delivered within the limitations of a particular context (e.g. session duration, expertise of counselors). The treatment manual for the quitline was very extensive, requiring counselors at times to deliver up to 28 BCTs per session. Sessions were found to last on average just 13 minutes. Therefore, for counsellors to achieve 100% fidelity in this context they would need to deliver on average 2 BCTs per minute, assuming the counsellor spoke for the entire session which is unlikely to be the case. This highlights the question as to whether 'more is better.' In this context, 100% fidelity may not represent an expectation that is feasible, relevant, or in fact beneficial or desirable in clinical practice as it may even be detrimental to the therapeutic relationship..

Furthermore, it is unclear whether 100% fidelity is actually needed to achieve effective outcomes. Counselors may work better if a specified degree of flexibility or adaptation to the local context or individual client is allowed for (Craig et al. 2008). For instance, the handbook for the widely implemented NHS Health Trainers intervention was specifically developed to be applied flexibly (Michie et al. 2008). The handbook provided information and explanations about psychological techniques and theories of behaviour change, alongside practical suggestions of how to use these techniques in practice. As lay Health Trainers delivering this intervention worked in a wide range of settings with individuals of varying health needs, the handbook was intended to be used flexibly and adapted to the local context (Michie et al. 2008). This flexible

approach may be more appropriate for smoking cessation behavioural support interventions, which are also widely implemented in a range of settings with variable client groups (Bauld et al. 2010).

For this reason, some argue in favour of flexibility in delivery rather than strict adherence to treatment manuals (Leventhal & Friedman, 2004). Reducing the number of BCTs that counsellors are expected to deliver with 100% fidelity may increase the likelihood of manual specifications being adhered to in practice. Proponents of a more flexible approach to fidelity argue that 100% fidelity should not be expected for the entire intervention content, but rather, only for those intervention components that are deemed to be evidence-based, essential and unique to the intervention (Collins et al. 2009; Waltz et a. 1993). Evidence for the effectiveness of individual smoking cessation BCTs is emerging (Lorencatto et al., 2012; Michie, Churchill, & West, 2011; West et al. 2011; West et al. 2010). Manuals are important for providing a benchmark for practice and promoting consistency and standards in service provision (Michie, 2008). In line with a more flexible approach, treatment manuals could therefore focus on recommending 100% fidelity for only these BCTs with an established evidence-base, and allow for flexibility in the delivery of those recommended BCTs for which individual effectiveness has not yet been established. Such an approach may represent a more feasible, realistic, and potentially beneficial, expectation of fidelity in the context of clinical practice.

Also, as advocated by the Medical Research Council framework for developing and evaluating complex interventions, it is necessary to consider implementation early in the initial stages of developing an intervention (Craig et al. 2008). The recently developed multi-phase optimisation strategy (MOST) for developing interventions argues in favour of a 'resource management principle' (Collins, Murphy, & Stretcher, 2007). That is, the 'treatment package' (i.e.

components) of an intervention should be carefully selected on the basis of theoeretical rationale, empricial evidence, and clinical experience whilst also maintaining a consideration of potential implementation constraints, such as what will be subsequently feasible to deliver in practice given limitations in time, financial and practical resources (e.g. number of counsellors available). Keeping fidelity and implementation in mind when developing future smoking cessation behavioural support interventions will potentially promote the design of interventions that are more feasible and realistic to implement in clinical practice than is currently the case, as displayed by the unrealistic expectations in the treatment manual of the telephone quitline.

The present study demonstrates the application of a reliable fidelity assessment method to telephone-delivered behavioural support. Observed findings in a UK national quitline illustrate the potential discrepancies and variability that may occur in delivering such support. Whilst the general findings of this study are consistent with previous evidence about the delivery of behavioural interventions, caution needs to be observed when generalising these results since the data have been drawn from the application of this method to only one quitline service. Assessing fidelity of delivery is one step towards identifying potential factors contributing to observed variability in outcomes and identifying targets for future interventions to improve service provision and outcomes. This study provides an example of a reliable fidelity assessment method for clinical practice settings and emphasizes the need to establish routine procedures for monitoring fidelity of delivery of behavioural support and other therapeutic interventions. The taxonomy of BCTs underpinning this method provides a consistent, common language with which to compare the content of manuals and sessions, and to reliably quantify fidelity of delivery. Evidence that it is possible to train novice coders to reliably apply this taxonomy (Lorencatto, West, Seymour, & Michie, *2013)* suggests this as a potentially useful tool for

service monitoring, evaluation and improvement. This method can be extended to other types of intervention, since taxonomies of BCTs are available for other health behaviours, such as healthy eating and physical activity (Michie, Ashford, Sniehotta, et al 2011) and excessive alcohol use (Michie, Whittington, West et al 2012), and a generic, integrative taxonomy has been developed on the basis of these previously published taxonomies (Michie, Johnston, Abraham, et al in press).

REFERENCES

Borrelli, B., Sepinwall, D., Ernst, D., Bellg, A., Czajkowski, S., Breger, R., et al. (2005). A new tool to assess treatment fidelity and evaluation of treatment fidelity across 10 years of health behaviour research. *Journal of Consulting and Clinical Psychology,* vol. 73 (5), p.852-860.

Borrelli, B. 2011. The assessment, monitoring and enhancement of fidelity in public health clinical trials. *Journal of Public Health Dentistry*, 71(s1): 52-63.

Boutron I, Moher D, Altman D, Scultz K, Ravaud P. 2008. Extending the CONSORT Statement to Randomized Trials of Non-pharmacologic Treatment: Explanation and Elaboration. *Annals of Internal Medicine*, 148:295-309.

Brose L.S., McEwen A & West R. (in press). Does it matter who you see to help you stop smoking? Short-term quit rates across specialist Stop Smoking Practitioners in England. Addiction, online first 9.5.2012. doi:
10.1111/j.1360-0443.2012.03935.x.

Brose, L. S., West, R., Michie, S., Kenyon, J. A., & McEwen, A. (2012). Effectiveness of an online knowledge training and assessment program for stop smoking practitioners. *Nicotine Tob Res, 14*(7), 794-800. doi: 10.1093/ntr/ntr286

Cabana, M. D., Rand, C. S., Powe, N. R., Wu, A. W., Wilson, M. H., Abboud, P. A., & Rubin, H. R. (1999). Why don't physicians follow clinical practice guidelines? A framework for improvement. *JAMA, 282*(15), 1458-1465.

Cohen, J. (1968). Weighted kappa: Nominal scale agreement with provision for scaled disagreement or partial credit. *Psychological Bulletin* 70 (4): 213–220. doi:10.1037/h0026256

Croghan, E. 2011. Local Stop Smoking Services: Service delivery and monitoring guidance 2011/2012. *Department of Health.* Accessed on 11-10-2012. Retrieved from:
http://www.dh.gov.uk/prod_consum_dh/groups/dh_digitalassets/documents/digitalasset/dh_125939.pdf

Curry, S., McBride, C., Grothaus, L., Louie, D., & Wagner, E. 1995. A randomized trial of self-help materials, personalized feedback and telephone counseling with non volunteer smokers. *Jounral of Consulting and Clinical Psychology,* vol. 63, p. 1005-1014.

Dane, A. & Schnieder, B. 1998. Program integrity in primary and early secondary prevention: are implementation effects out of control? *Clinical Psychology Review*, vol. 18, 22-45.

Dewing, S., Mathews, C., Cloete, A., Schaay, N., Shah, M., Simbayi, L., & Louw, J. (2013). From Research to Practice: Lay Adherence Counsellors' Fidelity to an Evidence-Based Intervention for Promoting Adherence to Antiretroviral Treatment in the Western Cape, South Africa. *AIDS Behav*. doi: 10.1007/s10461-013-0509-x

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: Theory and Practice,* 18(2): 237-256.

Eccles, M., Armstrong, D., Baker, R., Cleary, K., Davies, H., Davies, S., et al. 2009. An implementation research agenda. *Implementation Science,* vol. 4 (18), pp. 1-7.

Finlkestein, J., Lozano, P., Shulruff, R., Inui, T., Soumerai, S., Ng, M., et al. 2000. Self-reported physician practices for children with asthma: are national guidelines followed? *Pediatrics,* vol. 106, 4, pp. 886-896.

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. doi: 10.1186/1745-6215-14-15

Hardeman, W., Michie, S., Fanshawe, T., Prevost, T., Mcloughlin, K., & Kinmonth, A.L. 2008. Fidelity of a physical activity intervention: predictors and consequences. *Psychology and Health,* 23(1): 11-24.

Jones, T., Gerrity, M., & Earp, J. 1990. Written case simulations: do they predict physicians' behaviour? *Journal of Clinical Epidemiology,* vol. 8, p. 805-815.

Judge, K., Bauld, L., Chesterman, J., & Ferguson, J. (2005). The English smoking treatment services:

short-term outcomes. Addiction, vol. 100 (2), 46-58.

Lancaster T, Stead LF. 2005, Individual behavioural counseling for smoking cessation. *Cochrane Database Syst Rev*. CD001292.

Lawrence, C., Zhu, S., Nelson, D., Arikian, N., Nugent, S., Partin, M., et al. 2006. Benefits of telephone care over primary care for smoking cessation *Archives of Internal Medicine,* vol. 166, p. 536-542.

Leventhal, H., & Friedman, M. 2004. Does establishing fidelity of treatment help in understanding treatment efficacy? *Health Psychology,* vol. 23 (5), pp.452-456.

Lichtenstein, E., Glasgow, R., Lando, H., Ossip-Klein, D., & Boles, S. 1996. Telephone counseling for smoking cessation: rationales and meta-analytic review of evidence. *Health Education Research, Theory and Practice,* vol. 11 (2), p. 243-257.

Lorencatto, F., West, R., & Michie, S. 2012. Identifying evidence-based behaviour change techniques for smoking cessation in pregnancy. *Nicotine & Tobacco Research*, DOI: 10.1093/ntr/ntr324.

Lorencatto, F., West, R., Seymour, N., Kaplunov, E., & Michie, S. 2013. Developing a method for specifying components of behaviour change interventions delivered in practice: the example of smoking cessation. *Journal of Consulting & Clinical Psychology, In press*.

Lorencatto, F., West, R., Christopherson, C., & Michie, S. (*Under Review*). Assessing fidelity of delivery of smoking cessation behavioural support in practice.

McDermott, M., Thomson, H., West, R., Kenyon, J, & McEwen, A. (2012). Translating evidence-based guidelines into practice: a survey of practices of commissioners and managers of the English Stop Smoking Services. *BMC Health Services Research,* vol 12 (121).

Michie, S., Whittington C., Hamoudi Z., Zarani, F., Tober, G., & West R. 2012. Behaviour change techniques to reduce excessive alcohol use and their associations with outcome. *Addiction,* available online.

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

Michie, S., Ashford, S., Sniehotta, F.F., Dombrowski, S.U., Bishop, A. & French, D.P. 2011. A refined taxonomy of behaviour change techniques to help people change their physical activity and healthy eating behaviours – the CALO-RE taxonomy. *Psychology & Health, vol 26 (11).*

Michie, S., Walia, A., Hyder, N., West, R. 2010. Development of a taxonomy of behaviour change techniques used in individual behavioural support for smoking cessation. Addictive behaviours, 36(4): 315-319.

Michie, S., Fixsen, D., Grimshaw, J., & Eccles, M. 2009. Specifying and reporting complex behaviour change interventions: the need for a scientific method. Implementation science,vol. 4 (40), doi:10.1186/1748-5908-4-40.

Michie, S. 2008. What works and how? Designing more effective interventions needs answers to both questions. *Addiction,* vol. 103 (6), pp. 886-887.

Moncher, F., &Prinz, R. 1991. Treatment fidelity in outcome studies. *Clinical psychology review,* vol. 11, 247-266

Moy, J., Grant, E., Turner-Roan, K., Li, T., & Weiss, K. 1999. Asthma care practices, perceptions and beliefs of Chicago-area asthma specialists. *Chest,* vol 37, p. 78-82.

NHS Information Centre 2012. Statistics on NHS Stop Smoking Services: England, April 2011 to March 2012. NHS IC. Accessed on 27-11-2012. Retrieved from: http://www.ic.nhs.uk/statistics-and-data-collections/health-and-lifestyles/nhs-stop-smoking-services/statistics-on-nhs-stop-smoking-services--england-april-2011-to-march-2012.

Orleans, T. 2007. Increasing the demand for and use of effective smoking-cessation treatments: Reaping the full health benefits of tobacco-control science and policy gains--In our lifetime. *American Journal of Preventative Medicin*e, vol. 33 (6), p. 340-348. doi: 10.1016/j.amepre.2007.09.003

Shahab, L., & McEwen, A. 2009. Online support for smoking cessation: a systematic review of the literature. *Addiction,* vol. 104 (11), pp. 1792-1804.

Stead, L., Perera, R., & Lancaster, T. 2009. Telephone counseling for smoking cessation (review). *Cochrane Database of Systematic Reviews,* CD002850.

Stead LF, Lancaster T. (2005). Group behaviour therapy programmes for smoking cessation. *Cochrane Database Syst Rev*. CD001007. DOI: 10.1002/14651858.CD001007.pub2.

Steinijans, V., Diletti, E., Bomches, B., Greis, C., & Solleder, P. 1997. Interobserver agreement: Cohen's kappa coefficient does not necessarily reflect the percentage of patients with congruent classifications. *Int Journal Clinical Pharmacological Therapy,* vol. 35 (3), p.93-5.

Tober, G., Clyne, W., Finnegan, O., Farrin, A., & Russell, I., 2009. 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(6): 675-682.

Tyhrian, J., Panagiotakos, D., Polychronopoulos, E., West, R., Zatonski, W., & John, U. 2007. The relationship between smokers' motivation to quit and intensity of tobacco control at the population level: a comparison of five European countries. *BMC Public Health¸* vol. 8 (2).

West, R., Walia, A., Michie, S., & Walia, A. (2010). Behaviour change techniques used by English Stop Smoking Services and their associations with short- term quit outcomes. Nicotine & Tobacco Research, vol. 12 (7), pp. 742-747. doi:10.1093/ntr/ntq074

West, R., Lorencatto, F., Michie, S., Churchill, S., Willis, N., & McEwen, A.(2010). *NCSCT Training Standards: Learning outcomes for training stop smoking specialists.* NHS Centre for Smoking Cessatio and Training, Department of Health UK. Available at: http://www.ncsct.co.uk/Content/FileManager/documents/NCSCT_Training_Standard.pdf

West, R., & Stapleton, J. 2008. Clinical and public health significance of treatments to aid smoking cessation. *European Respiratory Review,* 17(110): 199-204

West, R., McNeill, A. & Raw, M. 2000. National smoking cessation guidelines for health professionals: an update. *Thorax,* 55(12): 987–999.

Zhu, S., Anderson, C., Tedeschi, G., Rosbrook, B., Johnson, B., Byrd, C., et al. 2002. Evidence of real-world effectiveness of a telephone quitline for smokers. *New England Journal of Medicine,* vol. 347 (14), p.1097-1093.

Zhu, S., Stretch, V., Balabanis, M., Rosbrook, B., Sadler, G., & Pierce, J. 1996. Telephone counseling for smoking cessation: effects of single-session and multiple session interventions. *Journal of Consulting and Clinical* Psychology, vol. 64, p. 202-11.

Table 1. Session characteristics and the proportion of BCTs specified in the treatment manuals delivered in behavioural support sessions; summarised across practitioner and session type.

Practioner ID	Session type (1=pre-quit; 2= Quit day; 3=Post-quit) (N sessions)	Duration (sec) (range)	No. of BCTs in Manual (according to session type)	Average No. of manual-specified BCTs delivered (%; range)	Average Total No. of BCTs in session (range)	Average No. of non-manual specified BCTs in session (%; range)
P01	1 (n=10)	1490 (414-3754)	22	9 (41%) (14-68%)	13 (4-23)	4 (31%) (17-53%)
P01	2 (n=2)	1481 (742-2221)	25	8 (32%) (28-40%)	15 (13-16)	6 (40%) (40-40%)
P01	3 (n=5)	716 (248-1105)	28	10 (36%) (24-52%)	11 (7-16)	0 (0%) (0-1)
P02	1 (n=7)	1653 (797-2375)	22	13 (59%) (41-82%)	17 (10-24)	4 (24%) (10-33%)
P02	2 (n=9)	931 (475-1607)	25	10 (40%) (24-60%)	17 (9-25)	7 (41%) (33-53%)
P02	3 (n=11)	947 (317-1521)	28	14 (50%) (14-62%)	16 (6-21)	2 (13%) (6-33%)
P03	1 (n=7)	877 (627-1357)	22	8 (36%) (18-45%)	9 (4-13)	1 (11%) (0-31%)
P03	2 (n=4)	324 (172-864)	25	6 (24%) (8-52%)	9 (3-20)	3 (33%) (17-38%)
P03	3 (n=5)	391 (154-594)	28	9 (32%) (18-55%)	14 (5-18)	2 (14%) (0-31%)
P04	1 (n=3)	1870 (1160-2531)	22	12 (54%) (41-73%)	16 (13-22)	5 (31%) (23-36%)
P04	2 (n=1)	288	25	7 (28%)	13	6 (46%)
P04	3 (n=0)	-	-	-	-	-

Table 2. Number of behavioural support sessions each BCT was delivered in according to manual specification across session types.

	Total No. sessions BCT delivered in according to manual specification
Set graded tasks	0 /16 (0%)
Prompt commitment from the client there and then	2/64 (3%)
Measure CO and explain the purposes of CO monitoring	1/64 (2%)
Explain the importance of abrupt cessation	4/64 (6%)
Provide reassurance	41/64 (64%)
Advise on avoiding social cues for smoking	5/64 (8%)
Prompt self-recording	2/21 (10%)
Advise on environmental restructuring	5/43 (12%)
Promote self-reward	2/16 (13%)
Advise on/facilitate use of social support	8/64 (13%)
Advise on conserving mental resources	3/21 (14%)
Facilitate action planning/ develop a treatment plan	7/48 (15%)
Advise on changing routine	3/16 (19%)
Facilitate relapse prevention and coping	8/37 (22%)
Distract from motivation to engage in behaviour	4/16 (25%)
Strengthen ex-smoker identity	17/64 (27%)
Emphasise choice	9/27 (33%)
Provide information on the health consequences of smoking and smoking cessation	17/48 (35%)
Facilitate barrier identification and problem solving	27/64 (42%)
Provide rewards contingent on not smoking	9/21 (43%)
Facilitate identification of reasons for wanting and not wanting to stop smoking	9/21 (43%)
Prompt review of set goals	9/21 (43%)
Ask about experiences of stop smoking medication that the smoker is currently using	17/37 (46%)
Build general rapport	39/64 (61%)
Provide normative information about others' behaviour and experiences	41/64 (64%)
Boost motivation and self-efficacy	42/64 (66%)
Provide rewards contingent on effort or progress	14/21 (67%)
Advise on stop smoking medication	49/64 (75%)
Facilitate goal setting	22/27 (81%)
General communication approaches	52/64 (81%)
Provide feedback on current behaviour and progress	18/21 (86%)
Give options for additional and later support	61/64 (95%)

Information gathering and assessment	57/64 (88%)
Explain how tobacco dependence develops	7/43 (16%)
Explain expectations regarding the treatment programme	28/43 (65%)
Provide information on withdrawal symptoms	11/64 (17%)
reflective listening	13/21 (62%)

Table 3. Percentage self-reported and actual use of sixteen evidence-based BCTs.

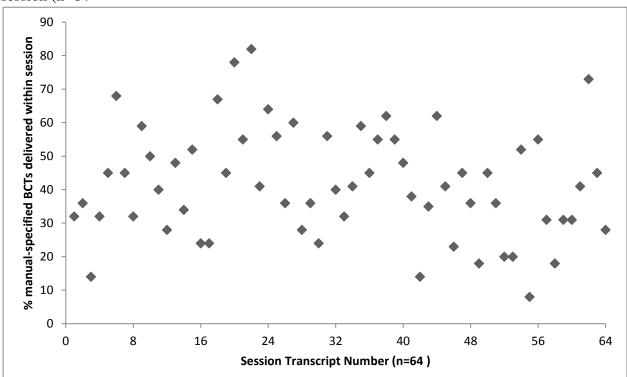| BCT Label	Cumulative score of self-reported use across counselors (%) (max=16; 100%)	Number (%) of sessions BCT actually delivered in across counselors (max= 64; 100%)
'Boost motivation and self-efficacy'	14 (88%)	45 (71%)
'Provide rewards contingent on not smoking'	16 (100%)	17 (27%)
'prompt commitment from the client there and then'	12 (75%)	2 (3%)
'strengthen ex-smoker identity'	11 (69%)	15 (24%)
'Measure CO'	5 (31%)	1 (2%)
'Facilitate relapse prevention and coping'	13 (81%)	8 (12%)
'Advise on changing routine'	14 (88%)	12 (19%)
'Advise on conserving mental resources'	14 (88%)	8 (13%)
'Advise on stop smoking medication'	13 (81%)	56 (88%)
'Advise on/facilitate use of social support'	13 (81%)	10 (15%)
'Ask about experiences of stop smoking medication that the smoker is currently using'	13 (81%)	27 (37%)
'Give options for additional and later support'	14 (88%)	57 (89%)
'General practitioner communication approaches'	11 (69%)	42 (65%)
'Provide reassurance'	13 (81%)	45 (71%)
'Provide information on withdrawal symptoms'	11 (69%)	13 (21%)
'Explain purpose of CO monitoring'	5 (31%)	1 (2%)

Figure 1. Percentage of manual-specified BCTs delivered (i.e. % fidelity) in each examined session (n=64