



Review

Reporting Treatment Fidelity in Behavioral Tobacco Treatment Clinical Trials: Scoping Review and Measurement Recommendations

Ramzi G. Salloum PhD^{1,✉}, Alana M. Rojewski PhD^{2,✉}, Megan E. Piper PhD^{3,✉}, Janice A. Blalock PhD⁴, Belinda Borrelli PhD^{5,✉}, Lindsay M. Boyce MLIS⁶, Jennifer A. Minnix PhD⁴, Omara Dogar PhD^{7,✉}, Rachel L. Tomko PhD^{2,✉}, Douglas E. Jorenby PhD³, Chris Kotsen PhD⁶, Jamie S. Ostroff PhD⁶; on behalf of the Treatment Fidelity Workgroup of the SRNT Treatment Network

¹Department of Health Outcomes and Biomedical Informatics, University of Florida, College of Medicine, Gainesville, FL, USA; ²Department of Public Health Sciences, Medical University of South Carolina, Charleston, SC, USA; ³Center for Tobacco Research and Intervention, University of Wisconsin-Madison, Madison, WI, USA; ⁴Department of Behavioral Science, University of Texas, MD Anderson Cancer Center, Houston, TX, USA; ⁵Center for Behavioral Science Research, Boston University, Henry M. Goldman School of Dental Medicine, Boston, MA, USA; ⁶Department of Psychiatry and Behavioral Science, Memorial Sloan Kettering Cancer Center, New York, NY, USA; ⁷Department of Health Sciences, University of York, York, United Kingdom

Corresponding Author: Jamie S. Ostroff, PhD, Department of Psychiatry & Behavioral Sciences, Memorial Sloan Kettering Cancer Center, 1275 York Avenue, New York, NY 10065, USA. Telephone: (646)888-0041; E-mail: ostroffj@mskcc.org

Abstract

Introduction: Adoption of rigorous standards for reporting treatment fidelity is essential for advancing discovery, validation, and implementation of behavioral treatments. Whereas the NIH Behavior Change Consortium (BCC) developed an assessment tool to assess the quality of reporting and monitoring of treatment fidelity across health behavior change interventions, it has not yet been applied specifically to treatment fidelity in behavioral tobacco treatment trials.

Aims and Methods: We conducted a scoping review of peer-reviewed, clinical trials of behavioral adult tobacco treatment interventions published in English between 2006 and 2018. Using the BCC treatment fidelity checklist, articles were coded for the presence or absence of various treatment fidelity strategies within each of 5 domains: Design, Training, Delivery, Receipt, and Enactment. Eligible articles ($N = 755$) were coded by two independent coders.

Results: The proportion of reporting strategies varied within the fidelity domains, ranging from 5.2% to 96.3% in Design, 1.9% to 24.9% in Training, 2.6% to 32.3% in Delivery, 5.2% to 44.3% in Receipt, and 6.7% to 43.2% in Enactment. The mean proportion of adherence to treatment fidelity strategies within each domain was: Design (68%), Training (14%), Delivery (15%), Receipt (16%), and Enactment (25%). Only 11 studies achieved $\geq 80\%$ reporting across >1 fidelity domain. There was no evidence for improvement in fidelity reporting across the 13-year time frame from the initial BCC publication to the present.

Conclusions: These findings illustrate the lack of consistency in fidelity reporting in tobacco treatment trials and underscore the challenges faced in evaluating rigor and reproducibility, as well as interpretation and dissemination of findings. Recommendations are made for improving fidelity reporting in tobacco treatment trials.

Implications: The SRNT Treatment Research Network sponsored a scoping review to summarize the current state of reporting treatment fidelity and make recommendations for best practices in reporting fidelity in tobacco treatment trials. The review identified a lack of consistency in fidelity reporting, illustrating the challenges faced in evaluating rigor, and reproducibility, as well as interpretation and dissemination of findings.

Introduction

Tobacco use remains the leading cause of preventable death and disease in the world and is responsible for more than 7 million deaths annually.¹ As highlighted in numerous tobacco treatment guidelines,²⁻⁵ systematic reviews and meta-analyses,⁶⁻¹¹ behavioral interventions have proven efficacious in promoting tobacco cessation, particularly when combined with pharmacotherapy. However, these reviews have also highlighted extensive variation in trial outcomes, even among trials purporting to test the same or similar treatments. For example, a meta-analysis of 28 motivational interviewing interventions for tobacco cessation found that quit rates varied from 0% to 60% across trials and noted that the high variability in the reporting of treatment delivery and fidelity limited the ability to draw generalizable conclusions about the effectiveness of treatment.⁹ Similarly, in multi-site trials, wide variation in tobacco cessation outcomes across sites has been observed despite testing the same behavioral intervention.¹²

Behavioral interventions are typically complex, consisting of multiple features¹³ including theory-based “active ingredients” of treatment (eg strategies to improve quitting self-efficacy, or avoid environmental stimuli associated with tobacco use) that influence the mechanisms hypothesized to produce behavioral change. Behavioral interventions also vary in delivery modality (eg group, individual, telehealth, web-based, with or without pharmacotherapy), setting (eg primary care, specialized outpatient program), provider characteristics (eg one or more clinicians, gender, personality, training, skill), and duration and dose of treatment. If delivered as intended, interventions have the greatest likelihood of producing the outcomes that they are hypothesized to produce.¹⁴ However, in practice, behavioral interventions are often delivered inconsistently resulting in variability in treatment outcomes.^{15,16}

Reducing variation in treatment effects and increasing confidence that changes in the dependent variable are due to the independent variable could be achieved by enhancing treatment fidelity.¹⁷ The US NIH Behavioral Change Consortium (BCC) defined treatment fidelity as methodological strategies used to monitor and enhance the reliability and validity of behavioral interventions.¹⁷ The overall goal of enhancing treatment fidelity is to increase scientific confidence that changes in the dependent variable are attributable to the independent variable and that behavioral interventions are implemented as described in the original protocol.^{17,18} Careful adherence to treatment protocols is one aspect of treatment fidelity that increases confidence in the interpretation of effect sizes by reducing random and unintended variability.^{15,17} Further, treatment fidelity enhances both internal validity (ie the treatment is delivered as intended) and external validity (ie the treatment can be replicated and applied in real-world settings).¹⁹ Failure to adequately assess and report the actual delivery of treatment features can undermine attempts to replicate research and to optimize treatment delivery in practice,^{17,20} and could limit the real-world impact of high-quality research findings.^{19,21}

To accelerate treatment progress, the NIH BCC developed a framework and measurement tool to assess, monitor, and enhance treatment fidelity across five domains: *Treatment Design* (ensuring that a study adequately tests its hypotheses in relation to its

underlying theoretical and clinical processes); *Provider Training* (ensuring that those who are delivering treatment are trained to criterion and that skills are maintained over time); *Treatment Delivery* (ensuring that the treatment is delivered as intended); *Treatment Receipt* (ensuring that the treatment was “received” by the participant; eg participant understood and was able to use the skills learned in treatment); and *Treatment Enactment* (ensuring that the participant used the skills and/or behaviors in real life settings).¹⁷ The BCC treatment fidelity assessment tool was later updated to reflect additional important considerations for treatment fidelity, such as selecting providers, measuring additional confounders, theory testing, and multicultural considerations.¹⁹

Building upon the earlier seminal paper using the NIH BCC treatment fidelity tool to assess broad health behavior change studies over a ten year period (1990–2000),¹⁷ the present review was undertaken by a workgroup of the Treatment Network of the Society for Research on Nicotine and Tobacco (SRNT) to examine treatment fidelity reporting among tobacco treatment studies of adult tobacco users published between the years of 2006 and 2018. Our goals were (1) to review the current state of reporting of assessment and monitoring of treatment fidelity in tobacco treatment clinical trials; and (2) to make recommendations for improving the reporting of tobacco treatment fidelity in clinical trials.

Methods

Eligibility Criteria

We included clinical trials of behavioral tobacco treatment interventions, published in English between 2006 and 2018, that enrolled adults who use any tobacco product (eg cigarettes, cigars, smokeless tobacco, e-cigarettes, and hookah) and included one or more tobacco treatment outcome (eg cessation, reduction). Non-peer reviewed articles, protocols, reviews, and systematic reviews were excluded. We included pilot trials, and non-randomized trials, as well as all trials testing pharmacotherapies delivered with a behavioral counseling component, regardless of whether the focus of the trial was on the behavioral treatment component(s).

Information Sources and Search Strategy

To identify potentially relevant articles, the following bibliographic databases were searched from January 2006 through July 2018: PubMed (Legacy), Embase, and Cochrane Central Register of Controlled Trials. The search strategies were drafted by a medical research informationist (LB) and further refined by the core workgroup. The final search strategy for PubMed can be found in [Appendix A](#). The final search results were exported into EndNote, and duplicates were removed using the Bramer method.²²

Characteristics of the Treatment Fidelity Coders

Ten of the coders held doctoral degrees and nine were masters- and bachelors-level research assistants at universities in the United States and England. Their areas of specialization included psychology, medicine, public health, and epidemiology. All doctoral-level coders

were members of the SRNT Treatment Network Treatment Fidelity Workgroup with expertise in tobacco treatment.

Article Selection

Articles were initially screened for inclusion by title and abstract by individual doctoral-level reviewers, using inclusion and exclusion criteria developed *a priori*. During these screening stages, Microsoft Excel files were used to capture reviewers' responses to the preset exclusion criteria. Any conflicts in screening articles for inclusion were resolved by consensus among the full group. Articles reporting on secondary analyses (eg subgroup analyses from a larger clinical trial) were noted by the reviewers and subsequently included in the coding of the parent article, following the same methods as the review conducted by the NIH BCC.¹⁹

Process of Coding the Articles

The BCC treatment fidelity checklist was used to code the articles.¹⁷ At the full-text coding stage, each article was coded by two independent reviewers, at least one of whom was a doctoral-level reviewer. Coders identified the article characteristics and indicated the presence or absence of each item related to treatment fidelity within each of the five domains of the BCC treatment fidelity assessment tool: treatment design, provider training, delivery of treatment, receipt of treatment, and enactment of treatment skills.¹⁷ Treatment fidelity information was judged to be either present – meaning the article mentioned use of a particular treatment fidelity strategy; absent – meaning treatment fidelity information was omitted, preventing the coder from accurately assessing the scientific validity of the protocol; or not applicable – whereby the particular treatment fidelity strategy was not applicable to the reviewed article (eg no provider characteristics for web-based interventions). A coding manual, which included definitions of the treatment fidelity reporting strategies and examples from Borrelli et al.¹⁷ was updated and refined for tobacco treatment trials and used by the coders. Conflicts in coding between the pairs were resolved by consensus between the two reviewers. Monthly conference calls were held with the full group to discuss coding issues and to resolve discrepancies. A total of 19 reviewers worked in pairs to screen and extract data from the articles, and these pairs were consistent throughout the review process.

Data Extraction and Analysis

Data from eligible studies were captured using a REDCap²³ survey tool that was developed by the coleading authors (AR, RS) based on the BCC treatment fidelity coding tool.¹⁷ Following the full-text coding, the percentage of articles reporting the use of each strategy was computed as the ratio of the number of articles that coders deemed as using the particular strategy to the total number of articles for which the strategy was considered applicable. Therefore, if the strategy was not applicable to a particular study design (eg training providers would not be relevant for an intervention that did not involve providers), the study was not included in the denominator. Next, we compared between-group differences in proportions across studies that had a pharmacotherapy component compared with those that included behavioral interventions only using chi-square tests in Stata v15.0 (StataCorp, College Station, TX). Finally, we aggregated the proportion of adherence to treatment fidelity strategies by domain and reported means, standard deviations, medians, and inter-quartile ranges. These analyses were conducted

for all articles and by publication year intervals to evaluate change in reporting over time (ie 2006–2009, 2010–2014, and 2015–2018) since the publication of the BCC framework. Changes over time in treatment fidelity reporting were evaluated using repeated measures ANOVA, and univariate contrasts were evaluated using Tukey's honestly significant difference test.

Results

Search Results

Using the article search strategy outlined above, the initial literature search generated 4052 articles, of which 2468 were excluded after title review, 498 were excluded after abstract review, and 334 were excluded after full text review with reasons for exclusion depicted in [Figure 1](#). A total of 755 articles met all inclusion criteria, including three related studies not indexed in any of the aforementioned bibliographic databases that were added during the full text review. The PRISMA chart is depicted in [Figure 1](#). A reference list for all 755 coded articles is included as [Supplementary Material](#).

Summary of Included Publications

[Table 1](#) summarizes the characteristics of included articles. The mean number of participants across included studies was 735.1 (standard deviation = 2707.5; range = 9–55 568). The large majority of studies included tobacco abstinence or cessation as the primary tobacco treatment outcome ($n = 710$, 94.0%), followed by tobacco use reduction ($n = 23$, 3.0%). The most common study populations were adults who use tobacco recruited from the general population ($n = 347$, 45.9%), those with behavioral health or medical comorbidities ($n = 206$, 27.3%), and those from low socioeconomic status populations ($n = 38$, 5.0%). The most common behavioral interventions were individual in-person counseling ($n = 564$, 74.8%), self-help ($n = 310$, 41.0%), and telephone counseling ($n = 231$, 30.6%). Approximately two-thirds of the studies ($n = 506$, 66.9%) used pharmacotherapy. These studies tended to focus on reporting the details of the pharmacotherapy component of treatment and used brief behavioral counseling (<5–15 min/visit). The vast majority of studies included interventions that targeted patients ($n = 742$, 98.1%) as opposed to providers or systems, employed two or more conditions in the trial design ($n = 710$, 94.0%), and were randomized at the patient level ($n = 641$, 84.9%).

Frequency of Reporting Treatment Fidelity Strategies

[Table 2](#) provides the percentage of articles reporting each of the treatment fidelity strategies within each of the five domains. The proportion of reporting strategies varied within the individual domains, ranging from 5.2% to 96.3% from the Design domain, 1.9% to 24.7% in the Training domain, 2.6% to 32.1% in the Delivery domain, 5.2% to 44.2% in the Receipt domain, and 6.6% to 43.0% in the Enactment domain. The five most commonly reported treatment fidelity strategies were: report of treatment dose in the intervention condition (96.3%; Design), content of treatment (88.0%; Design), duration of contact over time (87.3%; Design), report of treatment dose in the control condition (84.0%; Design), if there was more than one intervention condition they were all described equally well (81.8%; Design), and mention of potential confounders that limit the ability to make conclusions (81.1%; Design). The five least commonly reported treatment fidelity strategies were: assessment of fit between the provider and the intervention at the hiring

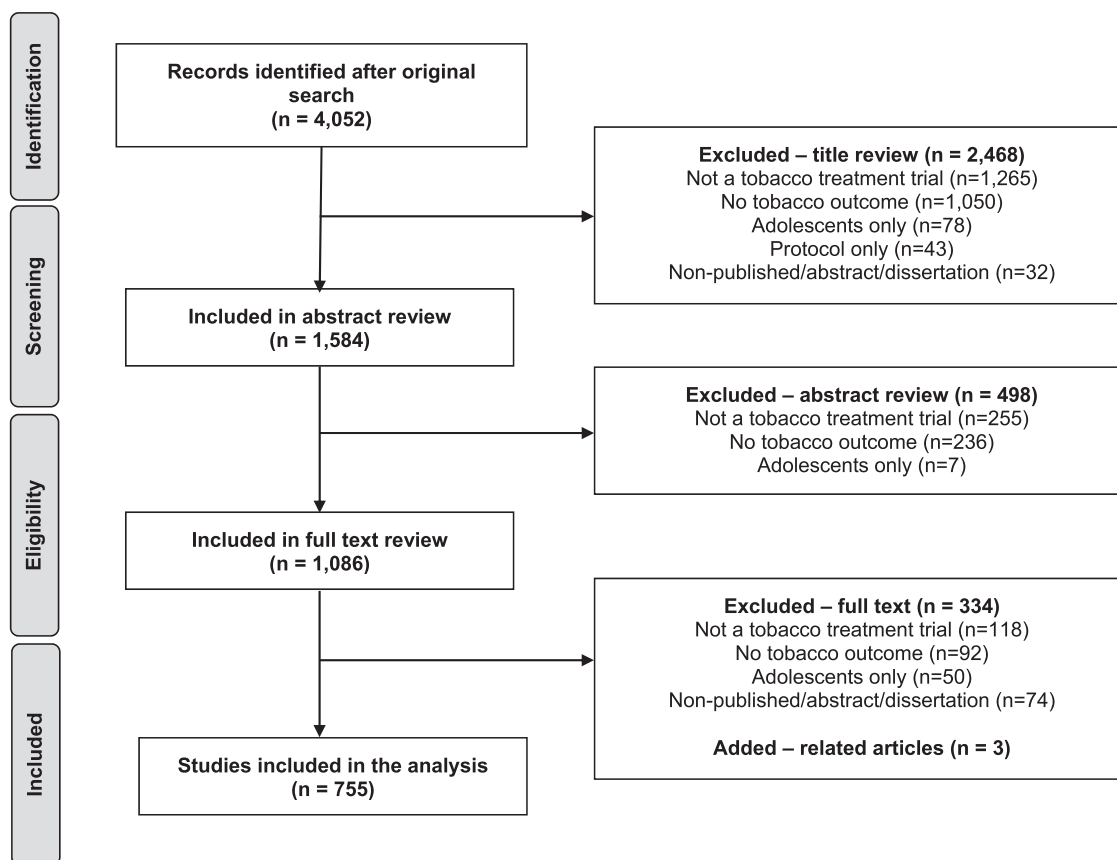


Figure 1. PRISMA flow diagram.

stage (1.9%; Training), non-specific treatment effects evaluated (2.6%; Delivery), mention of plans to address possible setbacks in implementation (5.2%; Design), strategy to improve subject comprehension of the intervention (5.2%; Receipt), and subject comprehension of the intervention assessed during the intervention period (6.6%; Receipt).

Studies that tested behavioral interventions alone had significantly higher proportions of treatment fidelity reporting compared with studies that also tested a pharmacotherapy component for the following strategies (p 's < .05, see Table 2): content of treatment (92.4% vs. 85.8%), mention of a theoretical model or clinical guideline as the basis for the treatment (73.6% vs. 56.9%), mention of potential confounders (86.4% vs. 78.5%), indication of how providers were trained (29.2% vs. 20.8%), a priori specification of treatment fidelity (eg providers adhere to delivering >80% of components; 16.9% vs. 7.9%), and reporting of post-implementation adaptations or modifications in treatment delivery (14.9% vs. 6.5%). Alternatively, studies with a pharmacotherapy component reported higher proportions than those without a pharmacotherapy component for the length (62.1% vs. 54.0%) and number (78.3% vs. 71.6%) of sessions in the intervention condition, as well as for the length of sessions in the control condition (47.0% vs. 35.6%).

Proportion Adherence to Treatment Fidelity Strategies Grouped by Domain

Table 3 displays the mean proportion adherence to reporting treatment fidelity strategies across all years and across domains. Across all included articles, the mean proportion of adherence to reporting

treatment fidelity strategies within the Design domain was 68%. The lowest mean proportion of adherence to reporting strategies was found in the Training domain, where, on average only 14% of applicable studies reported such fidelity strategies. The mean proportion of studies reporting strategies in the Delivery, Receipt, and Enactment domains was 15%, 16%, and 25%, respectively.

Table 3 also illustrates the change in reporting of treatment fidelity strategies by three different time blocks (2006–2009, 2010–2014, and 2015–2018). There was no statistically significant change in fidelity reporting observed across the 13-year time frame in which these studies were published.

High Levels of Treatment Fidelity Reporting

We identified a selection of exemplary articles that had 80% or greater adherence to the checklist in each domain as “high treatment fidelity reporting” studies, as was done previously.¹⁷ The number of articles that achieved 80% adherence to fidelity reporting or greater in a single domain was as follows: Design = 265 articles, Training = 4 articles, Delivery = 9 articles, Receipt = 4 articles, and Enactment = 23 articles. None of the included studies had 80% or greater adherence across all domains. Only eleven studies achieved 80% adherence across two or more domains (see Appendix B).

Discussion

Guided by the NIH BCC framework and using their tool for assessing and monitoring treatment fidelity in behavioral intervention trials, this scoping review specifically examined the adequacy of reporting

Table 1. Summary of Included Studies (*n* = 755)

	N	%
Number of subjects, mean (SD)	735.1 (2707.5)	
Primary (tobacco related) outcome:		
Abstinence or cessation	710	94.0
Reduction	23	3.0
Other	22	2.9
Study population		
General adults	347	45.9
People with behavioral health or medical comorbidities	206	27.3
Low socioeconomic status (income, education, etc.)	38	5.0
Racial and/or ethnic minorities	34	4.5
Pregnant women	33	4.4
Military and/or veterans	29	3.8
Women only	27	3.6
Youth and young adults	13	1.7
Older adults	8	1.1
Rural populations	6	0.8
Non-daily and/or light smoking	3	0.4
Sexual and gender minority	2	0.3
Intervention type		
Counseling – in-person	564	74.8
Pharmacotherapy	506	66.9
Self-help	310	41.0
Counseling – phone	231	30.6
Referral to quitline	67	8.9
Mobile health	60	7.9
Counseling – quitline	59	7.8
Contingency management	55	7.3
Education materials	34	4.5
Counseling – web and/or video	28	3.7
Exercise	23	3.0
System-level	14	1.8
Training	13	1.7
Harm reduction (eg e-cigarettes, low nicotine cigs)	12	1.6
Acupuncture or acupressure or hypnosis	10	1.3
Integrated voice response system	7	0.9
Electronic health record modification	7	0.9
Other	41	5.4
Intervention target		
Patient	741	98.1
Provider	8	1.1
System	3	0.4
Other	3	0.4
Trial design		
Two or more conditions	710	94.0
Single-arm	45	6.0
Unit of randomization		
Patient	641	85.0
Provider	10	1.3
Practice	25	3.3
Housing or neighborhood	6	0.8
No randomization	73	9.7
Year of publication		
2006–2009	209	27.6
2010–2014	312	41.3
2015–2018	234	31.0

strategies used to enhance treatment fidelity in tobacco treatment trials. Enhancing the rigor and reproducibility of tobacco treatment research is critical to achieving the goal of identifying effective interventions and translating them into real-world practice to reduce the personal and economic toll of tobacco across the globe. To that end, the SRNT Treatment Network Treatment Fidelity Workgroup identified 755

tobacco treatment articles that met inclusion criteria for this scoping review and used the NIH BCC checklist to code the articles' reporting of fidelity strategies across the five domains: Design, Provider Training, Treatment Delivery, Treatment Receipt, and Treatment Enactment.

The results suggest that while tobacco treatment researchers routinely demonstrate moderately high levels of reporting fidelity

Table 2. Assessment of Articles Reporting Use of Treatment Fidelity Strategies (*n* = 755)

	<i>n</i>	%	Medication, %		<i>p</i>
			No, <i>n</i> = 249	Yes, <i>n</i> = 506	
NIH BCC treatment fidelity strategy checklist					
Treatment design					
1. Provided information about treatment dose in the intervention condition(s)	755	96.3	96.0	96.4	.754
Length of contact session(s)	755	59.3	54.0	62.1	.030
Number of contacts	755	76.0	71.6	78.3	.040
Content of treatment	755	88.0	92.4	85.8	.009
Duration of contact over time	755	87.3	84.8	88.5	.141
2. Provided information about treatment dose in the control or comparison condition	755	84.0	85.6	83.2	.410
Length of contact session	755	43.2	35.6	47.0	.002
Number of contacts	755	56.7	54.0	58.1	.264
Content of treatment	755	74.3	76.8	73.1	.289
Duration of contact over time	755	71.3	69.2	72.3	.353
3. Mention of provider credentials	690	63.1	70.0	60.2	.016
4. Mention of theoretical model or clinical guidelines	755	62.4	73.6	56.9	<.001
5. Mention of potential confounders that limit the ability to make conclusions	755	81.1	86.4	78.5	.009
6. Mention of plans to address possible setbacks in implementation	755	5.2	4.8	5.3	.763
7. More than one intervention is described, they all described equally well	414	81.8	81.4	82.1	.858
Training providers					
1. Indication of how providers were trained	689	23.3	29.2	20.8	.017
2. Indication that provider training was standardized	689	21.3	22.9	20.7	.519
3. Measurement of provider skill acquisition post training (pre-field implementation)	689	7.4	10.0	6.3	.085
4. Indication of how provider skills were maintained	689	24.7	24.2	24.9	.839
5. A priori articulation of the characteristics being sought in a treatment provider	689	6.7	5.2	7.3	.317
6. Assessment of fit between the provider and the intervention at the hiring stage	689	1.9	1.0	2.3	.237
Delivery of treatment					
1. Method to ensure that the content of the intervention was being delivered as specified	755	24.2	27.3	22.7	.167
2. Method to ensure that the dose of the intervention was being delivered as specified	755	15.2	14.5	15.6	.678
3. Mechanism to assess if the provider actually adhered to the intervention plan	689	21.1	22.6	20.4	.502
4. Non-specific treatment effects (eg perceived warmth) evaluated	688	2.6	2.4	2.7	.784
5. Authors indicate that a treatment manual was used	688	32.1	30.5	32.8	.546
6. Authors indicate a plan for assessment of whether active ingredients were delivered	755	16.4	16.9	16.2	.817
7. Authors indicate a plan for assessment of whether proscribed components were excluded	755	7.8	8.4	7.5	.657
8. Authors indicate a plan for how will contamination between conditions be prevented	700	12.4	14.6	11.4	.220
9. Authors indicate a priori specification of treatment fidelity	755	10.9	16.9	7.9	<.001
10. Authors indicate any post-activation adaptations or modifications in treatment delivery	755	9.3	14.9	6.5	<.001
Receipt of treatment					
1. Subject comprehension of the intervention assessed during the intervention period	755	6.6	7.6	6.1	.435
2. Strategy to improve subject comprehension of the intervention	755	5.2	4.8	5.3	.763
3. Subject ability to perform the intervention skills assessed during the intervention period	755	44.2	44.2	44.3	.981
4. Strategy to improve subject performance of intervention skills during the intervention period	755	9.8	8.8	10.3	.531
5. Multi-cultural factors considered in the development and delivery of the intervention	755	13.8	13.6	13.8	.946
Enactment of treatment skills					
1. Subject performance of the intervention skills assessed	755	43.0	42.2	43.5	.733
2. Strategy to improve subject performance of the intervention	755	6.6	6.0	6.9	.643

Boldface indicates statistical significance (*p* ≤ .05).

Table 3. Proportion of Adherence to Treatment Fidelity Strategies Over Time

Domain	Design	Training	Delivery	Receipt	Enactment	Overall
Mean (SD)						
2006–2009	0.66 (0.31)	0.11 (0.09)	0.14 (0.09)	0.13 (0.14)	0.22 (0.22)	0.26 (0.28)
2010–2014	0.67 (0.30)	0.15 (0.10)	0.14 (0.08)	0.15 (0.15)	0.22 (0.24)	0.27 (0.28)
2015–2018	0.70 (0.29)	0.16 (0.11)	0.19 (0.11)	0.20 (0.19)	0.31 (0.31)	0.31 (0.28)
2006–2018	0.68 (0.30)	0.14 (0.10)	0.15 (0.09)	0.16 (0.16)	0.25 (0.26)	0.28 (0.28)
Median (IQR)						
2006–2009	0.81 (0.23)	0.12 (0.14)	0.13 (0.07)	0.07 (0.05)	0.22 (0.16)	0.14 (0.30)
2010–2014	0.80 (0.20)	0.16 (0.16)	0.13 (0.10)	0.09 (0.07)	0.22 (0.17)	0.14 (0.29)
2015–2018	0.81 (0.19)	0.15 (0.17)	0.17 (0.13)	0.13 (0.09)	0.31 (0.22)	0.20 (0.40)
2006–2018	0.81 (0.20)	0.14 (0.16)	0.14 (0.10)	0.10 (0.07)	0.25 (0.18)	0.16 (0.33)

IQR = inter-quartile range, SD = standard deviation.

strategies related to the description of the treatment condition (eg information about treatment dose, duration and content of active and control conditions; mention of provider credentials; theoretical or empirical underpinnings of treatment; potential confounders), fewer than half of the studies reported fidelity strategies related to the other four domains. There were also vastly more papers that reported at least 80% of the Design strategies ($n = 265$) compared to other strategies (range = 4–23). In contrast, there was a paucity of information reported on training and supervision of interventionists, and actual strategies used to improve delivery and uptake of the treatment condition. While 32% of studies reported that an intervention manual was used, checking fidelity of treatment delivery via audiotaping and coding counseling sessions was rarely reported. Interestingly, the variability in reporting of Design strategies varied based on whether pharmacotherapy was included in the treatment. This suggests that tobacco researchers may have different perspectives about what is necessary to report regarding behavioral treatment and this may differ by whether the study also involves pharmacotherapy.

The mean proportion of adherence to reporting treatment fidelity strategies was 36.8% and only 11 out of 755 (1.5%) studies reviewed achieved $\geq 80\%$ reporting across >1 treatment fidelity domain. In comparison, the mean proportion of adherence to treatment fidelity strategies in the broader literature of behavioral interventions from 1990 to 2000 as reported by the NIH BCC Consortium was 55%, and 15.5% of articles achieved greater than or equal to 80% adherence. Although the overall adherence to the BCC recommendation was higher in the broader behavioral interventions trials than in behavioral tobacco trials (15.5% vs. 1.5%), the pattern of adherence by domain was similar such that the mean proportion of reporting treatment fidelity strategies in the broader behavioral intervention trials was highest in terms of Design = 80% and relatively low across the other domains of Training = 22%; Delivery = 35%, Receipt = 49%, and Enactment = 57%.¹⁷

Despite more than 10 years since the publication of the NIH BCC Framework and checklist, these findings illustrate much variability and lack of consistency in reporting of strategies to improve treatment fidelity in tobacco treatment trials. We found no evidence of an improvement in fidelity reporting over time in behavioral tobacco treatment studies. This finding suggests the need for raising awareness and greater attention from the tobacco treatment field to the importance of assessing treatment fidelity and reporting these strategies in publications related to the design and main outcomes of such clinical trials.

It is unclear why there is a noteworthy deficit in reporting fidelity strategies in the behavioral tobacco treatment literature. It could be that the use of evidence-based “problem-solving” and “support” have

become such standards in the field that authors no longer see the value in reporting all of the specific details related to behavioral treatment fidelity. With only 59% of intervention studies (43% of control or comparison) specifying counseling contact time, this creates ambiguity in the exact dose of behavioral treatment, with even greater ambiguity in reporting dose of control or comparison conditions. For example, it is often unclear if patients are being delivered a brief 5As problem-solving intervention which may differ substantially compared to more in-depth cognitive behavioral treatment (CBT). The observed ambiguity regarding reporting of usual care control or comparison conditions is particularly important with the increased publication of pragmatic clinical trials conducted in real-world settings. Another explanation might be that studies evaluating pharmacotherapies are more focused on describing the pharmacotherapy intervention than the ancillary behavioral intervention. In more than two-thirds of the studies that included pharmacotherapy, the dose of the behavioral component was small, limiting the generalizability of this work in settings with tobacco treatment specialists and greater intensity of behavioral tobacco interventions. A third possibility is that as leading peer-reviewed journals have placed an increasing limitation on article length and study designs have become more complex, the critical issues of treatment fidelity may have been sacrificed. A final possibility is that many tobacco researchers may be less familiar than researchers in other behavioral treatment areas with recommendations for assessment and monitoring of treatment fidelity in clinical research trials. Regardless of the reason or reasons for this deficit, it is clear that tobacco treatment researchers are not consistently reporting key strategies for assessing treatment fidelity in tobacco treatment trials.

Given the findings of this scoping review, the SRNT Treatment Network Treatment Fidelity Workgroup has developed several recommendations intended to improve the reporting of behavioral tobacco treatment interventions:

- (1) *Journals should consider establishing minimum requirements for reporting adherence and treatment fidelity in cessation clinical trials.* In addition to completion of the appropriate CONSORT checklist, journals should consider requiring that clinical trial investigators use the NIH BCC checklist for greater transparency in reporting of treatment fidelity in behavioral tobacco treatment trials. Dissemination of the NIH BCC treatment fidelity checklist (Table 4)¹⁷ for authors and reviewers of tobacco-specialty journal manuscripts that report on findings from tobacco treatment clinical trials would assist editors in achieving high quality reporting of tobacco treatment trial results. Additionally, journals should consider requiring protocol papers to include pertinent information about their planned measurement of treatment fidelity.

Table 4. NIH BCC Consortium Treatment Fidelity Checklist with Tobacco Treatment-specific Examples Provided

Treatment fidelity strategy	Examples
Treatment design	
1. Treatment dose in the intervention condition(s)	
Length of contact session(s)	30-Minute counseling session
Number of contacts	6 Counseling sessions total
Content of treatment	Cognitive behavioral therapy
Duration of contact over time	1 Session per week for 5 weeks
2. Treatment dose in the control or comparison condition	
Length of contact session	10-Minute telehealth call
Number of contacts	3 Text messages per day for 2 weeks
Content of treatment	Motivational Interviewing
Duration of contact over time	12-Week treatment phase
3. Provider credentials	Master of Social Work (MSW), Tobacco Treatment Specialist (TTS)
4. Theoretical model or clinical guidelines	Self Regulation Theory, Transtheoretical Model, U.S. Clinical Practice Guidelines
5. Potential confounders that limit the ability to make conclusions	A change in clinics' tobacco treatment practices during the study implementation
6. Plans to address possible setbacks in implementation	With staff turnover, how will new providers be trained to ensure consistency
7. If more than one intervention, describe all equally well	
Training providers	
1. How providers were trained	Individual, vs. group, number of trainings over how many hours
2. Standardization of provider training	Standardized training manuals and/or materials, interventionists train together
3. Measurement of provider skill acquisition post training (pre-field implementation)	Written test or observation
4. How provider skills were maintained	Supervision, evaluation of recorded sessions, provider self-report
5. Characteristics being sought in a treatment provider	Education level, years experience
6. Fit between the provider and the intervention at the hiring stage	Someone who is psychoanalytic for a cognitive behavioral therapy trial
Delivery of treatment	
1. Method to ensure that the content of the intervention was being delivered as specified	Review of audio or video recordings
2. Method to ensure that the dose of the intervention was being delivered as specified	Track timing of interventions
3. Mechanism to assess if the provider actually adhered to the intervention plan	Review of audio or video recordings
4. Evaluation of non-specific treatment effects (eg perceived warmth)	Provider- and client-focused assessment of alliance and rapport; interpersonal relationship; interaction style
5. Treatment manual	Or checklist or protocol
6. Plan for assessment of whether active ingredients were delivered	Video or audio recording and coding
7. Plan for assessment of whether proscribed components were excluded	Video or audio recording and coding
8. Plan for how will contamination between conditions be prevented	Unique providers for each treatment arm
9. A priori specification of treatment fidelity	Audio recordings of sessions and agreement among coders of intervention delivered
10. Post-activation adaptations or modifications in treatment delivery	Reduction in number of text messages due to participant reports of excessive texts
Receipt of treatment	
1. Assessing subject comprehension of the intervention during the intervention period	Teach back activity
2. Strategy to improve subject comprehension of the intervention	Offered in oral and visual formats
3. Assessing subject ability to perform the intervention skills during the intervention period	Role playing
4. Strategy to improve subject performance of intervention skills during the intervention period	Specific skill practice assignments
5. Multi-cultural factors considered in the development and delivery of the intervention	Language intervention is provided in, incorporation of culturally relevant constructs (eg importance of family)
Enactment of treatment skills	
1. Subject performance of the intervention skills assessed	If using short-acting NRT is viewed as a coping skill for cravings in a risky situation, this could be better assessed and reported
2. Strategy to improve subject performance of the intervention	Setting a timer to prompt pharmacotherapy use

- (2) *Investigators should consider giving greater attention to the assessment and monitoring of treatment fidelity in the design and transparent reporting of the main outcomes of their clinical trials.* Encouraging greater compliance with the NIH BCC Framework and checklist and promoting best practices for measurement and reporting of multiple measures of treatment adherence and fidelity would likely improve the rigor and reproducibility of tobacco treatment trials. With respect to reporting treatment fidelity, if journal word limits do not allow for all components of the checklist to be included, the treatment fidelity details that would have the greatest impact on internal validity and Type 1 and Type 2 error should be included. The remainder of the treatment fidelity details could then be included in an appendix or supplement. Inclusion of all treatment fidelity details remains our recommendation, but we acknowledge that there may be word limitations that would preclude all treatment fidelity aspects from being reported in the main manuscript document.
- (3) *Grant reviewers for extramural funding agencies should consider these recommendations for measuring and reporting treatment fidelity when considering the rigor and reproducibility of proposed tobacco treatment trials.*

We posit that the broad adoption of consensus guidelines for reporting of tobacco treatment clinical trials will accelerate advances in the field of tobacco treatment and ultimately reduce the burden of tobacco use and dependence. There is growing concern regarding poor replicability of findings from behavioral interventions.²⁴ Treatment effects may be influenced and/or obscured by variations in the extent and quality of treatment delivery. The cause of such variability is difficult to interpret if treatment fidelity is neither assessed nor reported. Assessment and reporting of treatment fidelity may be particularly important when tobacco treatment is delivered by multiple providers in real world settings and when conducting multi-site trials to evaluate the potential for provider-level and site-level variation on outcomes,²⁵ as these trials may include more inherent natural variability and a higher likelihood for drift compared to single-site trials using study interventionists. These findings also underscore the importance of adequately measuring fidelity for every treatment arm (including the “control” and usual care arms) within a study to provide confidence that the observed outcomes are related to proposed treatment mechanisms rather than differences in the quality of the delivery itself or the presence or absence of unmeasured “active” ingredients in the control condition.¹⁹ Finally, we found very limited attention to the fidelity of brief behavioral treatment commonly delivered in pharmacotherapy trials. Greater attention to the reporting of brief behavioral treatment will likely assist with the interpretation of variation in effectiveness outcomes and advance understanding of the effectiveness of pharmacotherapies in clinical settings.

There are several noteworthy limitations to this review. First, this scoping review does not address the challenges of reporting trials when the primary intervention or outcome is provider behavior change or systems change due to the small number of such studies. Further attention is needed to establish meaningful domains for measuring and reporting implementation of provider and systems-level tobacco treatment interventions. Second, there has been much activity in the development and evaluation of trials testing digital or mobile behavioral interventions. Further attention is needed for advancing the understanding of implementation fidelity for digital or mobile health tobacco treatment interventions by considering specific measurement instruments and data collection methods that may be more appropriate for digital media. Third,

clinical trials reporting on interventions targeting adolescent tobacco users were excluded and therefore further consideration must be given to the adequacy of reporting implementation of interventions targeting adolescents. Similarly, it is plausible that treatment adaptations may be necessary for some special populations,²⁶ such as those with cooccurring behavioral health conditions who are largely understudied. However, only a few of the papers reviewed reported any systematic description of minor or major deviations from the clinical protocol. Greater attention to reporting of the nature and frequency of treatment adaptation would advance understanding of the extent to which a specific treatment can be modified for a given special population or setting and still maintain treatment effectiveness.²⁷ This latter point is particularly germane when reporting the effectiveness of treatments with proven efficacy in controlled research settings and subsequently delivered by practicing clinicians in real world settings. Fourth, the literature review for this project was conducted in July 2018; however, it is plausible that articles published in early 2018 may have inadvertently been excluded due to delays in database indexing. Finally, in summarizing study characteristics, we combined study populations of patients who smoke or use other tobacco products with behavioral health conditions and with those with medical comorbidities into one group; however, further attention is warranted for understanding any specific challenges of implementation fidelity for these two subgroups separately as these special populations may require adaptation in tobacco treatment delivery.

In conclusion, this scoping review found that there was limited reporting of behavioral tobacco treatment fidelity strategies during the last 15 years, even following the publication of the NIH recommendations in the form of the NIH BCC checklist. We see much value in disseminating these findings and the accompanying trial reporting recommendations via webinars and short course module sponsored by the SRNT University platform. The SRNT Treatment Network Treatment Fidelity Workgroup recommends that: (1) tobacco treatment researchers use the BCC checklist (Table 4) in their planning and reporting of tobacco treatment trials; (2) journals that publish tobacco treatment research and global funding institutions consider promoting the NIH BCC strategies as part of their review process. These changes have the potential to significantly improve the rigor and reproducibility of tobacco treatment research and thereby improve its impact on public health.

Supplementary Material

A Contributorship Form detailing each author’s specific involvement with this content, as well as any supplementary data, are available online at <https://academic.oup.com/ntr>.

Funding

This research was supported by National Cancer Institute grant K07CA214839 (AR) and a Core Core Facility of the Memorial Sloan Kettering Cancer Center Support Grant from the National Cancer Institute (P30CA008748).

Declaration of Interests

None.

Acknowledgment

The authors acknowledge Terese Kuba, Chris Ripley, Renae L. Borkowski, Cassandra J. Astemborski, Austin Boone, Haven Garcia, Haille Walker, Alexus

Brown, and Nora Kassis for their contribution to coding articles. The authors would also like to thank Andrea Weinberger, PhD of the Treatment Network Advisory Committee and Benjamin Toll, PhD of the SRNT Board of Directors for their helpful comments on a draft of this manuscript. RS and AR contributed equally to this work.

References

- World Health Organization. *Tobacco Fact Sheet*. <https://www.who.int/news-room/fact-sheets/detail/tobacco>. 2020. Accessed November 12, 2020.
- Fiore MC, Jaén CR, Baker TB, et al. *Treating tobacco use and dependence: 2008 update*. Rockville, MD: US Department of Health and Human Services; 2008.
- European Network for Smoking and Tobacco Prevention. *Guidelines for Treating Tobacco Dependence*. 2017. <http://elearning-ensp.eu/mod/page/view.php?id=36>. Accessed December 14, 2020.
- Public Health England. *Models of Delivery for Stop Smoking Services: Options and Evidence*. London, England: PHE Publications; 2017.
- United States Public Health Service, Office of the Surgeon General, National Center for Chronic Disease Prevention and Health Promotion, Office on Smoking and Health. *Smoking Cessation: A Report of the Surgeon General*. Washington (DC): US Department of Health and Human Services; 2020.
- Lancaster T, Stead LF. Individual behavioural counselling for smoking cessation. *Cochrane Database Syst Rev*. 2017;3:CD001292.
- Stead LF, Carroll AJ, Lancaster T. Group behaviour therapy programmes for smoking cessation. *Cochrane Database Syst Rev*. 2017;3:CD001007.
- Wray JM, Funderburk JS, Acker JD, Wray LO, Maisto SA. A meta-analysis of brief tobacco interventions for use in integrated primary care. *Nicotine Tob Res*. 2018;20(12):1418–1426.
- Lindson-Hawley N, Thompson TP, Begh R. Motivational interviewing for smoking cessation. *Cochrane Database Syst Rev*. 2015;(3):Cd006936.
- Mottillo S, Filion KB, Bélisle P, et al. Behavioural interventions for smoking cessation: a meta-analysis of randomized controlled trials. *Eur Heart J*. 2009;30(6):718–730.
- Oikonomou MT, Arvanitis M, Sokolove RL. Mindfulness training for smoking cessation: a meta-analysis of randomized-controlled trials. *J Health Psychol*. 2017;22(14):1841–1850.
- Siddiqi K, Khan A, Ahmad M, et al. Action to stop smoking in suspected tuberculosis (ASSIST) in Pakistan: a cluster randomized, controlled trial. *Ann Intern Med*. 2013;158(9):667–675.
- 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(5):587–592.
- Durlak JA, DuPre EP. 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(3-4):327–350.
- Lorenatto F, West R, Christopherson C, Michie S. Assessing fidelity of delivery of smoking cessation behavioural support in practice. *Implement Sci*. 2013;8:40.
- Hawe P, Shiell A, Riley T. Complex interventions: how “out of control” can a randomised controlled trial be? *BMJ*. 2004;328(7455):1561–1563.
- Borrelli B, Sepinwall D, Ernst D, et al. A new tool to assess treatment fidelity and evaluation of treatment fidelity across 10 years of health behavior research. *J Consult Clin Psychol*. 2005;73(5):852–860.
- Peters DH, Tran NT, Adam T. *Implementation Research in Health: A Practical Guide*. Geneva, Switzerland: World Health Organization; 2013.
- Borrelli B. The assessment, monitoring, and enhancement of treatment fidelity in public health clinical trials. *J Public Health Dent*. 2011;71(suppl 1):S52–S63.
- Lorenatto F, West R, Stavri Z, Michie S. How well is intervention content described in published reports of smoking cessation interventions? *Nicotine Tob Res*. 2013;15(7):1273–1282.
- Knittle K. Fidelity in intervention delivery: a rough field guide. *Eur Health Psychol*. 2014;16(5):190–195.
- Bramer WM, Giustini D, de Jonge GB, Holland L, Bekhuis T. De-duplication of database search results for systematic reviews in EndNote. *J Med Libr Assoc*. 2016;104(3):240–243.
- Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)—a metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform*. 2009;42(2):377–381.
- Maxwell SE, Lau MY, Howard GS. Is psychology suffering from a replication crisis? What does “failure to replicate” really mean? *Am Psychol*. 2015;70(6):487–498.
- Baer JS, Ball SA, Campbell BK, Miele GM, Schoener EP, Tracy K. Training and fidelity monitoring of behavioral interventions in multi-site addictions research. *Drug Alcohol Depend*. 2007;87(2-3):107–118.
- Steinberg ML, Weinberger AH, Tidey JW. Non-pharmacological treatments for tobacco users with mental health symptoms. *Nicotine Tob Res*. 2019;21(5):557–558.
- Wiltsey Stirman S, Kimberly J, Cook N, Calloway A, Castro F, Charns M. The sustainability of new programs and innovations: a review of the empirical literature and recommendations for future research. *Implement Sci*. 2012;7:17.