Journal of Psychiatry
ISSN: 2378-5756

Prowse, et al., J Psychiatry 2015, 18:2
DOI: 10.4172/2378-5756.1000258

Research Article Open Access

Treatment Fidelity Over the Last Decade in Psychosocial Clinical Trials Outcome Studies: A Systematic Review

Phuong-Tu D Prowse[1]*, Tricia Nagel[2], Graham N Meadows[3] and Joanne C Enticott[4]

[1]*Menzies School of Health Research, Department of Psychiatry, Monash University, Melbourne, Australia*
[2]*Menzies School of Health Research and Charles Darwin University, PO Box 40196, Casuarina, NT 0811, Australia*
[3]*Department of Psychiatry, Monash University, Melbourne, Australia*
[4]*Department of Psychiatry, School of Clinical Sciences at Monash Health, Monash University, Melbourne, Australia*

Abstract

Background: Treatment fidelity tools are frequently used in clinical trials, promoting treatment consistency and therefore validity of trial findings. However, treatment fidelity procedures have not been included within international clinical trial guidelines such as the Consolidated Standards of Reporting Trials (CONSORT 2010).

Aim: This study systematically reviews psychological clinical trials that incorporate Treatment Fidelity procedures and appraises their implementation using the Implementation of Treatment Integrity Procedures Scale (ITIPS).

Method: Using the PRISMA Checklist as a guide for systematic review, a comprehensive search of the Medline, PsychINFO, Ovid, Cochrane Library, Scopus, PUBMED databases for the period 2004 to 2014 resulted in retrieval of 3186 potential articles. Thirty-two studies meeting inclusion criteria were analysed against the ITIPS.

Results: Sixteen studies were assessed as 'approaching adequacy' in implementing Treatment Fidelity procedures as measured by the ITIPS scale, 8 studies were assessed as 'adequate' whilst a further 8 studies were deemed 'inadequate' against this assessment. Treatment Fidelity tools generally increased the intensity of the intervention or program within which they were used, resulting in improved levels of Treatment Fidelity.

Conclusion: Current evidence supporting the inclusion of Treatment Fidelity tools is limited since there have been relatively few published studies examining the effectiveness of Treatment Fidelity tools. Further research into the efficacy, feasibility and measurement of Treatment Fidelity in implementing treatments is recommended, in tandem with additions to the CONSORT Guidelines to better support the inclusion of Treatment Fidelity procedures within clinical trials.

Keywords: Treatment fidelity; Integrity; Adherence; Competence; Scales; Mental health research

Review Contributions to Existing Research

- Assesses current evidence and identifies areas for future research
- Highlights several key strengths of Treatment Fidelity in improving quality assurance and implementation strategies for clinical trials
- Finds limited use of Treatment Fidelity procedures within current clinical trials
- Provides evidence to support the inclusion of Treatment Fidelity tools in the CONSORT Guidelines.

Introduction

Treatment fidelity concepts

The last decade has witnessed rapid development in Treatment Fidelity research tailored to enhance therapy implementation including progress in terms of fidelity definitions, strategies, and approaches to maintenance. Accurate assessment of the effectiveness of therapy requires knowledge of the degree of Treatment Fidelity within the program under evaluation. Treatment Fidelity has been an important topic in the psychosocial research as it thus has important implications for clinical practices. It provides evidence as to whether the treatment being investigated was implemented in accordance to recommended protocols. Without this evidence it is difficult to ascertain the effectiveness of any given therapy. Early conceptualization of Treatment Fidelity, also referred to as 'treatment integrity' or 'treatment purity', was described as treatment delivered as intended [1]. Subsequently, 'treatment differentiation' gained favour amongst researchers as a descriptor of Treatment Fidelity; this referred to whether or not the treatment implemented differed from its intended manner [2-4]. Later again, 'treatment receipt' emerged as a separate element of Treatment Fidelity. Defined as whether the client comprehended and used the treatment skills taught during the sessions [5]. As the field evolved, 'treatment enactment' was identified as a Treatment Fidelity element that evaluated whether the client applied skills learnt in treatment to their daily life [6,7]. Leichsenring and colleagues [1], expanded further defining Treatment Fidelity as a means of exploring: (i) whether a treatment delivered is representative of the theoretical constructs and mechanisms presumed to underpin its purpose, (ii) the extent to which treatment effects are causally attributed to the treatment implemented and (iii) whether these methods are generalizable in the clinical setting.

***Corresponding author:** Phuong-Tu Prowse, Menzies School of Health Research, Department of Psychiatry, Monash University, Melbourne, Australia, Tel: 08 89228196; E-mail: ppro4@student.monash.edu

Received February 10, 2015; **Accepted** February 27, 2015; **Published** March 05, 2015

Citation: Prowse PTD, Nagel T, Meadows GN, Enticott JC (2015) Treatment Fidelity Over the Last Decade in Psychosocial Clinical Trials Outcome Studies: A Systematic Review J Psychiatry 18: 258. doi:10.4172/2378-5756.1000258

Fidelity tools

The introduction of treatment manuals allowed interventions under empirical investigation to be operationalised to best support therapy delivery in line with designed treatment structures [8,9]. Manualisation enabled an intervention to be monitored for Treatment Fidelity levels according to the research protocols [4,7]. In the area of psychology and therapy implementations, a number of researchers have described manuals as a reliable and cost effective mechanism to support Treatment Fidelity to maximize targeted outcomes [10-13]. However, several researchers argue mutualized treatments do not ensure effective delivery of the treatment [14-18]. This helps to explain the introduction of adherence and competence scales to enhance the assessment and measurement of Treatment Fidelity.

The value of measuring adherence and competence to determine the quality of Treatment Fidelity is a fundamental consideration [19-21]. Adherence is expressed as the delivery of a key component or technique of the treatment [9]. In contrast, competence measures assess the skills or accuracy with which the treatment is implemented [22,23]. Quantitative and qualitative research methods are pivotal tools for the investigation of these complex Treatment Fidelity phenomena [24-27]. Despite advances in the field of Treatment Fidelity, studies of treatment adherence and competence continue to reveal inconsistencies between treatment and outcome [28]. Perepletchikova and Kazdin [29] suggested that adequate Treatment Fidelity measures are essential in research settings to explain such inconsistencies. Treatment Fidelity procedures can assist in exploration of associations between outcomes and features of the intervention, or the therapist.

Treatment fidelity in randomised clinical trials

Randomised controlled trials (RCT) are a rigorous means of describing and determining the existence of cause and effect relationships between treatment and outcome, and assessing the cost-effectiveness of a treatment. When clinical trials are designed, delivered, analysed and interpreted, generalisation and attribution of findings are possible. However, published clinical trials can yield biased results, lack methodological rigour, and may provide incomplete reporting, thus limiting the opportunity for replication of studies [30]. In 1998, these known limitations were a catalyst for the development of the Consolidated Standards of Reporting Trials (CONSORT), which were subsequently updated in 2001 and 2010 [31,32]. The CONSORT were developed to help improve the quality of clinical trials in terms of accuracy, clarity, transparency, research design and findings [31,33].

Members of the National Institute of Health (NIH) Behavioural Change Consortium (BCC) offered a comprehensive Treatment Fidelity Framework that included a five-part theoretical model for Treatment Fidelity in clinical trials [34]. The model suggested the following factors be considered when designing a trial; (i) study design; (ii) training: specific competencies required for successful delivery of the intervention for training design; (iii) delivery: processes that monitor and maintain quality of delivery; (iv) receipt: processes that ensure that participants understand the information provided in the intervention; and (v) enactment: processes to monitor and improve the ability of participants to perform treatment-related cognitive and behavioural strategies in their daily lives [34,35].

Measuring and assessing Treatment Fidelity provides a method to document deviations within and from an intended model and enhances internal and external validity, and reliability of behavioural research interventions [5,36-38]. Whilst the assessment of Treatment Fidelity is important, it can also be resource-intensive [39]. It has the potential to add an enhanced dimension to clinical trial implementation [40,41]. Perepletchokova and colleagues [23] emphasised that fidelity procedures and measures are central to the delivery of successful clinical trials. They highlighted four key areas of Treatment Fidelity in clinical trials that included: establishment of fidelity (e.g. specification of protocol, structured training of therapists and continued monitoring of therapist's adherence to the prescribed procedures); assessment of fidelity (e.g. assessed via direct observations in areas of treatment adherence, therapist competence and treatment differentiation), evaluation of fidelity (e.g. use of adherence and competence scales) and reporting of fidelity (e.g. overall integrity of treatment the extent to which all components were correctly implemented according to the manual, and component integrity- consistently implementing all treatment component across sessions). This measure was referred to as the Implementation of Treatment Integrity Procedures Scale (ITIPS) and was designed to evaluate the extent clinical trials addressed these four defined areas [23].

Treatment fidelity and CONSORT

At present many elements of Treatment Fidelity are absent from the revised CONSORT statements and explanations [42]. Persche and Page [25] highlighted that though the CONSORT Guidelines are regularly used in clinical trials, they are deficient in the area of supporting the attainment of high levels of Treatment Fidelity, potentially detracting from the effectiveness of the delivered intervention. Identified deficiencies of the CONSORT Guidelines include: inability to capture the provision of clinician professional development and the inability to record the attainment of clinically significant results. Moreover, the guidelines are ineffective in terms of assessing treatment delivery, treatment receipt and treatment enactment [34,43]. The identified weaknesses within the present CONSORT guidelines have direct implications for client care, as the attainment of high levels of Treatment Fidelity is often critical for program goals and replication across multiple sites [44].

The inclusion of Treatment Fidelity tools can assist with adequate testing of a proposed hypothesis, and can enhance statistical power for measures of internal validity, From the point of view of translational research, it enhances the ability to replicate the treatment in other studies, promptly disseminate the treatment, and potentially to maximise successful patient/client outcomes [45,46]. In contrast, a lack of attention to Treatment Fidelity implementation may lead to poor standardisation within and across treatments in clinical trials and will contribute to an inflated error variance, decreased statistical power and increased likelihood of a Type II Error [47,48].

Aim

This systematic review aims to identify how Treatment Fidelity has been implemented in clinical trials to help contribute to improved understanding of current trends.

The review appraises psychosocial clinical trials that specifically investigate Treatment Fidelity over the last decade. It uses the Implementation of Treatment Integrity Procedures Scale (ITIPS) [23] to critically appraise and synthesize evidence in terms of:

1. Use of Treatment Fidelity procedures within clinical trials of psychological interventions

2. Alignment of clinical trial Treatment Fidelity procedures with the Implementation of Treatment Integrity Procedures Scale.

Method

Types of studies

Only randomised control trials published in English between 2004 and 2014 were included. When a trial did not report randomisation but was described as "double-blind" and the demographics details of each group were similar, the trial was deemed to be randomised. The authors excluded quasi-randomised studies, but studies that employed "cluster randomisation" (such as randomisation by clinician or practice) were included.

Search strategy

Prior to commencing the systematic review, a preliminary search of the Database of Abstracts of Reviews of Effects (DARE), and the Cochrane Database of Systematic Reviews (CDSR) confirmed no similar systematic reviews had been published. The following electronic library databases were investigated by two research assessors using the PRISMA (2009) Model of Systematic Review: The Cochrane Central Register of Controlled Trials on the Cochrane Library (January 2004-January 2014); PubMed (January 2004-January 2014); Ovid Medline (January 2004-January 2014); PsychINFO (January 2004-January 2014); CINAHL Plus (January 2004-January 2014); Scopus (January 2004-January 2014). Search terms included treatment fidelity, integrity, intervention integrity, adherence, competence, and implement, scale, assessment, and monitor and outcome measure.

Two assessors (Clinical Psychologist Ted Graham and first author-PP) independently screened titles and abstracts based on the research questions, study design, specified population, intervention, and outcome(s). Each individual article was assessed using the inclusion criteria of randomised control trials, mental health treatments, English and full text articles. The exclusion criteria consisted of poor quality results, non-randomised control trial, not peer reviewed, meta-analysis/systematic reviews, trials not explicitly assessing Treatment Fidelity, letters, opinions, inadequate considerations of confounders, development of scale articles, tools used not validated/reliable and qualitative studies. The reference lists of included studies and reviews were searched to help identify further relevant studies. If the assessors agreed an assessed trial did not supply sufficient pre-requisite data, that study was omitted from the review. Publications were viewed individually and any double reporting recorded. Figure 1 shows the results of different search engines [49].

Data extraction

The two assessors independently extracted data from the selected 32 articles using the ITIPS. To maintain data integrity this information was cross-checked by each assessor. In the event consensus could not be reached pertaining to the rating of items within the scale, Tricia Nagel (TN-second author) assisted in making a final determination. Decisions requiring clarification or data extraction challenges were documented for future discussion.

Measurement used for data management

The PRISMA (2009) Checklist for Systematic Review [31] is a structured way to summarise literature reviews, which was further complemented by the inclusion of the Implementation of Treatment Procedures Scale (ITIPS). The ITIPS was designed by Perepletchikova and colleagues [23] to promote a common language to best position researchers to understand, measure and define Treatment Fidelity. It provides a framework to systematically evaluate and code Treatment Fidelity in clinical trials [23,50]. The ITIPS consists of 22-items covering domains of establishment (use of treatment manuals), assessment (treatment adherence, therapist competence, evaluation (therapist reactivity (e.g. therapist performance altered due to awareness of being observed) and reporting (professional development of therapists and raters of Treatment Fidelity in outcomes studies). Each of the 22-items has a potential rating scale of four points. Total scores range from 22 to 88. Higher scores indicate more adequate implementation of Treatment Fidelity procedures (e.g., "Training strategies of therapists," where 1: not trained, 2: authors mentioned that therapists were trained but no other information was provided, 3: used indirect strategies, and 4: used direct strategies). The establishing treatment fidelity domain (6 items) refers to how researchers conceptualize fidelity (e.g., in terms of adherence and/or competence), as well as the extent to which they provide a detailed treatment manual to therapists, train and supervise them. The assessing treatment fidelity domain (7 items) refers to the assessment of treatment fidelity via direct, indirect, or hybrid strategies; measurement of therapist treatment adherence as well as competence; and employment of fidelity measures with good psychometric properties (i.e., validity and reliability). The evaluating treatment integrity domain (5 items) refers to procedures such as ensuring the accuracy of the representation of the obtained fidelity data, training of raters, assessing inter-rater reliability, and controlling for measure reactivity. The reporting treatment fidelity domain (4 items) refers to procedures such as reporting numerical data; reporting overall, component and session fidelity; and reporting the implementation of various fidelity procedures. Therapist treatment adherence and therapist competence aspects of fidelity (6 items each) encompass how the terms were defined, assessed, evaluated, and reported [23]. To reduce the risk of

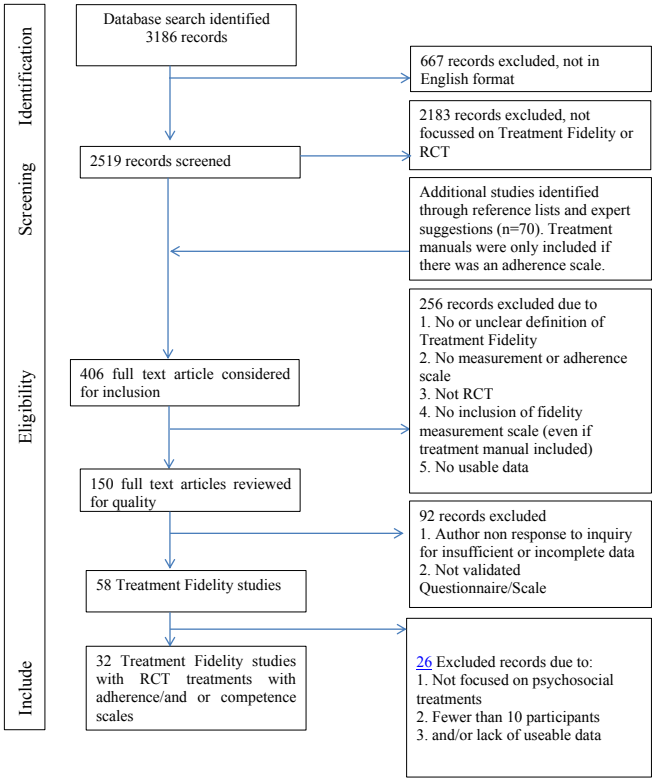

Figure 1: Clinical trials vetting process employed using The PRISMA Systematic Review flow diagram.

bias associated in over-estimating the effects assessors rated the articles independently, prior to assigning ITIPS scores. Assessor scores were averaged and recorded on an Excel Spreadsheet for analysis.

Treatment Fidelity procedures were categorised as determined by the ITIPS guidelines: adequate (AD); approaching adequate (AA); and inadequate (IA). For each item a score of 1 or 2 was assumed to reflect inadequate implementation of integrity procedures as the clinical trial showed either no evidence or talked only in broad terms of Treatment Fidelity; a score of 3 indicated that implementation approached adequacy as the clinical trial had provided some data that measured Treatment Fidelity ; and a score of 4 designated adequate implementation of fidelity procedures as the clinical trial had provided detailed data that measured Treatment Fidelity. Because there were 22 items on the ITIPS, studies were classified as implementing fidelity procedures (a) inadequately if the study's total score ranged between 22 and 44; (b) in a manner approaching adequacy if the total score ranged between 45 and 66; and (c) adequately if the total score exceeded 66. This strategy was also utilized for evaluation of the adequacy of the Treatment Fidelity procedures for the four domains and the two aspects of fidelity. Statistical analysis was conducted using SPSS for Windows.

Results

Figure 1 illustrates the flow of information through the different phases of the systematic review and maps out the number of records identified, included, and reasons for exclusions. Thirty two studies met all selection criteria for inclusion. 3224 records were excluded due to: non-RCT design, treatment focus not psychosocial, lack of appropriate fidelity measures or treatment fidelity definition, fewer than 10 participants, and/or lack of useable data.

Establishing treatment fidelity

The majority of studies (71.9%) reported use of a treatment manual to support the therapist as shown in Table 1. However, more than a quarter (28.1%) of studies made no reference to the use of a treatment manual. Of the 23 studies which used a manual, 14 studies (43.5%) referred to the use of a specific manual, nine studies (28.1%) mentioned the general use of a manual within the treatment process.

Assessment of adherence and competence procedures of the treatment

Table 1 also shows that most studies (90.6%, n=29) approached or achieved adequacy in terms of use of treatment adherence procedures. Competence measures were less frequently employed. More than half of the studies (56.2%, n=18) approached or achieved adequacy of therapist implementing competence procedures while 14 studies (43.8%) did not refer to the use of therapist competence procedures.

Evaluating treatment fidelity

Clinician adherence and competence: Of the 32 studies that assessed Treatment Fidelity, most reported conceptualisation of Treatment Fidelity data. Data included competence raters, employed methodology and key properties of the treatment. While the majority of these studies defined Treatment Fidelity specifically in terms of adherence and/or competence (59.4%, n=19), one third of studies appraised Treatment Fidelity in only general terms (34%, n=11), (Table 2). A majority of studies (71.9%, n=23) assessed Treatment Fidelity using adherence and/or competence concepts, and half of the studies (50%, n=16) used specific adherence or competence tools to measure Treatment Fidelity.

On the other hand, more than a quarter of the studies (28%, n=9) only assessed Treatment Fidelity indirectly or failed to provide detail of the assessment of Treatment Fidelity (28.2%, n=9). Most of these (six studies) referred to Treatment Fidelity indirectly with the remaining three studies not reporting the use of any fidelity tools. Explaining and assessing (observational data) adherence and competence can involve direct, indirect or hybrid strategies. Direct observations are generally conducted by trained staff present in the treatment setting, viewing sessions through one-way mirror or via monitors and/or videotapes. Indirect methods include self-report, rating scales, interviews and permanent products (eg. written homework assignments or data collection sheets) of treatment implementation.

Rating adherence and competence: Evaluating Treatment Fidelity through the use of adherence and competence scales requires the use of raters who assess and score treatment delivered by clinicians or researchers. Half of the 32 studies (n=16) did not report any training of raters, (Table 3). In a third of the studies (31.3%, n=10) training provided to raters was indirect or not specific to the treatment. Over two thirds of the studies (68.7%, n=22) measured adherence and/or competence.

A third of studies (31.3%, n=10) relied on indirect measures to assess fidelity. Reactivity to Treatment Fidelity (clinicians altering their performance or behaviours due to the awareness that they are being observed) was controlled within 21 studies (65.6%) with a further 11 studies (34.4%) controlling for reactivity indirectly. That is, observations were conducted at randomised times without prior notice being provided to clinicians.

Reporting treatment fidelity: Nine studies provided informative data related to therapist adherence, eight of which provided numerical data related to measurement of Treatment Fidelity using competence scales, (Table 4). Of the remaining 23 studies using adherence scales most (71.9%) did not provide detailed informative data of treatment adherence levels although they assessed treatment adherence and provided numerical data. Of the 24 which did not provide information

	Therapist treatment adherence procedures			Therapist competence procedures			Use of the manual			
Variable (ITIPS range)	IA (6-12)	AA (13-18)	AD (19-24)	IA (6-12)	AA (13-18)	AD (19-24)	Manual not mentioned	Manual only mentioned	Manual is general	Manual is specific
Overall *(N)*	3	17	12	14	9	9	9	14	4	5
Overall (%)	9.4%	53.1%	37.5%	43.8%	28.1%	28.1%	28.1%	43.8%	12.5%	15.6%
Mean Score	12	15.4	21	7.6	15.3	21.8				
SD	0	1.8	1.7	1.5	1.9	1.9				
Median	12	15.0	20.5	7	16	22				
Min-Max	12-12	13-18	19-24	6-11	13-18	19-24				

Table 1: Implementation of therapist treatment adherence and therapist competence procedures and use of the manual Note: IA=inadequate; AA=approaching adequacy; AD=adequate. Total studies N=32.

about competence scales 12 studies (37.5%) provided no data related to competence levels. The remaining 12 studies (37.5%) provided some numerical data, however it was not sufficient to allow determination of competence levels.

Implementation of treatment fidelity procedures across the four domains of the ITIPS: Table 5 shows the adequacy levels across the four domains of establishing, assessing, evaluating, and reporting fidelity.

I) Establishment: Less than half (40.6%, n=13) of the studies established procedures for ascertaining Treatment Fidelity and nearly a quarter (21.9, n=7) of the studies approached adequacy in this domain. However, more than one third of studies (37.5%, n=12) did not establish adequate Treatment Fidelity procedures.

II) **Assessment**: In terms of assessing Treatment Fidelity about one third of studies (31.3%, n=10) approached adequacy, a further third of studies implemented adequate procedures (34.4%, n=11), while the remaining third of studies (34.3%, n=11) scored within the inadequate range on the ITIPS scale.

III) **Evaluation**: Over a half of studies (53.1%, n=17) approached adequacy in terms of methods of evaluation of Treatment Fidelity, with markedly fewer (18.8%, n=6) achieving scores indicating adequate implementation. More than a quarter of the studies (28.1%, n=9) achieved scores which indicated inadequate evaluation procedures.

IV) **Reporting**: Approximately a third of studies approached adequacy (31.3%, n=10) in terms of reporting of Treatment Fidelity procedures and findings with even more (46.9%; n=15) demonstrating adequate reporting. However, seven studies (21.9%) did not report Treatment Fidelity procedures adequately.

Total levels of treatment fidelity implemented: Overall on each of the 22 items, a score of 1 or 2 was assumed to reflect inadequate implementation of integrity procedures; a score of 3 indicated that implementation approached adequacy; and a score of 4 reflected adequate implementation of integrity procedures. The overall score of each clinical trial was calculated using a combination of the percentage of treatments implementing integrity procedures with (a) inadequately if the study's total score ranged between 22 and 44; (b) in a manner approaching adequacy if the total score ranged between 45 and 66; and (c) adequately if the total score exceeded 66. The percentage of treatments implementing integrity procedures within each range of scores was calculated and shown in Table 6.

Of the 32 clinical studies within this systematic review, a quarter of the studies (25%, n=8) did not adequately implement Treatment Fidelity. Three quarters of the reviewed studies either approached adequacy (50%, n=16) or were adequate (25%, n=8) in implementing Treatment Fidelity procedures.

Discussion

Whilst solid advances in Treatment Fidelity research continue to be made, several opportunities to strengthen this approach are yet to be realised. This review assessed 32 clinical trials of which only fifty percent adequately addressed Treatment Fidelity.

	No	Indirect	Adherence or competence	Adherence and competence	No	Indirect	Adherence or Competence	Adherence and Competence
%	6.3	34.4	37.5	21.9	9.4	18.8	50.0	21.9
N	2	11	12	7	3	6	16	7

Table 2: Treatment Fidelity in terms of treatment adherence and therapist competence.

	Training raters			Assessment inter-rater reliability				Control for measure reactivity		
	Yes	No	Indirect	Yes (adherence or competence)	Yes(both adherence and competence	No	Indirect	Yes	No	Indirect
%	18.7	50.0	31.3	53.1	15.6	0	31.3	65.6	0	34.4
N	6	16	10	17	5	0	10	21	0	11

Table 3: Measuring Treatment Fidelity.

	Adherence			Competence		
	No	Not informative	Informative	No	Not informative	Informative
%	0	71.9	28.1	37.5	37.5	25.0
N	0	23	9	12	12	8

Table 4: Provision of numerical data of treatment adherence and competence in clinical trials.t

	Establishing			Assessing			Evaluating			Reporting		
Variable (ITIPS range)	IA (6-12)	AA (13-18)	AD (19-24)	IA (7-14)	AA (15-20)	AD (21-28)	IA (5-10)	AA (11-15)	AD (16-20)	IA (4-8)	AA (9-12)	AD (13-16)
Overall *(N)*	12	13	7	11	10	11	9	17	6	7	15	10
Overall (%)	37.5%	40.6%	21.9%	34.4%	31.3%	34.4%	28.1%	53.1%	18.8%	21.9%	46.9%	31.3%
Mean Score	10.8	15.2	21.0	12.1	16.7	24.1	8.9	12.9	17.5	7.3	10.7	14.5
SD	1.5	1.7	1.3	1.5	1.2	2.3	0.8	1.4	1.4	0.5	1.0	1.1
Median	11.5	15.0	22.0	12.0	16.5	24.0	9.0	13.0	17.5	7	11	14.5
Min-Max	8-12	13-18	19-22	10-14	15-19	21-28	8-10	11-15	16-19	7-8	9-12	13-16

Note: IA=inadequate; AA=approaching adequacy; AD=adequate. Total studies N=32.

Table 5: Adequacy levels across the four domains of establishing, assessing, evaluating, and reporting fidelity.

Use of Treatment Fidelity Procedures within Clinical Trials of Psychological Interventions

This review found gaps across all four domains of the ITIPS scale, with 12 studies attaining inadequate scores for establishment, 11 for assessment, nine for evaluation, and seven for reporting of Treatment Fidelity. Overall, a third of the selected articles showed inadequate implementation of Treatment Fidelity. In some cases, the poor rating may reflect insufficient reporting rather than a lack of procedures, while in other cases the establishment, measurement, or evaluation of Treatment Fidelity may have been inadequate.

Treatment Fidelity in psychosocial research is implemented in accordance with theoretical and procedural models of adherence and competence measurements. A small number of studies addressed both the use of adherence and competence measures and their assessment. However, more value was placed on the assessment of adherence to Treatment Fidelity than on therapist's competence levels. Goense and colleagues [49] identified a similar trend in their review. The concept of adherence within Treatment Fidelity may be better suited to quantitative measurement. Adherence measures assess how frequently and to what degree the therapist ensures treatment "purity" [50]. In contrast, it is difficult to provide a quantitative measurement of therapist competence in implementing the treatment without relying heavily upon the clinical judgement and expertise of selected assessors.

Alignment of Clinical Trial Treatment Fidelity Procedures with the Implementation of Treatment Integrity Procedures Scale (ITIPS)

Across the four domains of Treatment Fidelity in the ITIPS, we found that methods for establishing and assessing fidelity scored below fifty percent on average, whilst the evaluating and reporting of results approached average in a number of studies. Only seven studies (21.9%) had adequately established procedures for ascertaining Treatment Fidelity. Typically, a manual was provided for the therapist when implementing the treatment. However, it is noteworthy that only five studies reported providing therapists with a specific treatment manual. Additionally, not all studies provided therapists with training and/or supervision related to use of the manual and implementation of treatment. One explanation could be that the therapy was well known and that guidelines and training were already broadly available. For example, several studies involved the use of the widely practiced treatment of Cognitive Behavioural Therapy.

In contrast, it would be expected that new therapies and emerging interventions would place greater reliance on the provision of a specific manual. It should be noted that whilst a clinician can be very skilled in delivering a treatment, they may not necessarily be adhering to pertinent techniques contained within a prescribed manual. In terms of assessment, in most studies indirect methods of fidelity assessment were more commonly used than specific adherence and competence scales. A key limitation of such indirect methods is the lack of capacity in measuring the quality of the delivered intervention or treatment [50].

Total Treatment Fidelity			
Variable (ITIPS range)	IA (22-44)	AA (45-66)	AD (67-88)
Overall (N)	8	16	8
Overall (%)	25	50	25
Mean score	38.75	54.81	76.75
SD	2.49	5.59	4.83
Median	38.50	54.00	73.50
Min-Max	35-42	47-64	70-83

Note: IA=inadequate; AA=approaching adequacy; AD=adequate. Total studies N=32.

Table 6: Adequacy levels of the total implementation of Treatment Fidelity procedures in 32 clinical trials studies.t.

The fourth and final domain, reporting, scored poorly overall. It appeared the reporting of Treatment Fidelity was influenced by the lack of establishment and implementation of rigorous assessment procedures across several studies. This rendered interpretation of fidelity data difficult. Clinician adherence was generally more adequately reported than clinician competence with very few studies reporting both adherence and competence measures. Numerous studies were deficient in reporting critical details of their evaluation, such as inter-rater reliability. Moreover, when reporting Treatment Fidelity data, many studies provided little detail of their fidelity measurement findings. A possible explanation may be authors attached more weight to treatment outcomes than to the importance of assessing and reporting Treatment Fidelity. Nevertheless the failure to implement Treatment Fidelity strategies limits the available conclusions to be drawn from the study and the overall generalizability of the findings.

The findings of this systematic review suggested a need for guidelines to better detail the key Treatment Fidelity of establishment, evaluation, assessment and reporting. The clinical trials in this review had implemented and reported according to the CONSORT Guidelines [31]; however, the quality of fidelity processes was generally inadequate with a majority of studies insufficiently reporting both therapist adherence and competence measurements. Whilst it is recognised that conceptualising and evaluating Treatment Fidelity is critical in understanding the validity of research results, clinical trials can still achieve this 'gold standard' without evidence of high quality Treatment Fidelity procedures. Perrepletchikova and colleagues [23] argued that Treatment Fidelity needs to be elevated in prominence and to be viewed as fundamental for empirical research. Our review supports this argument. This review highlights an opportunity for strengthened program implementation through adopting enhanced procedures for future clinical trials to maximise outcomes, a view shared by Perepletchivova and colleagues [23] who stated that, "guidelines of empirical testing of psychological treatment require re-evaluation" (p838). This review provides further evidence in support of this recommendation. Moreover, changes are recommended to the Guidelines for reporting of clinical trials for psychological research to better communicate procedures for establishing assessing, evaluating, and reporting of fidelity. A review of the CONSORT guidelines [32] would provide an opportunity for amendments to establish procedures for ascertaining Treatment Fidelity across each of the four ITIPS domains. Whilst the guidelines provide a solid starting point for the reporting of clinical trials, they remain vague in terms of reporting on Treatment Fidelity. Borrelli and colleagues [34] suggested the addition of a fidelity framework to best support treatment design, therapist training, treatment delivery, treatment receipt and treatment enactment.

The benefits of revising the CONSORT Guidelines include: establishing a structured Treatment Fidelity focus, increased usage of treatment manuals, enhanced consistency of clinician treatment implementation, and greater statistical power achieved through a more standardised measurement method. This review suggests the best suited tools to deliver high levels of Treatment Fidelity were adherence scales, competence scales and specific manuals. These tools provide a constant reference point to better support consistency in the delivery of a program across multiple sites by different clinicians.

The ITIPS domains of establishment, evaluating, assessing, and reporting Treatment Fidelity supported the goals of this systematic review. However, some minor refinement may prove beneficial for future research studies. The intentional broadness of the domains makes it difficult to effectively drill down to specific elements of interest. For example, it is not possible to readily capture the specific Treatment Fidelity tools used or professional development opportunities provided to clinicians. Accordingly, the option to include additional questions within domains may help overcome this identified limitation.

Limitations

There are five main limitations to this study. Firstly, the assessors were not blind to the authors of the selected clinical trials allowing for the potential of observer bias. Secondly, inclusion and exclusion criteria may have been too restrictive with specific search terms having unintentionally excluded valid clinical trials. Thirdly, the small number of clinical trials included affects the generalisability of the current study. Perhaps having a longer time frame and including studies published prior to 2004 may have led to a better understanding of the evolving nature of Treatment Fidelity.

Fourthly, expanding the review parameters from clinical trials to include systematic reviews and meta-analyses may have provided a more heterogeneous field for comparison. Finally, Treatment Fidelity was identified as not commonly being the specific focus of the clinical trials, and these studies were often undertaken by a core group of authors. For example, Hogue contributed to three separate included trials. This makes it problematic to generalise findings given the potential for bias linked with this research group.

Conclusion and future directions

Mental health professionals working within the fields of psychology and psychiatry seek evidence-based treatments to underpin their day to day clinical work. High quality Treatment Fidelity within clinical trials provides robust evidence for effectiveness of a given therapy. This systematic review, however, found under-usage of fidelity measures within clinical trials, contributing to limited quality of Treatment Fidelity, and consequent limited evidence for effectiveness of therapy. The review also found that those studies, which included multi-method approaches to Treatment Fidelity using adherence and competence measures and treatment manuals, achieved higher ratings of Treatment Fidelity quality. This provides clear direction for future research, suggesting that inclusion of such procedures will promote improved Treatment Fidelity. While the development of the CONSORT Guidelines has contributed to improved quality of clinical trials and hence strengthened the evidence base for specific therapies, these protocols do not presently provide detailed recommendations related to Treatment Fidelity procedures. Thus one means of improving quality of Treatment Fidelity in clinical trials would be the inclusion of such detailed guidance.

This review evaluated Treatment Fidelity quality in four procedural domains; establishment, evaluation, assessment, and reporting. Integrating procedures within these domains in future clinical trials will provide data, which promotes greater understanding of treatment implementation and strengthened evidence for treatment effectiveness. Reaching agreement on core measures and fidelity tools to support improved levels of Treatment Fidelity in psychosocial research will promote quality of Treatment Fidelity, and consistency across mental health disciplines, allowing future research to provide more robust evidence in support of better client outcomes.

Acknowledgment

This project would have not been possible without the assistance of Ted Graham (Clinical Psychologist and Research Assistant) and Anne Young (Librarian, Monash University Melbourne).

References

1. Leichsenring F, Salzer SJ, Hilsenroth M, Leibing E (2011) Treatment integrity: An unresolved issue in psychotherapy research. Current Psychiatry Reviews 7: 313-321.
2. Kazdin AE (1986) Comparative outcome studies of psychotherapy: Methodological issues and strategies. J Consult Clin Psychol 54: 95-105.
3. Yeaton WH, Sechrest L (1981) Critical dimensions in the choice and maintenance of successful treatments: Strength, integrity, and effectiveness. J Consult Clin Psychol 49: 156-167.
4. Moncher FJ, Prinz FJ (1991) Treatment fidelity in outcome studies. Clinical Psychology Review 11: 247-266.
5. Riedel BW, Lichstein KL (1994) Behavioral assessment and treatment of insomnia: a review with an emphasis on clinical application. Behavior Therapy 25: 659-688.
6. Burgio L, Corcoran M, Lichstein K, Nichols L, Czaja S, et al. (2001) Judging outcomes in psychosocial interventions for dementia caregivers: The problem of treatment implementation. The Gerontologist 4: 481-489.
7. Lichstein KL, Riedel BW, Grieve R (1994) Fair tests of clinical trials: A treatment implementation model. Advances in Behavior Research and Therapy 16: 1-29.
8. Wilson GT (1996) Manual-based treatments: The clinical application of research findings. Behaviour Research and Therapy 34: 295-314.
9. Town JM, Diener MJ, Abbass A, Leichsenring F, Driessen E, et al. (2012) A meta-analysis of psychodynamic psychotherapy outcomes: Evaluating the effects of research-specific procedures. Psychotherapy 49: 276.
10. Chorpita BF, Taylor AA, Francis SE, Moffitt C, Austin AA (2004) Efficacy of modular cognitive behavior therapy for childhood anxiety disorders. Behavior Therapy 35: 263-287.
11. Lopata C, Thomeer ML, Volker MA, Nida RE, Lee GK (2008) Effectiveness of a manualized summer social treatment program for high-functioning children with autism spectrum disorders. J Autism Dev Disord 38: 890-904.
12. Moretti MM, Obsuth I (2009) Effectiveness of an attachment-focused manualized intervention for parents of teens at risk for aggressive behaviour: The Connect Program. Journal of Adolescence 32: 1347-1357.
13. Smith SW, Daunic AP, Taylor GG (2007) Treatment fidelity in applied educational research: Expanding the adoption and application of measures to ensure evidence-based practice. Education and Treatment of Children 30: 121-134.
14. Binder JL (1993) Is it time to improve psychotherapy training? Clinical Psychology Review 13: 301-318.
15. Binder JL (1999) Issues in teaching and learning time-limited psychodynamic psychotherapy. Clin Psychol Rev 19: 705-719.
16. Strupp HH, Anderson T (1997) On the limitations of therapy manuals. Clinical Psychology Science and Practice 4: 76-82.
17. Hilsenroth MJ, Defife JA, Blagys MD, Ackerman SJ (2006) Effects of training in short-term psychodynamic psychotherapy: Changes in graduate clinician technique. Psychotherapy Research 16: 293-305.
18. Strupp HH, Anderson T (1997) On the limitations of therapy manuals. Clinical Psychology, Science & Practice 4: 76-82.
19. Hogue A, Henderson CE, Dauber S, Barajas PC, Fried A, et al. (2008) Treatment adherence, competence, and outcome in individual and family therapy for adolescent behavior problems. J Consult Clin Psychol 76: 544.
20. Waltz J, Addis ME, Koerner K, Jacobson NS (1993) Testing the integrity of a psychotherapy protocol: assessment of adherence and competence. Journal of consulting and clinical psychology 61: 620.

21. Webb CA, DeRubeis RJ, Barber JP (2010) Therapist adherence/competence and treatment outcome: A meta-analytic review. Journal of Consulting and Clinical Psychology 78: 200.

22. Fixsen DL, Naoom SF, Blase KA, Friedman RM, Wallace F (2005) Implementation research: A synthesis of the literature. Tampa, FL: University of South Florida, Louis de la Parte Florida Mental Health Institute, The National Implementation Research Network (FMHI Publication #231).

23. Perepletchikova F, Treat TA, Kazdin AE (2007) Treatment integrity in psychotherapy research: analysis of the studies and examination of the associated factors. J Consul Clin Psychol 75: 829.

24. Carpenter JS, Burns DS, Wu J, Yu M, Ryker K, et al. (2013) Strategies used and data obtained during treatment fidelity monitoring. Nurs Res 62: 59.

25. Persch AC, Page SJ (2013) Protocol development, treatment fidelity, adherence to treatment, and quality control. Am J Occup Ther 67: 146-153.

26. Resnick B, Galik E, Enders H, Sobol K, Hammersla M, et al. (2011) Pilot testing of function-focused care for acute care intervention. J Nur Care Qual 26: 169-177.

27. Robb SL, Burns DS, Docherty SL, Haase JE (2011) Ensuring treatment fidelity in a multi-site behavioral intervention study: implementing NIH behavior change consortium recommendations in the SMART trial. Psycho-Oncology 20: 1193-1201.

28. Garbacz L, Brown D, Spee G, Polo A, Budd K (2014) Clinical Child and Family Psychology Review 17: 230-247.

29. Perepletchikova F, Kazdin AE (2005) Treatment integrity and therapeutic change: Issues and research recommendations. Clinical Psychology: Science and Practice 12: 365-383.

30. Dickersin K, Chan SS, Chalmersx TC, Sacks HS, Smith H (1987) Publication bias and clinical trials. Control clin trials 8: 343-353.

31. Moher D, Liberati A, Tetzlaff J, Altman D (2009) The PRISMA Group-Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement.

32. Mohera D, Hopewellb S, Schulzc KF, Montorid V, Gøtzschee PC, et al. (2010) CONSORT 2010 Explanation and Elaboration: updated guidelines for reporting parallel group randomised trials. Journal of Clinical Epidemiology: 63

33. Rennie D (1996) How to report randomized controlled trials: the CONSORT statement. JAMA 276: 649-649.

34. Borrelli B, Sepinwall D, Belllg AJ, Breger R, DeFrancesco C, et al. (2005) A New Tool to Assess Treatment Fidelity and Evaluation of Treatment Fidelity Across 10 Years of Health Behaviour Research. J Consult Clin Psychol 73: 852-860.

35. Bellg AJ, Borrelli B, Resnick B, Hecht J, Minicucci DS, et al. (2004) Enhancing treatment fidelity in health behaviour change studies: best practices and recommendations from the NIH Behavior Change Consortium. Health Psychol 23: 443-451.

36. Bond GR, Becker DR, Drake RE (2011) Measurement of Fidelity of Implementation of Evidence-Based Practices: Case Example of the IPS Fidelity Scale. Clinical Psychology: Science and Practice 18: 126-141.

37. Carroll C, Patterson M, Wood S, Booth A, Rick J, et al. (2007) A conceptual framework for implementation fidelity. Implementation Science 2: 1-9.

38. Mowbray CT, Holter MC, Teague GB, Bybee D (2003) Fidelity criteria: Development, measurement, and validation. American Journal of Evaluation 24: 315-340.

39. Stein KF, Sargent JT, Rafaels N (2007) Intervention research: establishing fidelity of the independent variable in nursing clinical trials. Nurs Res 56: 54-62.

40. MacPherson H, Altman DG, Hammerschlag R, Youping L (2010) Revised standards for reporting interventions in clinical trials of acupuncture (STRICTA): extending the CONSORT statement. Journal of Evidence-Based Medicine 3: 140-155.

41. Schulz KF, Altman DG, Moher D (2010) CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. BMC Medicine 8: 18.

42. Gutman, Murphy (2012).

43. Davidson KW, Goldstein M, Kaplan RM, Kaufmann PG, Knatterud GL, et al. (2003) Evidence-based behavioral medicine: what is it and how do we achieve it? Ann Behav Med 26: 161-171.

44. Sloan JA, Cella D, Frost M, Guyatt GH, Sprangers M, et al. (2002) Assessing clinical significance in measuring oncology patient quality of life: Introduction to the symposium, content overview and definition of terms. Mayo Clin Proc 77: 367-370.

45. Godwin M, Ruhland L, Casson I, MacDonald S, Delva, D, et al. (2003) Pragmatic controlled clinical trials in primary care: the struggle between external and internal validity. BMC Med Res Methodol 3: 28.

46. Mark MM (1983) Treatment implementation, statistical power, and internal validity. Evaluation Review 7: 543-549.

47. Stone-Romero EF, Alliger GM, Aguinis H (1994) Type II error problems in the use of moderated multiple regression for the detection of moderating effects of dichotomous variables. Journal of Management 20: 167-178.

48. Litière S, Alonso A, Molenberghs G (2007) Type I and Type II Error Under Random-Effects Misspecification in Generalized Linear Mixed Models. Biometrics 63: 1038-1044.

49. Goense P, Boendermaker L, van Yperen T, Stams G J, van Laar J (2014) Implementation of treatment integrity procedures: An analysis of outcome studies of youth interventions targeting externalizing behavioral problems. Zeitschrift für Psychologie 222: 12.

50. Weck F, Weigel M, Richtberg S, Stangier U (2011) Reliability of adherence and competence assessment in psychoeducational treatment influence of clinical experience. J Nerv Ment Dis 199: 983-986.