

Treatment fidelity in clinical trials

Chad E. Cook^{1,2,3}, Bryan O'Halloran⁴, Steve Karas⁵, Mareli Klopper⁶, Jodi L. Young⁷

¹Department of Orthopaedics, Duke University School of Medicine, Duke University, Durham, North Carolina - USA

²Department of Population Health Sciences, Duke University, Durham, North Carolina - USA

³Duke Clinical Research Institute, Duke University, Durham, North Carolina - USA

⁴Division of Physical Therapy, Medical University of South Carolina, Charleston, South Carolina - USA

⁵Department of Physical Therapy, Chatham University, Pittsburgh, Pennsylvania - USA

⁶Department of Physical Therapy, Graceland University, Independence, MO - USA

⁷Department of Physical Therapy, Bellin College, Green Bay, Wisconsin - USA

ABSTRACT

In the context of clinical trials, treatment fidelity (TF) has traditionally referred to the extent to which an intervention or treatment is implemented by the clinicians as intended by the researchers who designed the trial. Updated definitions of TF have included an appropriate design of the intervention that was performed in a way that is known to be therapeutically beneficial. This requires careful attention to three key components: (1) protocol and dosage adherence, (2) quality of delivery, and (3) participant adherence. In this viewpoint, we describe several cases in which TF was lacking in clinical trials and give opportunities to improve the deficits encountered in those trials. We feel that along with quality, risk of bias, and certainty of evidence, TF should be considered an essential element of the veracity of clinical trial.

Keywords: Clinical trials, Fidelity, Quality, Treatment

Introduction

How many times have you read a research study and either: (1) had no idea what the treatment intervention consisted of; or (2) realized that the “intervention” that was used in the study was nothing like what you would apply in clinical practice? If you’ve encountered these two situations while reading literature, you may have been witness to limitations of treatment fidelity (TF).

Despite its importance, TF is often poorly reported in clinical trials (1-3). This is especially the case in behavioral-based studies that require some degree of clinician interpretation of the patient’s progress and a modification based on that interpretation (4). It may also be because the definition of TF can vary across studies and contexts. Although TF generally refers to the extent to which an intervention is delivered as intended, ensuring consistency and reliability, terms such as “adherence,” “integrity/veracity,” or “implementation fidelity” are commonly used, which may not be anchored to the same underlying concept.

In this viewpoint, we focus on perspectives that have a “clinical context” (with a goal of improving clinician interpretation of TF) and provide a modern definition of TF, by

describing two key components of TF (adherence and veracity), discuss examples in the literature in which TF was lacking, and provide methods to improve the implementation of interventions in clinical trials. We hope to show that in addition to commonly measured constructs such as quality, risk of bias, and certainty of evidence, TF should be assessed when interpreting the meaningfulness of a clinical trial.

A modern definition of TF

In the context of clinical trials, TF has historically referred to the extent to which an intervention or treatment adheres to the implementation parameters intended by the researchers who designed the trial (5). Indeed, appropriate implementation is critical as TF is essential in ensuring that the results of the trial accurately reflect the treatment effects of the intended intervention, with no additions or omissions. Adequate TF improves one’s interpretation of the outcome data in research studies, improves the likelihood of reproducibility (if studied again), and is essential for clinical translation (5,6). This demands appropriate reporting of treatment structure used in the trial. Perhaps most importantly, TF is one of the few elements in a clinical trial that equally represents components of internal and external validity (5).

Adherence of TF routinely measures protocol and dosage adherence. Adherence can be considered as “did the researchers do as they indicated they would do?” Protocol and dosage adherence reflect the extent to which the intervention was delivered as planned. It involves an assessment of whether the treatment protocol was followed closely, including the dosage, frequency, and duration of the intervention.

Received: May 13, 2024

Accepted: August 18, 2024

Published online: September 16, 2024

Corresponding author:

Chad E. Cook

email: chad.cook@duke.edu



Investigators in clinical trials should demonstrate an effort to show that they have optimized dosage capacity by incorporating known parameters of therapeutic effectiveness and an application that is similar to that provided in clinical practice.

Recent consensus-based work has markedly widened the scope of topics that reflect TF (7). In addition to whether the intervention is delivered with a high degree of adherence, TF should include efforts to ensure that the application of the intervention is performed in a way that is known to be therapeutically beneficial (4) (Fig. 1). In other words, was the veracity of the intervention performed and implemented in a manner that should allow someone to improve if performed in a similar clinical situation? To ensure the veracity of TF in a clinical study, one must consider: (a) the quality of delivery and (b) participant adherence.

Quality of delivery assesses both the therapeutic potency of the interventions and the competency of the individuals delivering the intervention. Therapeutic potency reflects whether the clinical parameters such as dosage, time, etc., are performed in a way that allows optimal therapeutic recovery. In a pharmaceutical trial, it would reflect whether

the research participant received the appropriate dosages of the medications at appropriate time intervals. Additionally, quality of delivery involves evaluating whether the research administrators have the necessary skills, training, and expertise to deliver the treatment effectively, and ensuring consistency in the provision of interventions between those delivering the treatment.

Participant adherence refers to the extent to which participants engage with and respond to the intervention. It involves monitoring participants' adherence to the intervention protocol, their understanding of the intervention, and their willingness to participate. The selection of appropriately responsive measures that actually assess patient engagement and change in outcomes within the targeted domain is requisite to ensure these measures have meaning.

Examples and recommendations involving TF in clinical trials

Although critical, it is important to recognize that assessment and implementation of TF procedures in a trial is a challenging process (2). There are numerous studies that

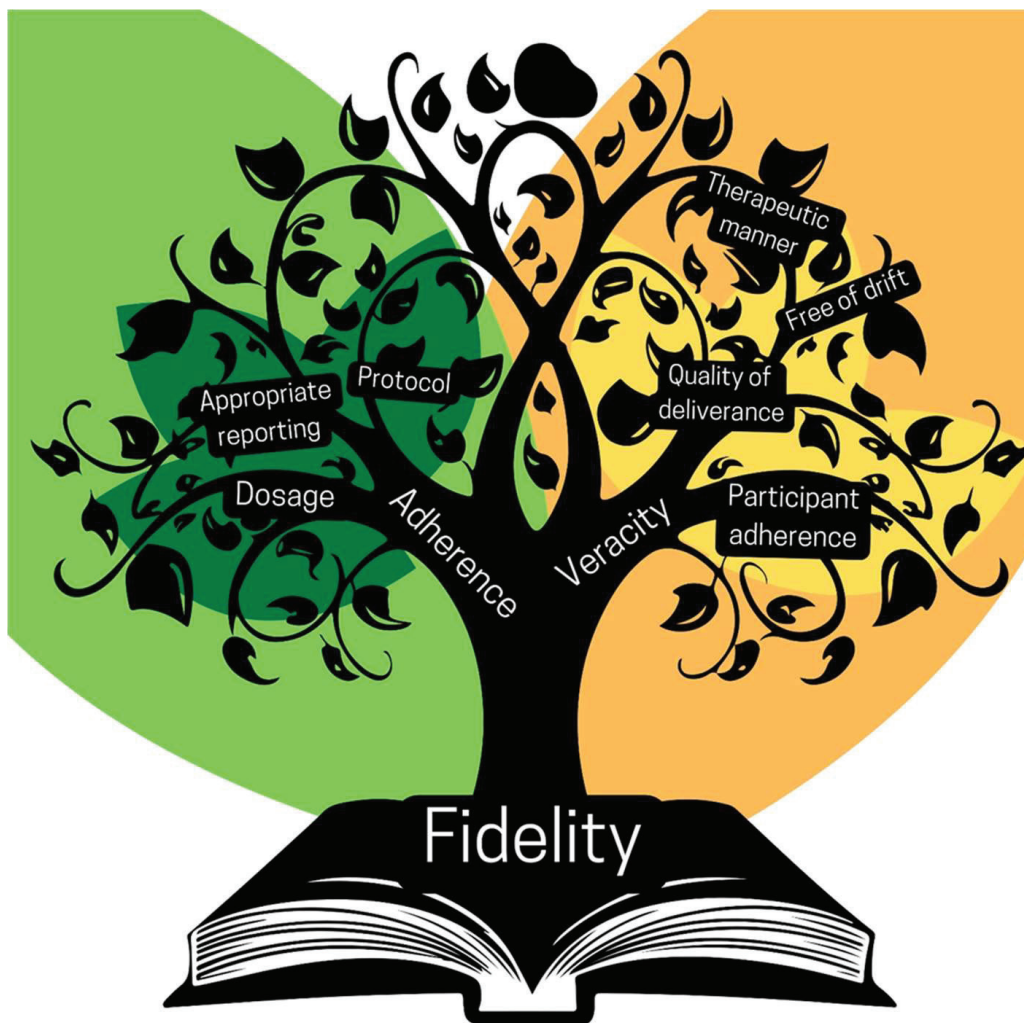


FIGURE 1 - Knowledge tree reflecting the elements of treatment fidelity.

have either experienced or highlighted TF concerns. In this section, we outline examples of TF limitations and provide options for improvement in future studies.

Procedural drift (implementation drift)

Procedural drift is a subcomponent of TF that may influence how a clinician delivers a specific intervention over the course of treatment. It occurs when a clinician chooses the most appropriate intervention based on recommendations at the onset of treatment, and then “drifts” away from using adequate intervention over an episode of care, likely due to their personal beliefs, training, and/or lack of motivation to deviate from their typical model of practice (4). A potential example of procedural drift is the recently published TARGET trial. The TARGET trial (8) reported limited TF in the implementation of a psychologically informed physiotherapy approach, despite initial agreement and formalized training among study clinicians.

Options for improvement

Adding in checklists or manuals that clinicians and researchers can use to improve the quality of specific interventions provided is recommended to limit procedural drift, but adherence to checklists may not always be an easy task due to lack of time, experience, and the belief that the checklists are unnecessary (2). Direct supervision and feedback, videotaping and structured meetings to discuss interventions, along with checklists/manuals, may reinforce the need to limit procedural drift. Early training sessions for clinicians, along with “booster” sessions, to guide the use of appropriate and meaningful interventions may also limit procedural drift in clinical practice. Implementing regular supervised performance reviews with clinicians may assist in determining when adjustments should be made to increase TF (3). Lastly, pretests and the use of specific technologies designed to minimize procedural drift may lend value as well.

Quality and dosage of treatments

A 2021 systematic review (9) was published involving manual therapy interventions vs. sham treatment approaches. In the review, 11 of the 24 reviewed studies (46%) included one visit involving only one technique, applied once. This is not reflective of clinical application nor is it considered to be therapeutic. Further, in many cases, the treatment was applied without interactions with the participants, which did not reflect the contextual aspect of a treatment domain.

Options for improvement

To examine the full treatment effect, including contextual factors and how these are intricately tied to a specific treatment, one must provide the same unique characteristics and components of the intervention, including interpersonal interactions (10). In addition, careful effort should be made to apply the treatment in a manner that is similar to clinical practice and one that reflects clinical practice guideline recommendations.

Vague treatment applications

Recent systematic reviews have found that research reporting and quality of TF remains low across trials investigating exercise therapy and manual therapy for chronic pain, neck pain, and low back pain (11-13). Possible reasons for this deficit include increased time, additional cost, real-world feasibility, and “provider fatigue” from prescriptive and possibly clinician-limiting research designs (14).

Options for improvement

The aforementioned studies exhibited TF limitations, despite the fact that several reporting and fidelity checklists have been developed to monitor the quality of interventions provided in randomized controlled trials (RCTs) for various musculoskeletal conditions. These include the Template for Intervention Description and Replication (TIDieR) and the Consensus on Exercise Reporting Template (CERT), which were both designed to improve the reporting of interventions used in RCTs to assist with methodological transparency and reproducibility of interventions, ultimately leading to improved TF (12,13). In addition to the experimental group, it is imperative that the interventions received by the control group are well described and “controlled.” This is commonly an issue in trials and has been identified as a major area of confusion when describing the somewhat innocuous but confusing term of “usual care” (15).

There are also fidelity checklists that have been developed but their effectiveness is questionable. Fidelity checklists are cumbersome, lack succinctness for application, and often include only some of the areas (typically intervention only) that are deemed important to assess (2), frequently failing to address areas such as expertise level of the clinician or procedural drift.

Quality of delivery

In trials that do demonstrate quality reporting of interventions and provide descriptive information on the training and experience of the practitioners’ clinical decision-making even while adhering to a strict protocol, TF may still be variable between clinicians. The grade of application in manual therapy, the intensity of resistance in exercise therapy, and the content of the patient instruction including whether to respect or ignore pain are all inherent in physiotherapy interventions. Without consistency of application of these constructs, the same apparent interventions may be applied in a vastly different fashion masking treatment effect.

Options for improvement

One can improve the quality of delivery by training the study providers, and adhering to guiderails of care that are predesigned and incorporated into the training process. This process should be used in both prescriptive and pragmatic clinical trials.

Participant adherence

The recently published PEERC trial (16) is a good example of how participant adherence may have eroded the effect

of one of the treatment arms. In the study, participants with shoulder impingement received a phone-based cognitive behavioral intervention. The authors of the study indicated that there were several instances in which participants took calls: “1) while the patient was driving a car, 2) attending or coaching their youth’s sporting events, 3) while at work, 4) while cooking dinner, or 5) during other activities in which they multi-tasked the cognitive behavioral strategies of the PEERC with other daily activities.” A cognitive behavioral intervention requires careful attention and active participation to optimize benefits; both of these were absent in many cases in the PEERC trial.

Options for improvement

The necessity of participant adherence should be discussed during the study initiation, and emphasized during the trial. Further, the use of a sensitivity analysis based on those who did and did not adhere to prescribed treatment planning is an option to measure its potential effect.

Unique challenges of TF for physiotherapy and rehabilitation approaches

Measuring TF in physiotherapy can be challenging compared to other areas of healthcare, such as a pharmacological intervention that uses objective laboratory values to determine a treatment regimen, because the nature of physiotherapy is multifaceted, interventions are often clinician-dependent, and interactions between the clinician and patient are uniquely individual (17). Multiple elements impact the delivery, receipt, and enactment of a prescribed physiotherapy treatment intervention and TF may be impacted by the clinician, the patient, or the actual treatment itself (18). The skill of the physiotherapist, the individual needs of the patient, and the distinct interventions required for each individual widely vary across the physiotherapy field, which can lead to significant difficulties in measuring TF.

Multiple covariates associated with the delivery of physiotherapy or other rehabilitation services, such as the time spent with the patient, the setting, and the therapeutic alliance between the patient and provider, can influence TF (18). Because there is so much variation in physiotherapy, a specific checklist may not allow for enough latitude, leading to an unclear interpretation of how high the TF truly is (19). Adaptability within a research protocol, or “flexible fidelity” (20), allows the adjustment of protocol components in response to individual patient differences, such as tailoring exercises based on an individual’s pain response or strength. In this context, fidelity can be viewed as adherence to the underlying theory outlined in a treatment protocol, rather than to specific activities or behaviors.

Conclusion

In this viewpoint, we outline the components of TF and provide examples in the literature where TF was lacking. We argue that TF is critical to establishing the evidence base of interventions and determining the circumstances under

which an intervention is most effective. Interventions need to be delivered with a high degree of TF, which will allow for greater confidence that the outcomes observed are truly driven by the specific intervention. When TF is not adhered to in clinical research, we may rightly be left to wonder what effect minor modifications of the protocol had on patient outcomes. We suggest that there is a risk that minor modifications could potentially erode the true effect of the treatment and influence clinical outcomes, leading to “evidence” that is erroneously adopted into evolving clinical paradigms.

Disclosures

Conflict of interest: The authors report no conflicts of interest.

Financial support: This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Authors’ contributions: All authors participated in the study concept, writing, and final manuscript preparation.

Data Availability Statement: There are no data associated with this viewpoint.

References

- Ribeiro LP, Curriel-Montero F, Rodrigues-de-Souza DP, Camargo PR, Albuquerque-Sendin F. Assessment of description and implementation fidelity of clinical trials involving exercise-based treatment in individuals with rotator cuff tears: a scoping review. *Braz J Phys Ther.* 2024;28(2):101062. [CrossRef PubMed](#)
- Ginsburg LR, Hoben M, Easterbrook A, Anderson RA, Estabrooks CA, Norton PG. Fidelity is not easy! Challenges and guidelines for assessing fidelity in complex interventions. *Trials.* 2021;22(1):372. [CrossRef PubMed](#)
- Baker J, Stringer H, McKean C. Ensuring treatment fidelity in intervention studies: developing a checklist and scoring system within a behaviour change paradigm. *Int J Lang Commun Disord.* 2024;59(1):379-395. An M. [CrossRef PubMed](#)
- Cook CE, Beneciuk JM, George SZ. Procedural drift: an underappreciated element of clinical treatment fidelity. *J Orthop Sports Phys Ther.* 2022;52(2):63-66. [CrossRef PubMed](#)
- An M, Dusing SC, Harbourne RT, Sheridan SM; START-Play Consortium. START-Play Consortium. What really works in intervention? Using fidelity measures to support optimal outcomes. *Phys Ther.* 2020;100(5):757-765. [CrossRef PubMed](#)
- Feely M, Seay KD, Lanier P, Auslander W, Kohl PL. Measuring fidelity in research studies: a field guide to developing a comprehensive fidelity measurement system. *Child Adolesc Social Work J.* 2018;35(2):139-152. [CrossRef](#)
- Sousa Filho LF, Farlie MK, Haines T, et al. Developing an international consensus Reporting guideline for intervention Fidelity in Non-Drug, non-surgical trials: the ReFiND protocol. *Contemp Clin Trials.* 2024;142:107575. [CrossRef PubMed](#)
- Delitto A, Patterson CG, Stevans JM, et al. Stratified care to prevent chronic low back pain in high-risk patients: the TARGET trial. A multi-site pragmatic cluster randomized trial. *EClinicalMedicine.* 2021;34:100795. [CrossRef PubMed](#)
- Lavazza C, Galli M, Abenavoli A, Maggiani A. Sham treatment effects in manual therapy trials on back pain patients: a systematic review and pairwise meta-analysis. *BMJ Open.* 2021;11(5):e045106. [CrossRef PubMed](#)
- Testa M, Rossetini G. Enhance placebo, avoid nocebo: how contextual factors affect physiotherapy outcomes. *Man Ther.* 2016;24:65-74. [CrossRef PubMed](#)



11. Adams SC, McMillan J, Salline K, et al. Comparing the reporting and conduct quality of exercise and pharmacological randomised controlled trials: a systematic review. *BMJ Open*. 2021;11(8):e048218. [CrossRef PubMed](#)
12. McConnell R, Klopper M, Rhon DI, Young JL. The influence of exercise therapy dosing on pain and functional outcomes in patients with subacromial pain syndrome: a systematic review. *Shoulder Elbow*. 2024;16(1)(suppl):42-58. [CrossRef PubMed](#)
13. Kucksdorf JJ, Bartley J, Rhon DI, Young JL. Reproducibility of exercise interventions in randomized controlled trials for the treatment of rotator cuff-related shoulder pain: a systematic review. *Arch Phys Med Rehabil*. 2024;105(4):770-780. [CrossRef PubMed](#)
14. Fuller T, Pearson M, Peters J, Anderson R. What affects authors' and editors' use of reporting guidelines? Findings from an online survey and qualitative interviews. *PLoS One*. 2015;10(4):e0121585. [CrossRef PubMed](#)
15. Pascoe SC, Spoonmore SL Jr, Young JL, Rhon DI. Proposing six criteria to improve reproducibility of "usual care" interventions in back pain trials: a systematic review. *J Clin Epidemiol*. 2022;149:227-235. [CrossRef PubMed](#)
16. Myers H, Keefe FJ, George SZ, et al. Effect of a Patient Engagement, Education, and Restructuring of Cognitions (PEERC) approach on conservative care in rotator cuff related shoulder pain treatment: a randomized control trial. *BMC Musculoskelet Disord*. 2023;24(1):930. [CrossRef PubMed](#)
17. Whyte J, Hart T. It's more than a black box; it's a Russian doll: defining rehabilitation treatments. *Am J Phys Med Rehabil*. 2003;82(8):639-652. [CrossRef PubMed](#)
18. Toomey E, Matthews J, Hurley DA. Using mixed methods to assess fidelity of delivery and its influencing factors in a complex self-management intervention for people with osteoarthritis and low back pain. *BMJ Open*. 2017;7(8):e015452. [CrossRef PubMed](#)
19. Hall AM, Ferreira PH, Maher CG, Latimer J, Ferreira ML. The influence of the therapist-patient relationship on treatment outcome in physical rehabilitation: a systematic review. *Phys Ther*. 2010;90(8):1099-1110. [CrossRef PubMed](#)
20. Palmer JA, Parker VA, Barre LR, et al. Understanding implementation fidelity in a pragmatic randomized clinical trial in the nursing home setting: a mixed-methods examination. *Trials*. 2019 Nov 28;20(1):656. [CrossRef PubMed](#)