Treatment Fidelity in 94 Randomized Controlled Trials of Physical Rehabilitation in the ICU: A Scoping Review*

Christopher Farley, PT, MSc[1]
Anastasia N.L. Newman, PT, PhD[1]
Jen Hoogenes, PhD, MSc, MS[1]
Dina Brooks, PT, PhD[1]
Mark Duffett, PhD[2]
Michelle E. Kho, PT, PhD[1,3]

OBJECTIVES: Recent reviews demonstrated discordant effects of ICU-based physical rehabilitation on physical function. These inconsistencies may be related to differences in treatment fidelity—the extent to which a protocol is delivered as planned. Before evaluating the association of fidelity with outcomes, we must first understand the extent of treatment fidelity reporting in ICU-based physical rehabilitation randomized controlled trials (RCTs).

DATA SOURCES: Six electronic databases from inception to December 2022.

STUDY SELECTION: We included RCTs enrolling adults or children admitted to the ICU, if greater than or equal to 50% were invasively mechanically ventilated greater than 24 hours, and underwent an ICU-based physical rehabilitation intervention, with no limitation to comparators or outcomes.

DATA EXTRACTION: We screened and extracted data independently and in duplicate, with a third reviewer as needed. Extracted data included study characteristics, treatment descriptions, and the presence of National Institutes of Health Behaviour Change Consortium (NIH-BCC) treatment fidelity tool components. Treatment fidelity scores were calculated as the proportion of reported (numerator) out of total NIH-BCC components (denominator). We calculated scores across studies and by treatment group (intervention vs. comparator). We used linear regression to assess for a time trend in study treatment fidelity scores.

DATA SYNTHESIS: Of 20,433 citations, 94 studies met inclusion criteria. Authors reported a median (first–third quartiles) of 19% (14–26%) of treatment fidelity components across studies. Intervention group scores were higher than comparator groups (24% [19–33%] vs. 14% [5–24%], $p < 0.01$). We found a mean increase in study treatment fidelity scores by 0.7% (0.3 points) per year.

CONCLUSIONS: Only 19% of treatment fidelity components were reported across studies, with comparator groups more poorly reported. Future research could investigate ways to optimize treatment fidelity reporting and determine characteristics associated with treatment fidelity conduct in ICU-based physical rehabilitation RCTs.

KEYWORDS: critical care; critical illness; patient compliance; rehabilitation; reproducibility of results; respiration, artificial

Approximately 40% of people with critical illness develop ICU-acquired weakness (ICU-AW) (1), with related physical function impairments that can persist for 5 years after ICU discharge (2). Physical rehabilitation in the ICU may be an effective means to reduce physical function disability (3). Schweickert and colleagues conducted a 104-patient randomized controlled trial (RCT) of early ICU physical rehabilitation, which allocated mechanically ventilated patients to early mobilization by occupational therapists and physical therapists or standard occupational and physical therapy (4). This

*See also p. 842.

DOI: 10.1097/CCM.0000000000006192

KEY POINTS

Question: To what extent is treatment fidelity reported in ICU-based physical rehabilitation randomized controlled trials (RCTs)?

Findings: Authors reported only 19% of treatment fidelity components in the 94 included studies. Comparator groups were more poorly reported than intervention groups (14% vs. 24%). Treatment fidelity reporting is improving over time; however, is still suboptimal.

Meaning: Researchers can use this work to guide study protocol development and article preparation to optimize treatment fidelity conduct and reporting, and readers can use this work to guide interpretation of ICU-based physical rehabilitation RCTs with a consideration of treatment fidelity.

was the first RCT of early ICU rehabilitation. Patients in the intervention group started rehabilitation activities within 1.5 days of intubation, and received 19 minutes of rehabilitation per day while receiving mechanical ventilation. In contrast, the control group did not start until 7.4 days after intubation and did not receive rehabilitation activities during mechanical ventilation. Those randomized to early mobilization were 2.7 times more likely to return to independent function by hospital discharge (59% [29/49] vs. 35% [19/55], $p = 0.02$) (4). Schweickert et al (4) provided their intervention on 87% of planned days.

Since this publication, results of ICU-based physical rehabilitation trials on physical function outcomes have been mixed (3, 5–9). Inconsistent results may be attributed to differences in study samples, intervention and comparator group components (10), or outcomes assessed; however, these findings may also be related to differences in treatment protocol implementation (11).

Treatment fidelity is the extent to which a protocol is delivered as planned (11); however, fewer than half of 117 adult ICU-based physical rehabilitation intervention studies reported this metric (12). PICU-based physical rehabilitation studies have shown similar limitations in treatment fidelity reporting (13). A clearly described protocol is important to understand what therapeutic elements were intended to be delivered in each group. Furthermore, understanding how well the intervention and comparator groups' components were delivered is important to appreciate how rigorously the intervention was tested. For example, a systematic review of 60 ICU-based physical rehabilitation trials found that studies in which comparator groups received a higher dose of physical rehabilitation were less likely to show significant outcomes (6). In a neutral trial, treatment fidelity can help distinguish whether an intervention was poorly implemented or ineffective (14). We aimed to describe treatment fidelity reporting in physical rehabilitation trials started in the ICU (15).

We used the National Institutes of Health Behavioral Change Consortium (NIH-BCC) treatment fidelity tool to guide reporting assessment (11, 16). The NIH-BCC focuses on treatment fidelity reporting and was developed in the behavioral sciences. ICU-based physical rehabilitation interventions are similar to behavior change interventions. When researchers test novel treatments, healthcare providers and patients engage in new behaviors. The use of the NIH-BCC allowed us to understand the ICU-based treatment fidelity from both healthcare provider and patient perspectives.

Our research question was: "To what extent is treatment fidelity reported in ICU-based physical rehabilitation RCTs?" The primary objective was to determine the proportion of treatment fidelity components reported in ICU-based physical rehabilitation RCTs. We hypothesized that ICU-based physical rehabilitation RCT treatment fidelity reporting would not exceed 70% of components, which has been used as a benchmark for adequate reporting quality in ICU-based physical rehabilitation studies (10, 12). Second, we aimed to identify: 1) the proportion of treatment fidelity components reported by randomization group, 2) the proportion of treatment fidelity components reported by treatment fidelity domain, and 3) if a time trend existed for the proportion of treatment fidelity components reported. We hypothesized that the proportion of treatment fidelity components reported would be: 1) lower in comparator groups (12), 2) lower in the training provider domain (described further below) (17), and 3) higher in more recent publications (18).

MATERIALS AND METHODS

This review was completed with standard scoping review methodology according to the Joanna

Briggs Institute (19) and is reported according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews (PRISMA-ScR; **eTable 1**, http://links.lww.com/CCM/H486) (20). The methods were specified a priori in a published protocol (**Appendix B**, http://links.lww.com/CCM/H486) (21) and highlights are described below.

Inclusion and Exclusion Criteria

We included RCTs enrolling adults or children admitted to an ICU, where 50% of participants or more required invasive mechanical ventilation (IMV) for greater than 24 hours, since IMV for greater than 24 hours is associated with respiratory muscle weakness (22).

Studies were included if they were published in a peer-reviewed journal, assessed a planned physical rehabilitation intervention, were initiated in the ICU, and were intended to optimize physical functioning and reduce physical disability (23). Eligible interventions included passive or active exercises alone, inspiratory muscle training, progressive mobility, cycle ergometry, neuromuscular electrical stimulation (NMES), or any combination thereof (multicomponent) (12). There was no limitation of comparator type or outcome measures included. We excluded airway clearance interventions, non-English studies, crossover and cluster designs, gray literature, and abstracts.

Search Strategy

Six databases were searched from inception through December 2022 (OVID Medline, OVID Excerpta Medica Database, OVID Allied and Complementary Medicine Database, OVID Emcare, EBSCOhost Cumulative Index to Nursing and Allied Health Literature, and the Physiotherapy Evidence Database). Search terms for the three concepts ("ICU," "physical rehabilitation," and "RCT") were developed through review of the literature and in consultation with a research librarian (12, 24–26). The Peer Review of Electronic Search Strategies guidelines (27) were employed by a different team of research librarians to approve the final search strategy, which are fully reported in our published protocol (Appendix B, http://links.lww.com/CCM/H486) (21). Reference lists of included studies and systematic reviews of physical rehabilitation interventions found in the search were also screened for additional citations.

Document Selection and Data Abstraction

Citations were imported into Covidence systematic review software (2020, Veritas Health Innovation, Melbourne, VIC, Australia), a web-based platform which also automatically removed duplicate citations. Titles and abstracts were screened independently and in duplicate, with prior piloting of the eligibility criteria with 25 random citations to ensure greater than 75% agreement amongst reviewers (19). Full texts were then assessed independently and in duplicate for eligibility criteria, with similar piloting prior to commencement; a third reviewer arbitrated disagreements.

Data abstraction was completed independently and in duplicate for study characteristics (e.g., country, trial phase, funding), study design, participant characteristics (e.g., age, sex, severity of illness), intervention and comparator group descriptions, and treatment fidelity components (described below) (16). Data were extracted from a trial's parent full-text publication which described results, as well as any associated supplemental files, cited protocols, or registrations. When conflicting information existed among associated publications, we prioritized information from the full-text publication which described results. With the focus of this review on completeness of reporting, authors were not contacted for additional information.

Treatment Fidelity

The NIH-BCC fidelity tool includes five domains of treatment fidelity (study design, training providers, delivery of treatment, receipt of treatment, and enactment of treatment) encompassing 21 components to assess whether specific characteristics were present, absent, or not applicable in a study (**Table 1**) (16). The NIH-BCC tool has high inter-rater reliability (percent agreement of 78%), and studies more highly scored for treatment fidelity reporting were more likely to obtain statistically significant results (28). Guidelines for implementing each component of the treatment fidelity tool and score calculations are published elsewhere (21).

The following treatment fidelity sum scores were calculated for each study: 1) overall study, 2) by

TABLE 1.
Overview of Each Domain of the National Institutes of Health Behavioral Change Consortium Treatment Fidelity Tool (16)

Domain	Description	Components
Study design	Ensures a clearly planned treatment	1. Length of contact session(s) 2. Number of contacts 3. Content of treatment 4. Duration of contact over time 5. Provider credentials 6. Theoretical model or clinical guidelines on which the intervention is based
Training providers	Focus on features that ensure providers are competent to carry out the intervention as planned	1. Description of provider training 2. Standardized provider training 3. Provider skill acquisition post-training 4. Description of provider skill maintenance over time
Delivery of treatment	To monitor and optimize provision of the treatment as it was planned	1. Method to ensure that the contact of the intervention delivered as specified 2. Method to ensure dose of the intervention delivered as specified 3. Mechanism to assess if the provider adhered to the intervention plan 4. Assessed nonspecific treatment effects 5. Provider used treatment manual during session
Receipt of treatment	Focus on whether a participant was able to understand and perform the treatment skill	1. Assessed subject's comprehension of the intervention during the intervention period 2. Strategy to improve subject's comprehension of the intervention above and beyond what is included in the intervention 3. Assessed subject's ability to perform the intervention skills during the intervention period 4. Strategy to improve subject's performance of intervention skills during the intervention period
Enactment of treatment	To assess the participant's use of the intervention skill in a real-life setting (defined in this review as any time outside the physical rehabilitation sessions directly related to the intervention of a particular study)	1. Assessed subject's performance of the intervention skills in setting in real-life setting 2. Assessed strategy to improve subject's performance of the intervention skills in settings in which the intervention might be applied

intervention and comparator group, 3) by domain, and 4) domain score by intervention and comparator group. In studies with more than two groups, we scored the group with the most physical rehabilitation components as the intervention group. We assumed all components were applicable and thus the denominator for each calculation was the total number of NIH-BCC components. Accuracy of treatment fidelity score calculations was ensured through quality assurance procedures (**eTable 2**, http://links.lww.com/CCM/H486). Recently, ICU-based physical rehabilitation reviews of study reporting used 70% as the benchmark for adequate reporting (10, 12); we modified our hypothesized proportion from 80% reported in our protocol to 70% (**eTable 3**, http://links.lww.com/CCM/H486).

Statistical Analysis

We collated all data and calculated treatment fidelity scores (Microsoft Excel, 2019, Microsoft Corporation, Redmond, WA). Nominal variables were summarized as counts and percentages. Continuous variables were reported as mean and SD when normally distributed, or median and quartiles (first and third), when nonnormally distributed.

Simple linear regression was used to determine the association between overall study treatment fidelity scores and publication year. Evaluation of regression assumptions was completed with regression diagnostics. We compared both overall study scores and study domain scores by intervention type using one-way analysis of variance (ANOVA) or Kruskal-Wallis test with post hoc pairwise analyses. All statistical tests were conducted using an alpha level of 0.05 or 0.05 corrected with the Bonferroni method. Stata (v. 16.1 for Mac, StataCorp LP, College Station, TX) was used to conduct all statistical analyses.

We made two changes to our prespecified statistical plan (eTable 3, http://links.lww.com/CCM/H486). We originally planned to conduct independent comparisons of two groups; however, upon reflection that group scores would be related due to similarities concerning the authors' research approach, we decided to use paired assessments for group comparisons (eTable 3, http://links.lww.com/CCM/H486). We thus conducted pairwise comparisons with the paired *t*-test if normally distributed, or the sign rank test if non-normally distributed. When three or more groups were compared, the one-way ANOVA, if normally distributed, or Kruskal-Wallis test, if non-normally distributed, were used. We originally planned to conduct pairwise comparison correction using Tukey correction; however, this test is not applicable to nonparametric assessments. Thus, post hoc, we conducted pairwise comparisons using the Bonferroni correction with the independent *t*-test if normally distributed, or the Wilcoxon Rank Sum test if non-normally distributed (eTable 3, http://links.lww.com/CCM/H486).

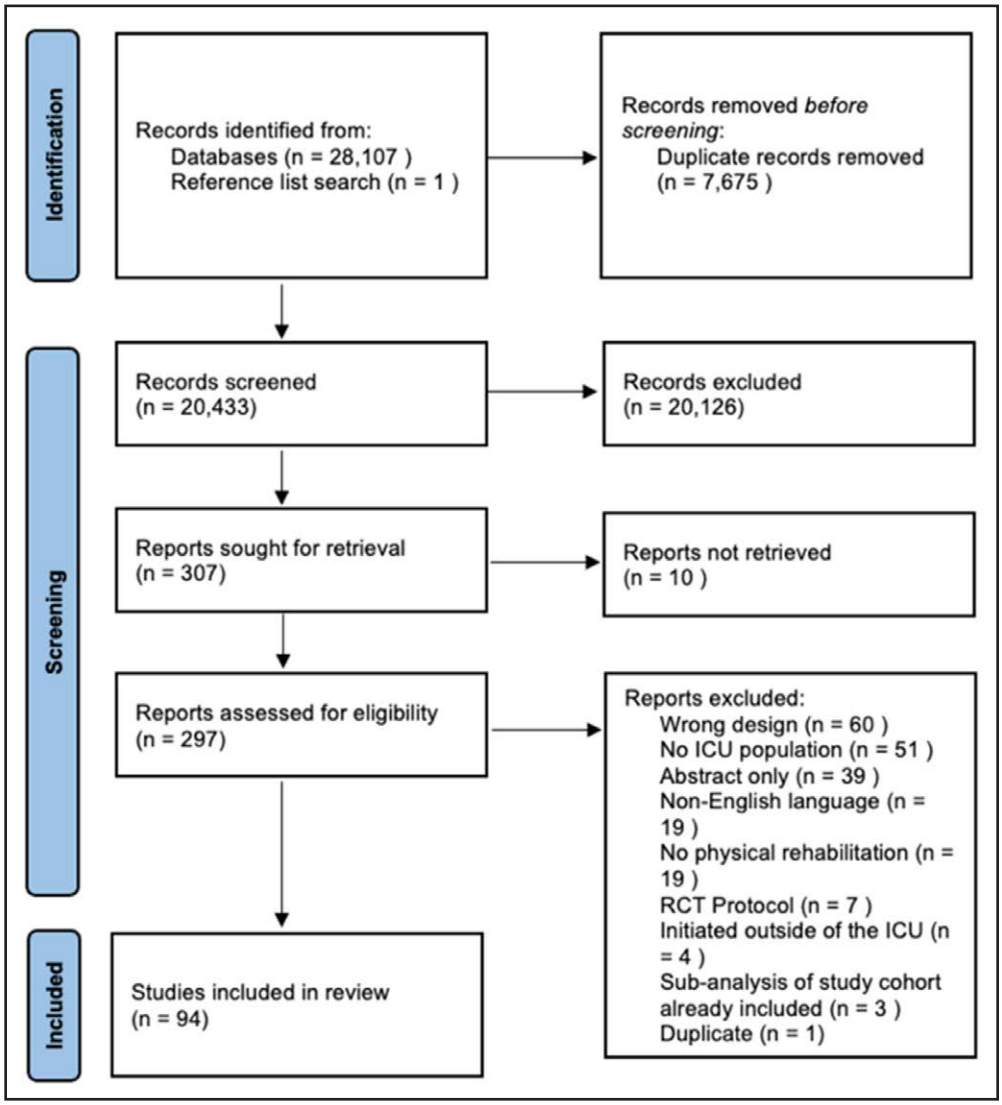

Figure 1. Preferred Reporting Items for Systematic Reviews and Meta-Analyses Extension for Scoping Reviews flow diagram (20).

RESULTS

After duplicates were removed, 20,433 were screened for inclusion. Ninety-four met the inclusion criteria (**Fig. 1**). Most studies were single-center (86% [*n* = 81]) and did not specify trial phase (65% [*n* = 61]) (**Table 2**). The median (first–third quartiles) number of randomized participants was 60 (32–104) across all trials (**eTable 4**, http://links.lww.com/CCM/H486). Ninety-one studies enrolled adults with a median age of 59 (54.8–64.0) years, and three studies enrolled pediatric participants with a median age of 8.5 (7.8–10.6) years (eTables 4 and **5**, http://links.lww.com/CCM/H486). Progressive mobility interventions were most common (29% [*n* = 27]), followed by NMES

TABLE 2.
Characteristics of Included Studies

Study Characteristics	All Studies	Adult Studies	Pediatric Studies
Included studies, *n* studies (%)	94 (100)	91 (97)	3 (3)
Countries, *n* studies (%)			
Brazil	19 (20)	18 (20)	1 (33)
China	11 (12)	11 (12)	0 (0)
Australia	10 (11)	10 (11)	0 (0)
United States	9 (10)	8 (9)	1 (33)
Greece	5 (5)	5 (5)	0 (0)
Egypt	4 (4)	4 (4)	0 (0)
Belgium	4 (4)	4 (4)	0 (0)
Others[a]	32 (34)	31 (34)	1 (33)
Number of centers, *n* studies (%)			
Single-center	81 (86)	79 (87)	2 (67)
Multi-center	13 (14)	12 (13)	1 (33)
Median number of centers (first, third quartiles)	1 (1, 1)	1 (1, 1)	1 (1, 3)
Trial phase, *n* studies (%)[b]			
Pilot	28 (30)	26 (29)	2 (67)
Full	5 (5)	5 (5)	0 (0)
Not specified	61 (65)	60 (66)	1 (33)
Intervention type, *n* studies (%)			
Progressive mobility	27 (29)	26 (29)	1 (33)
Neuromuscular electrical stimulation	23 (24)	23 (25)	0 (0)
Multicomponent	16 (17)	15 (16)	1 (33)
Inspiratory muscle training	15 (16)	15 (16)	0 (0)
Cycle ergometry	12 (13)	11 (12)	1 (33)
Passive or active exercise alone	1 (1)	1 (1)	0 (0)
Date of publication, *n* studies (%)			
2005–2009	5 (5)	5 (5)	0 (0)
2010–2014	19 (20)	19 (21)	0 (0)
2015–2019	40 (43)	38 (42)	2 (67)
2020–2022	30 (32)	29 (32)	1 (33)
ICU type, *n* studies (%)			
Medical/surgical	14 (15)	14 (15)	0 (0)
Mixed	14 (15)	14 (15)	0 (0)
General	7 (7)	7 (8)	0 (0)
Surgical	3 (3)	3 (3)	0 (0)
Unspecified	51 (54)	48 (53)	3 (100)
Other[c]	5 (5)	5 (5)	0 (0)

[a]Other: adult trials—Japan (*n* = 3), Turkey (*n* = 3), United Kingdom (*n* = 3), Austria (*n* = 2), Canada (*n* = 1), France (*n* = 2), Germany (*n* = 2), Iran (*n* = 2), Israel (*n* = 2), Taiwan (*n* = 2), Argentina (*n* = 1), Colombia (*n* = 1), Czech Republic (*n* = 1), Denmark (*n* = 1), Iceland (*n* = 1), India (*n* = 1), The Netherlands (*n* = 1), Switzerland (*n* = 1), Thailand (*n* = 1); Pediatric trial—Canada (*n* = 1).
[b]Trials classified as pilot if self-identified as phase I, phase II, pilot, or feasibility trials. Trials classified as full if self-identified as phase III, phase IV trials, and/or cited their own previous pilot work (42).
[c]Other includes medical (*n* = 2), neurology (*n* = 2), and thoracic (*n* = 1).

(24% [n = 23]) and multicomponent interventions (17% [n = 16]).

Treatment Fidelity Reporting

The median proportion of treatment fidelity components reported across studies was 19% (14–26%) (**Table 3**), which was consistent with our primary hypothesis. By group, fidelity scores were higher among intervention groups than comparators (Table 3 and **Fig. 2**). All group domain scores were significantly greater in the intervention group than the comparator group except for the enactment of treatment domain, which was greater in the intervention group but not statistically different (Table 3; and **eFigs. 1–5**, http://links.lww.com/CCM/H486). By domain, study design components were most commonly reported in the intervention and comparator groups, with median proportions of 67% (50–85%) and 33% (17–50%), respectively (eFig. 1, http://links.lww.com/CCM/H486). Across all studies, study design was the most comprehensively reported domain (**eTable 6**, http://links.lww.com/CCM/H486), with the number of contacts being the most frequently reported component across both groups (Table 3; **eTable 7**, http://links.lww.com/CCM/H486).

Fidelity Reporting Over Time

By linear regression, recent publications were more likely to have higher treatment fidelity scores (**Fig. 3**), with the regression model accounting for 8.6% of the variance in the study treatment fidelity scores (constant = –607.1; 95% CI, –1024.1 to –190.0, $p = 0.005$; $\beta_{year} = 0.3$, 95% CI, 0.1–0.5, $p = 0.004$). For each year, reported treatment fidelity components increased by 0.7% (0.3 treatment fidelity points). Progressive mobility interventions were more comprehensively reported than NMES interventions in the delivery of treatment ($p = 0.010$) and the receipt of treatment ($p < 0.001$) domains with no differences among other intervention types (**eTable 8,** http://links.lww.com/CCM/H486). Domain scores by study (**eTable 9**, http://links.lww.com/CCM/H486), domain scores by study and randomization group (**eTable 10**, http://links.lww.com/CCM/H486), and summary of component reporting by study (**eTables 11-15**, http://links.lww.com/CCM/H486) are available in the supplemental file.

DISCUSSION

We identified important gaps in treatment fidelity reporting of ICU-based physical rehabilitation interventions in this scoping review of 94 unique studies published between 2005 and 2022. Only 19% (8/42) of components were reported across studies, well below the 70% target (10, 12, 29, 30). Intervention group reporting was more comprehensive than comparator groups. Specifically, study design, training providers, delivery of treatment, and receipt of treatment domains were reported better among intervention groups than comparators. Treatment fidelity reporting improved over time. This modest improvement may be attributed to the use of reporting guidelines with similar items to the NIH-BCC, specifically the Consensus on Exercise Reporting Template (CERT) (31) and the Template for Intervention Description and Replication (TIDieR) (32). Ultimately, clinicians and researchers are interested in whether investigators delivered an intervention as planned. Known barriers that may hinder ICU-based physical rehabilitation, include the presence of an endotracheal tube, sedation, staffing ratios (33), concerns for patient safety (e.g., falls), and uncooperative patient behavior (34). However, assessing how these barriers affect fidelity is difficult without understanding the planned and delivered interventions. Transparent reporting is essential.

Treatment fidelity reporting is important because it helps to identify the extent to which the outcomes of a trial represent the true results of the study sample and are not due to methodological error (35). Furthermore, the application and replication of study findings are hindered when treatment protocols are poorly described, inconsistently implemented, and inadequately reported (36).

Intervention group treatment fidelity is important to understand if the intervention was assessed as intended. In contrast to previous reviews which used the NIH-BCC tool to evaluate intervention groups (17, 37), we assessed the presence of treatment fidelity components in both the intervention and comparators groups. Our intervention group results concluding poor reporting are similar to other studies, where none exceeded the 70% component reporting threshold. A systematic review of 22 studies published between 1998 and 2013 enrolled participants undergoing group-based osteoarthritis or chronic low back pain

TABLE 3.
National Institutes of Health Behavioral Change Consortium Treatment Fidelity Reporting Scores by Domain and Overall

	Study	Intervention	Comparator
Treatment fidelity			
Overall score	8 (6, 11)$^{<0.01}$	5 (4, 7)	3 (1, 5)
Overall components reported (%)	19 (14, 26)	24 (19, 33)	14 (5, 24)
Domains			
Study design			
Domain score	6 (4, 8)$^{<0.01}$	4 (3, 5)	2 (1, 3)
Components reported (%)	50 (33, 67)	67 (50, 83)	33 (17, 50)
Training providers			
Domain score	0 (0, 0)$^{<0.01}$	0 (0, 0)	0 (0, 0)
Components reported (%)	0 (0, 0)	0 (0, 0)	0 (0, 0)
Delivery of treatment			
Domain score	1 (0, 2)$^{<0.01}$	1 (0, 1)	0 (0, 1)
Components reported (%)	10 (0, 20)	20 (0, 20)	0 (0, 20)
Receipt of treatment			
Domain score	0 (0, 2)$^{<0.01}$	0 (0, 1)	0 (0, 1)
Components reported (%)	0 (0, 25)	0 (0, 25)	0 (0, 25)
Enactment of treatment			
Domain score	0 (0, 0)$^{0.08}$	0 (0, 0)	0 (0, 0)
Components reported (%)	0 (0, 0)	0 (0, 0)	0 (0, 0)

All values reported as median (first, third quartiles). *p* values for group comparisons included in superscript of domain and study scores.

self-management interventions resulted in an average study score of 35% (17). Another review of 10 studies that enrolled 1154 participants undergoing manual therapy and exercise management for shoulder subacromial pain found overall scores ranged from 9% to 56% across all studies, with an average of 29% (37). Both of these studies reported slightly better scores than our median ICU-based intervention group score of 24%.

Comparator group treatment fidelity is crucial because characteristics of these groups can directly impact the conclusions of studies. For example, a systematic review and meta-analysis of 60 physical rehabilitation ICU studies found that those studies with a comparator group that received fewer than 5 days of physical rehabilitation per week were more likely to show a decrease in mechanical ventilation duration and ICU length of stay (6). However, studies in which comparator groups received 5 or more days of physical rehabilitation demonstrated no difference in mechanical ventilation duration and ICU length of stay (6). Inconsistency in describing the intervention and comparator group plan and conduct can impede accurate interpretation of trial results (10, 38). To comprehensively interpret a physical rehabilitation trial, we must understand what was planned, how providers were trained to carry out that plan, and how providers and participants were able to deliver and receive the planned treatments. Future trialists can use this scoping review to guide comprehensive treatment fidelity reporting and consult the examples of component reporting when determining how to optimize their own treatment fidelity planning and conduct (eTables 11–15, http://links.lww.com/CCM/H486).

We considered using other tools for this review. An et al (39) described a prospective intervention-specific fidelity tool for physical therapy studies. This approach did not suit our current study objectives because we sought a standardized fidelity assessment across different studies and interventions. TIDieR (32) groups

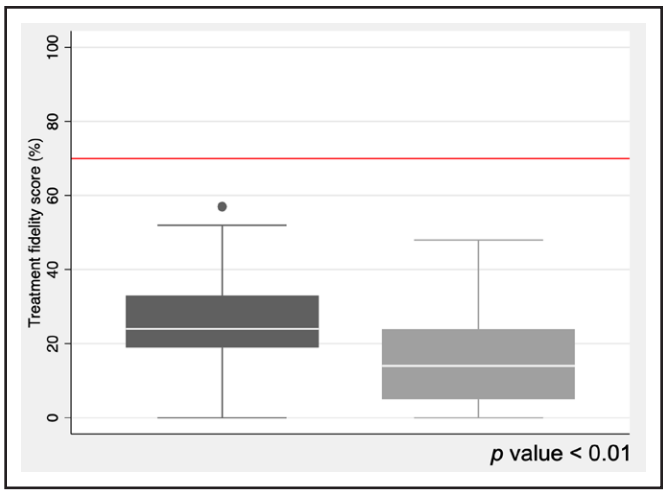

Figure 2. Box plots of intervention (*dark gray*) and comparator (*light gray*) treatment fidelity scores. ICU-based physical rehabilitation reviews of study reporting used 70% as the benchmark for adequate reporting (10, 12), indicated by *solid horizontal line*. *p* value for group comparisons included in bottom of plot. CERT = Consensus on Exercise Reporting Template, TIDieR = Template for Intervention Description and Replication.

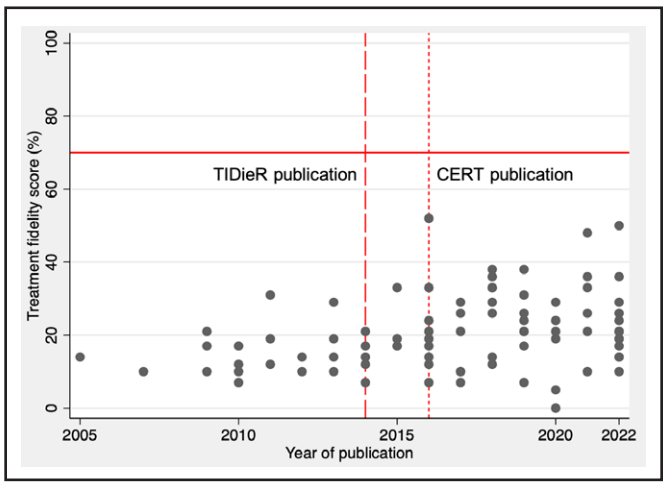

Figure 3. Scatter plot of study treatment fidelity scores by year of publication. ICU-based physical rehabilitation reviews of study reporting used 70% as the benchmark for adequate reporting (10, 12), indicated by *solid horizontal line. Long dashed vertical line* indicates publication year of the Template for Intervention Description and Replication (TIDieR), 2014 (32). *Short dashed red line* indicates publication year of the Consensus on Exercise Reporting Template, 2016 (31).

several important fidelity components under a single item, whereas the NIH-BCC provides more granular detail. For example, TIDieR item 5: Who provided: "For each category of intervention provider (e.g., psychologist, nursing assistant), describe their expertise, background and any specific training given" is represented by four distinct components in the NIH-BCC. Although CERT includes two fidelity components in one item (e.g., item 5: Detailed description of how adherence to exercise is measured and reported) (31), it does not consider how faithfully the providers delivered the intended treatment or how the participant received the treatment. The NIH-BCC captures these concepts in two domains. These details are important for ICU-based physical rehabilitation because a provider may have a rehabilitation plan that requires alteration on some days due to a patient's medical acuity.

Our study has limitations. We excluded non-English publications which may have excluded applicable studies. The capacity of the treatment fidelity tool only allowed for each component to be scored as present (score of 1) or absent (score of 0). The report of multiple unique descriptions per treatment fidelity component provides a more diverse representation of how consistent implementation was with the planned protocol, which may not be captured by a single description. For example, Moss et al (40), reported nine unique methods to ensure the intervention dose was delivered as specified. We only included two group treatment fidelity scores per study, even in those that had more than two groups (7% [n = 7]). To ensure a comprehensive evaluation of treatment fidelity in studies with greater than two groups, we scored the group with the most physical rehabilitation components as the intervention group. We sought all cited protocols, registrations, and supplemental materials associated with each study's primary publication. Although we reviewed over 250 unique documents associated with the 94 included studies, we may have overlooked some eligible documents; however, given our approach of conducting data extraction independently and in duplicate, we have confidence that we likely did not miss applicable documents.

The NIH-BCC treatment fidelity tool was originally developed for use in behavioral sciences research. As such, some of the 21 treatment fidelity components, such as the assessment of nonspecific treatment effects or the use of a treatment manual during a session, may be less applicable to ICU-based physical rehabilitation RCTs. Additionally, the aggregate scoring of the tool assumes all components are equally important. We identified a modified NIH-BCC tool which includes 15 additional

components focused on theoretical and cultural considerations (41). At the time we conducted this review, we did not identify a publication that reported the psychometric properties of the modified tool; thus, we proceeded with the original NIH-BCC tool. Future studies could consider use of the modified NIH-BCC tool, with these limitations in mind. Due to the limited number of PICU studies (n = 3) compared with adult ICU studies (n = 91), we did not conduct any exploratory analyses between pediatric and adult trials. Fidelity reporting components are universal to improving research transparency.

Our review also has several strengths. To our knowledge, this is the first review to comprehensively study treatment fidelity of intervention and comparator groups in ICU-based physical rehabilitation RCTs. We applied all 21 components of the NIH-BCC across both the intervention and comparator groups, which advances the science of treatment fidelity. We used rigorous scoping review methods including a comprehensive search strategy and two reviewers working independently and in duplicate for screening and data extraction. Screening and data extraction were completed by experienced ICU physical therapists with formal research training, which supported clinical understanding of the physical rehabilitation interventions described in the included studies. We included quality assurance procedures to ensure the accuracy of the calculated treatment fidelity scores. Finally, our published protocol (21) includes clear guidelines of how each component of the treatment fidelity tool was assessed, which will allow future researchers to replicate our work.

CONCLUSIONS

A median of only 19% of treatment fidelity components were reported among ICU-based physical rehabilitation trials. Study design was the best reported domain. Few studies reported both planned and provided ICU-based rehabilitation activities in intervention and comparison groups. These gaps limits readers' ability to interpret study results and apply them to clinical practice. Further, it hinders the design of future clinical trials. Authors need to improve treatment fidelity reporting for intervention and comparator groups to optimize the interpretation and application of study results.

Treatment fidelity reporting is the first step to understanding the conduct of ICU-based physical rehabilitation. Researchers can use this work to hone the development of study protocols and manuscripts to optimize treatment fidelity reporting and conduct. Similarly, readers can use this work to guide the treatment fidelity interpretation in ICU-based physical rehabilitation RCTs. Future research could use a qualitative approach to provide insight into solutions to optimize treatment fidelity reporting and determine characteristics associated with treatment fidelity conduct in ICU-based physical rehabilitation.

ACKNOWLEDGMENTS

We thank the McMaster Health Sciences Library team for its support in developing our search strategy and the Hamilton Health Sciences Library Services team for conducting the PRESS assessment. We thank Professor Paul Stratford for his aid in developing the statistical analysis plan.

1 *School of Rehabilitation Sciences, Faculty of Health Sciences, McMaster University, Hamilton, ON, Canada.*

2 *Departments of Pediatrics and Health Research, Evidence, and Impact, McMaster University, Hamilton, ON, Canada.*

3 *Physiotherapy Department, St. Joseph's Healthcare, Hamilton, ON, Canada.*

Supplemental digital content is available for this article. Direct URL citations appear in the printed text and are provided in the HTML and PDF versions of this article on the journal's website (http://journals.lww.com/ccmjournal).

Mr. Farley, Dr. Brooks, Dr. Duffett, and Dr. Kho were involved in conceptualization. Mr. Farley and Dr. Newman were involved in data curation. Mr. Farley, Dr. Hoogenes, and Dr. Kho were involved in formal analysis. Dr. Kho was involved in funding acquisition. Mr. Farley and Dr. Newman were involved in investigation. Mr. Farley, Dr. Newman, Dr. Hoogenes, Dr. Brooks, Dr. Duffett, and Dr. Kho were involved in methodology. Mr. Farley and Dr. Kho were involved in project administration. Mr. Farley and Dr. Kho were involved in resources. Drs. Brooks, Duffett, and Kho were involved in supervision. Mr. Farley, Dr. Newman, and Dr. Kho were involved in validation. Mr. Farley and Dr. Kho were involved in visualization. Mr. Farley and Dr. Kho were involved in writing the original draft. Mr. Farley, Dr. Newman, Dr. Hoogenes, Dr. Brooks, Dr. Duffett, and Dr. Kho were involved in reviewing and editing the writing.

Dr. Kho held a Canada Research Chair in Critical Care Rehabilitation and Knowledge Translation from the Canadian Institutes of Health Research during this study. Mr. Farley received support for article research from the Canadian Institutes

of Health Research. The remaining authors have disclosed that they do not have any potential conflicts of interest.

For information regarding this article, E-mail: farlec3@mcmaster.ca

REFERENCES

1. Appleton RT, Kinsella J, Quasim T: The incidence of intensive care unit-acquired weakness syndromes: A systematic review. *J Intensive Care Soc* 2015; 16:126–136
2. Herridge MS, Tansey CM, Matte A, et al; Canadian Critical Care Trials Group: Functional disability 5 years after acute respiratory distress syndrome. *N Engl J Med* 2011; 364:1293–1304
3. Anekwe DE, Biswas S, Bussières A, et al: Early rehabilitation reduces the likelihood of developing intensive care unit-acquired weakness: A systematic review and meta-analysis. *Physiotherapy* 2020; 107:1–10
4. Schweickert WD, Pohlman MC, Pohlman AS, et al: Early physical and occupational therapy in mechanically ventilated, critically ill patients: A randomised controlled trial. *Lancet* 2009; 373:1874–1882
5. Waldauf P, Jiroutkova K, Krajcova A, et al: Effects of rehabilitation interventions on clinical outcomes in critically ill patients: Systematic review and meta-analysis of randomized controlled trials. *Crit Care Med* 2020; 48:1055–1065
6. Wang YT, Lang JK, Haines KJ, et al: Physical rehabilitation in the ICU: A systematic review and meta-analysis. *Crit Care Med* 2021; 50:375–388
7. Doiron KA, Hoffmann TC, Beller EM: Early intervention (mobilization or active exercise) for critically ill adults in the intensive care unit. *Cochrane Database Syst Rev* 2018; 2018:Cd010754
8. Castro-Avila AC, Seron P, Fan E, et al: Effect of early rehabilitation during intensive care unit stay on functional status: Systematic review and meta-analysis. *PLoS One* 2015; 10:e0130722
9. Tipping CJ, Harrold M, Holland A, et al: The effects of active mobilisation and rehabilitation in ICU on mortality and function: A systematic review. *Intensive Care Med* 2017; 43:171–183
10. O'Grady HK, Reid JC, Farley C, et al: Comparator groups in ICU-based studies of physical rehabilitation: A scoping review of 125 studies. *Crit Care Explor* 2023; 5:e0917
11. Bellg AJ, Borrelli B, Resnick B, et al; Treatment Fidelity Workgroup of the NIH Behavior Change Consortium: Enhancing treatment fidelity in health behavior change studies: Best practices and recommendations from the NIH behavior change consortium. *Health Psychol* 2004; 23:443–451
12. Reid JC, Unger J, McCaskell D, et al: Physical rehabilitation interventions in the intensive care unit: A scoping review of 117 studies. *J Intensive Care* 2018; 6:80
13. Zorko DJ, Reid JC, Unger J, et al: Measurement and reporting of physical rehabilitation interventions in pediatric critical care: A scoping review. *Disabil Rehabil* 2021; 43:3417–3424
14. Moore GF, Audrey S, Barker M, et al: Process evaluation of complex interventions: Medical research council guidance. *BMJ* 2015; 350:h1258
15. Arksey H, O'Malley L: Scoping studies: Towards a methodological framework. *Int J Soc Res Methodol* 2005; 8:19–32
16. 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; 73:852–860
17. Toomey E, Currie-Murphy L, Matthews J, et al: Implementation fidelity of physiotherapist-delivered group education and exercise interventions to promote self-management in people with osteoarthritis and chronic low back pain: A rapid review part II. *Man Ther* 2015; 20:287–294
18. Pussegoda K, Turner L, Garritty C, et al: Systematic review adherence to methodological or reporting quality. *Syst Rev* 2017; 6:131
19. Peters M, Godfrey C, McInerney P, et al: Chapter 11: Scoping reviews (2020 version). In: *JBI Manual for Evidence Synthesis*. Aromataris E, Munn Z (Eds). JBI, 2020, pp 406–449. Available at: https://doi.org/10.46658/JBIMES-20-12. Accessed January 12, 2024
20. Tricco AC, Lillie E, Zarin W, et al: PRISMA extension for scoping reviews (PRISMA-ScR): Checklist and explanation. *Ann Intern Med* 2018; 169:467–473
21. Farley C, Newman A, Hoogenes J, et al: Treatment fidelity in randomized controlled trials of physical rehabilitation in the intensive care unit: A protocol for a scoping review. 2022. Available at: https://osf.io/6egfm. Accessed January 12, 2024
22. Levine S, Nguyen T, Taylor N, et al: Rapid disuse atrophy of diaphragm fibers in mechanically ventilated humans. *N Engl J Med* 2008; 358:1327–1335
23. World Health Organization: Rehabilitation in Health Systems. Geneva, World Health Organization, 2017. Available at https://www.who.int/publications/i/item/9789241515986. Accessed January 12, 2024
24. Glanville J, Kotas E, Featherstone R, et al: Which are the most sensitive search filters to identify randomized controlled trials in MEDLINE? *J Med Libr Assoc* 2020; 108:556–563
25. Glanville J, Foxlee R, Wisniewski S, et al: Translating the cochrane EMBASE RCT filter from the ovid interface to embasecom: A case study. *Health Info Libr J* 2019; 36:264–277
26. Glanville J, Dooley G, Wisniewski S, et al: Development of a search filter to identify reports of controlled clinical trials within CINAHL Plus. *Health Info Libr J* 2019; 36:73–90
27. McGowan J, Sampson M, Salzwedel DM, et al: PRESS peer review of electronic search strategies: 2015 guideline statement. *J Clin Epidemiol* 2016; 75:40–46
28. Johnson-Kozlow M, Hovell MF, Rovniak LS, et al: Fidelity issues in secondhand smoking interventions for children. *Nicotine Tob Res* 2008; 10:1677–1690
29. Braga LH, McGrath M, Easterbrook B, et al: Quality of reporting for randomized controlled trials in the hypospadias literature: Where do we stand? *J Pediatr Urol* 2017; 13:482.e1–482.e9
30. Hadi MA, McHugh GA, Conaghan PG: Quality of reporting of harms in randomised controlled trials of pharmacological interventions for rheumatoid arthritis: A systematic review. *Evid Based Med* 2017; 22:170–177
31. Slade SC, Dionne CE, Underwood M, et al: Consensus on Exercise Reporting Template (CERT): Explanation and elaboration statement. *Br J Sports Med* 2016; 50:1428–1437

32. Hoffmann TC, Glasziou PP, Boutron I, et al: Better reporting of interventions: Template for Intervention Description and Replication (TIDieR) checklist and guide. *BMJ (Clin Res Ed)* 2014; 348:g1687
33. Parry SM, Knight LD, Connolly B, et al: Factors influencing physical activity and rehabilitation in survivors of critical illness: A systematic review of quantitative and qualitative studies. *Intensive Care Med* 2017; 43:531–542
34. Cooper D, Gasperini M, Parkosewich JA: Nurses' perceptions of barriers to out-of-bed activities among patients receiving mechanical ventilation. *Am J Crit Care* 2021; 30:266–274
35. Patino CM, Ferreira JC: Internal and external validity: Can you apply research study results to your patients? *J Bras Pneumol* 2018; 44:183
36. Moncher FJ, Prinz RJ: Treatment fidelity in outcome studies. *Clin Psychol Rev* 1991; 11:247–266
37. Ribeiro DC, Spiers K, Thomas L, et al: Monitoring, implementation and reporting of interventions in a selection of trials assessing exercise therapy for the management of shoulder subacromial pain: A cross-sectional investigation. *BMJ Open* 2021; 11:e044462
38. Kho ME, Berney S, Connolly B: Physical rehabilitation in the intensive care unit: Past, present, and future. *Intensive Care Med* 2023; 49:864–867
39. An M, Dusing SC, Harbourne RT, et al; START-Play Consortium: What really works in intervention? Using fidelity measures to support optimal outcomes. *Phys Ther* 2020; 100:757–765
40. Moss M, Nordon-Craft A, Malone D, et al: A randomized trial of an intensive physical therapy program for patients with acute respiratory failure. *Am J Respir Crit Care Med* 2016; 193:1101–1110
41. Borrelli B: The assessment, monitoring, and enhancement of treatment fidelity in public health clinical trials. *J Public Health Dent* 2011; 71:S52–S63
42. Thabane L, Ma J, Chu R, et al: A tutorial on pilot studies: The what, why and how. *BMC Med Res Methodol* 2010; 10:1