



Quality and Reporting of Cluster Randomized Controlled Trials Evaluating Occupational Therapy Interventions: A Systematic Review

OTJR: Occupation, Participation and Health
1–11
© The Author(s) 2015 
Reprints and permissions:
sagepub.com/journalsPermissions.nav
DOI: 10.1177/1539449215618625
otj.sagepub.com


Ema Tokolahi¹, Clare Hocking¹, Paula Kersten¹,
and Alain C. Vandal^{1,2}

Abstract

Growing use of cluster randomized control trials (RCTs) in health care research requires careful attention to study designs, with implications for the development of an evidence base for practice. The objective of this study is to investigate the characteristics, quality, and reporting of cluster RCTs evaluating occupational therapy interventions to inform future research design. An extensive search of cluster RCTs evaluating occupational therapy was conducted in several databases. Fourteen studies met our inclusion criteria; four were protocols. Eleven (79%) justified the use of a cluster RCT and accounted for clustering in the sample size and analysis. All full studies reported the number of clusters randomized, and five reported intercluster correlation coefficients (50%): Protocols had higher compliance. Risk of bias was most evident in unblinding of participants. Statistician involvement was associated with improved trial quality and reporting. Quality of cluster RCTs of occupational therapy interventions is comparable with those from other areas of health research and needs improvement.

Keywords

cluster randomized controlled trials, methods, occupational therapy, research design, systematic review

Background

There is growing use of cluster randomized control trials (RCTs) in health care research, with systematic reviews of cluster RCTs emerging over recent decades as the methodology gained credibility and prominence (Diaz-Ordaz, Froud, Sheehan, & Eldridge, 2013; Eldridge & Kerry, 2012; Isaakidis & Ioannidis, 2003). No one has reviewed the quality and reporting of cluster RCTs within occupational therapy, although methodological evaluations of occupational therapy RCTs have been published (Bennett, Hoffmann, McCluskey, Coghlan, & Tooth, 2013; Bennett et al., 2007; Kim, Yoo, Jung, Park, & Park, 2012; Moberg-Mogren & Nelson, 2006; Norton-Mabus & Nelson, 2008). It is important to evaluate the quality of current studies to ensure robust research of this kind is being conducted within the occupational therapy profession, to identify the unique challenges and opportunities created by the evaluation of occupational therapy interventions using this design and to inform future research design.

Cluster RCT Design

Cluster RCTs are defined as having “groups or clusters of individuals rather than individuals themselves . . . randomized to intervention arms” (Eldridge & Kerry, 2012, p. 3).

There are several pragmatic reasons for selecting this design. In interventions or population studies, randomizing by cluster can limit the potential for contamination between study arms. For example, if someone in a helper role was given additional training, it would be impractical to randomly allocate their clients to receive or not receive the benefits of that training as the new information cannot be “unlearned”; this scenario would also be ethically questionable (Barbui & Cipriani, 2011). Researchers may be interested in outcomes at the cluster level (i.e., focus on change within and between clusters), the individual level (i.e., focus on change within and between individuals), or both (Eldridge & Kerry, 2012). Furthermore, there are convenience and cost benefits to investigating an intervention in clusters (Isaakidis & Ioannidis, 2003).

Cluster designs have several implications for sample size calculations and data analysis. The individuals within clusters are likely to be more homogeneous than those from a

¹Auckland University of Technology, New Zealand

²Health Intelligence & Informatics, Ko Awatea, Auckland, New Zealand

Corresponding Author:

Ema Tokolahi, Centre for Person-Centred Research, Auckland University of Technology, Private Bag 92006, Auckland 1142, New Zealand.
Email: etokolah@aut.ac.nz

sample drawn from the general population. For example, members of a family are likely to have shared eating habits, similar attitudes to health and recovery, or comparable exposure to environmental factors. These similarities necessitate a greater sample size than individually randomized trials, with adjustments made to the data analysis plan that account for clustering (Campbell, Thomson, Ramsay, MacLennan, & Grimshaw, 2004; Eldridge, Ashby, & Kerry, 2006; Teerenstra, Eldridge, Graff, de Hoop, & Borm, 2012). Failing to account for clustering increases the risk of Type 1 error (e.g., finding a significant difference where there is not one). The need to address and report these additional considerations was formalized in the Consolidated Standards of Reporting Trials extension to cluster randomized trials (CONSORT extension), which highlights the key areas to report on when utilizing this research design, over and above the reporting guidelines for a standard RCT (Campbell, Elbourne, & Altman, 2004).

Occupational Therapy and Cluster RCTs

Occupational therapy focuses on “the nature, balance, pattern and context of occupations and activities in the lives of individuals, family groups and communities” (Creek, 2003, p. 8). The main aim of therapy is to enable the individual (family or community) to make occupational choices that “maintain, restore or create a match, beneficial to the individual, between the abilities of the person, the demands of her/his occupations in the areas of self-care, productivity and leisure, and the demands of the environment” (Creek, 2003, p. 8). There is significant potential for occupational therapy interventions to be implemented and investigated using a cluster design, such as with naturally occurring groups (e.g., nursing home residents, children in schools) or with therapists who are randomly allocated to deliver a novel intervention. Consequently, researchers need to be informed about the conduct of robust cluster RCTs.

This review investigated the characteristics and quality of conduct and reporting of cluster RCTs evaluating occupational therapy interventions. Guidance is provided for future research to promote robust design, conduct, and analysis of high-quality cluster RCTs in occupational therapy. Potential factors that may influence study quality are identified.

Method

Inclusion and Exclusion Criteria

We included cluster RCTs conducted prior to July 2015, reported in peer-reviewed journals, that evaluated an occupational therapy intervention. Studies were identified as cluster RCTs if sufficient detail was provided to determine that participants were randomized by groups (clusters) and the study was not quasi-experimental or non-randomized. Interventions were defined as occupational therapy if explicitly labeled as

such or implicitly identifiable as such due to the developer or facilitator being an occupational therapist and/or the underlying theory being drawn primarily from occupational therapy and science literature. Interventions conducted by occupational therapists *or* those from other disciplines were included if the intervention was recognized as a valid occupational therapy intervention; however, interventions that were conducted by an occupational therapist *and* those from other disciplines were considered inter/multi-disciplinary and excluded. Articles reporting protocols and findings were included. No date limits were applied, full-text articles available in English were included, and no studies were excluded on the basis of quality because one of the review objectives was to provide a description of quality.

Data Sources and Search Methods

Databases were searched in January 2015 and again in July 2015 to ensure completeness. Search terms were used in combination to systematically search titles, abstracts and keywords from the following databases: EBSCO—includes Cumulative Index of Nursing and Allied Health Literature (CINAHL) Plus, MEDLINE, Health Business Elite, Psychological and Behavioural Sciences Collection; OVID—includes Allied and Complementary Medicine Database (AMED), Cochrane Health Databases, Evidence Based Medicine (EBM) Reviews, Educational Resources Information Centre (ERIC), PsychInfo; PubMed; and SCOPUS. Specific key words and phrases used were “RCT,” “random* control* trial*,” “random* clinical trial*,” “OT,” “occupational therapy,” “cluster*,” “nest*,” and “group” within three words from “random*.” The first author also conducted a manual search of all references from included articles to identify additional eligible reports.

Data Extraction

The titles and abstracts were screened by the first author, who obtained the full text of those definitely or possibly meeting the inclusion criteria. All full-text articles, for which eligibility was clear or unclear, were screened by all authors and a decision made by consensus whether or not to include it. All included reports were reviewed by E.T., 10 of which were independently reviewed among C.H., P.K., and A.C.V.; one was reviewed by a fifth reviewer, completely independent of the authors as it was a paper written by them. Disagreements or queries about the reviewed articles were discussed between the authors and resolved by consensus agreement.

Data Extracted

To describe the range of trials, and potential moderators of quality (Diaz-Ordaz et al., 2013), we extracted data about the journal, sample, and trial design characteristics. Journal characteristics extracted were journal name, publication year, and

endorsement of the CONSORT statement. Endorsement was rated as “present” if adherence to the CONSORT statement was indicated as compulsory or a preference in the journal’s guideline for authors and “absent” if not mentioned. Sample characteristics extracted were trial location, clinical problem, and intervention. Trial design characteristics extracted were unit of randomization, number of clusters randomized, average cluster sizes (as reported, calculated, or planned), and whether or not a statistician was a co-author. Statistician co-authorship was deemed an indicator of active statistician involvement, and this criterion was satisfied if an author was identified as from a department/unit of biostatistics/epidemiology or mathematics or was clearly designated as a statistician or epidemiologist. When this was not possible to ascertain, this was recorded as absent.

To assess the quality of trials, we used five design and analysis recommendations, as reported and used by Eldridge, Ashby, Feder, Rudnicka, and Ukoumunne (2004) in a systematic review of cluster randomized trials in primary health. The five items were as follows: justifies the use of a cluster design; includes at least four clusters per intervention group; allows for clustering in sample size calculation; uses matching, stratification, or an alternative means of reducing chance imbalances at baseline; and allows for clustering in analysis. Data were extracted to evidence whether or not each recommendation had been satisfied. The authors deemed the original sixth item, “allows for confounding in analysis,” to be superfluous as there were already recommendations to stratify or match the clusters if potential confounders were predetermined and to account for clustering in the analysis. It was agreed that any further chance confounding would not be specific to the cluster-RCT nature of the design.

To assess the quality of trial reporting, the 12 recommendations from Eldridge et al. (2004) were adopted, which incorporate requirements from the CONSORT extension. Recommendations were as follows: cluster RCT identified in the title, includes an estimate of an intercluster correlation coefficient (ICC), lists number of clusters randomized, describes baseline comparison of clusters and individuals, lists average cluster size, explains whether analysis is conducted at the cluster or individual level, reports on loss to follow-up of clusters, and reports on loss to follow-up of individuals. The last recommendation from Eldridge et al. (2004) stipulates reporting loss to follow-up of individuals from *within clusters*. However, a more recent article is less specific in this regard (Ivers et al., 2011), and the CONSORT extension specifies only reporting of “losses and exclusions for both clusters and individual participants” (Campbell, Elbourne, et al., 2004, p. 10). If articles reported on loss to follow-up of individuals *between study arms*, then that was considered adequate. Data were extracted to evidence whether or not each recommendation had been satisfied in full, partially, or not at all. Due to overlap between three of the reporting recommendations and the conduct recommendations (i.e., justification of a cluster design, explaining how the sample size and analysis account

for between-cluster variations), these have been reported once only to avoid duplication. Furthermore, due to protocol-only reports having limited data available to report, they were reviewed separately with credit given where there was an explicit plan to address the reporting recommendations.

Risk of bias within the trials was assessed using a tool developed by the Cochrane Collaboration to measure low, unclear, or high risk of bias within and across trials (Higgins et al., 2011). Data were extracted to evidence these ratings. When the term *randomized* was absent but a randomization procedure was referred to, such as “drawing of lots,” this was considered sufficient to confirm a randomization sequence had occurred. When allocation concealment was difficult to interpret, but it was evident recruitment of individuals within the clusters (not just recruitment of the clusters) was completed prior to randomization, this was considered a low risk of bias to the trial findings. When therapists providing the intervention were unblinded, but not participants, this was not considered to increase risk of bias as they were part of the intervention. However, therapists and participants receiving the intervention were considered to entail unblinding and to contribute to increased risk of bias. Potential risk of bias from incomplete outcome data was considered low if there was limited data loss (<5%; IBM, 2011), if reasons for missing outcome data were unlikely to be related to the outcome, and if an intention-to-treat (ITT) analysis was conducted.

Analysis of quality, reporting, and risk of bias focused on the primary outcome(s) for each study. When more than one article reported on the same study (e.g., protocol and findings), the combined data were used and the best data examples extracted.

Analysis

We present descriptive statistics of trial characteristics, grouped by journal, sample, and trial design. For each of the quality and reporting items, we report the outcomes for each study and the percentage of studies that fulfilled each criterion. Risk of bias judgments is presented across and within studies. To explore potential moderators of quality, we planned to present year of publication (pre/post publication of the CONSORT extension), endorsement of the CONSORT statement and involvement of a statistician, and the number (and percentage) of studies adhering to three key items reflecting major methodological issues with the validity of cluster RCTs. These items were accounting for clustering in sample size, accounting for clustering in data analysis, and potential selection bias (Diaz-Ordaz et al., 2013), determined by risk of bias related to allocation concealment.

Results

Trial Characteristics and Interventions

In this review, 18 manuscripts were included reporting on 14 clinical trials from seven different countries, and from a range

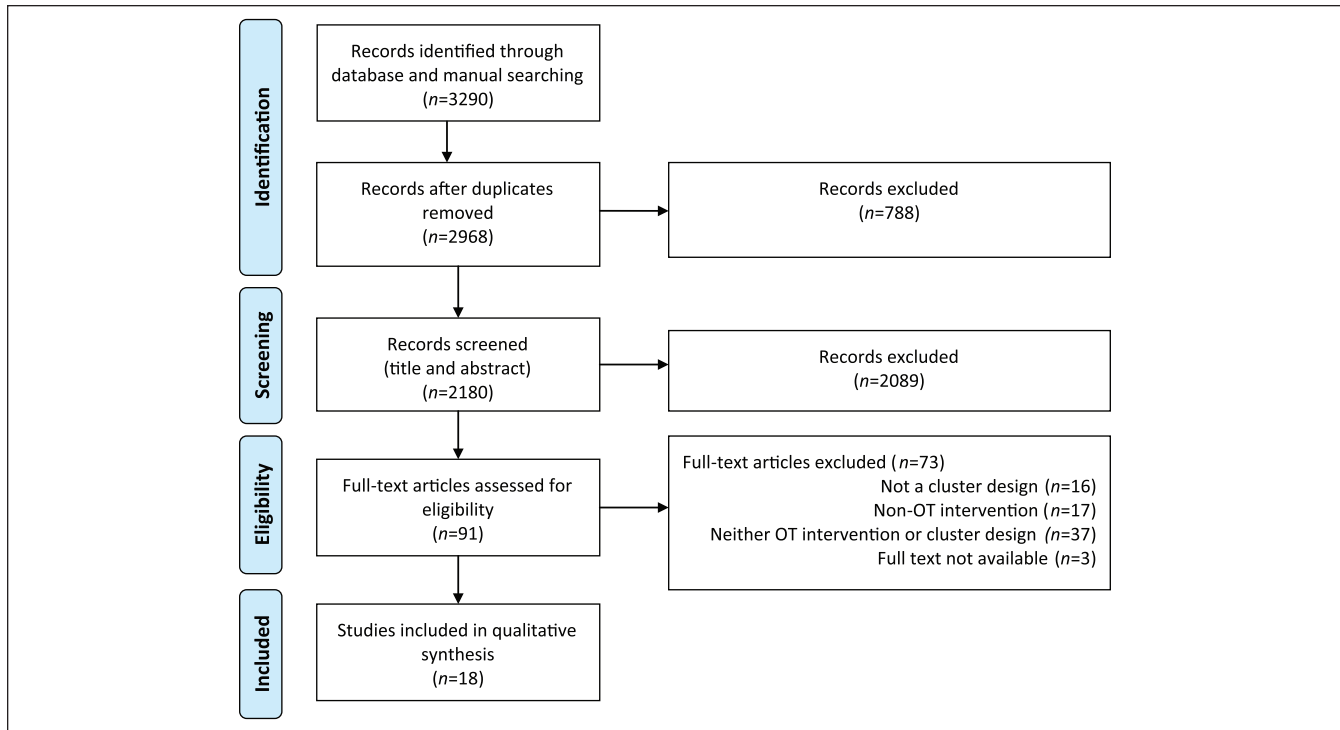


Figure 1. PRISMA flow diagram of the identification process for the sample of 18 articles describing cluster RCTs included in this review. Note. OT = occupational therapy; PRISMA = preferred reporting items for systematic reviews and meta-analyses; RCTs = randomized control trials.

of allied health, medical, and condition-oriented journals (Figure 1). Four were protocols for studies with findings yet to be reported. Studies were published from 2006 onward, 12 in journals that endorsed the CONSORT statement and 11 had statistician involvement (Table 1).

Three of the interventions were enriched day programs aimed at increasing engagement and participation in meaningful occupations; three involved educating staff about new treatment approaches. The median number of clusters randomized was 15 (range = 2-228) with the median of the average cluster sizes being 11.5 (range = 1.77-54; \bar{x} = 13.8).

Due to differing expectations between what can and should be reported in a full-trial report ($n = 10$) and a protocol ($n = 4$), these are separated out in each of the tables.

Trial Quality

Table 2 presents adherence of the studies to the conduct recommendations. Having at least four clusters per study arm was achieved in 86% of the studies ($n = 12$), and stratification or matching was achieved in 79% ($n = 11$). The remaining three recommendations—justification for using a cluster design, allowing for clustering in the sample size calculation and in the analysis—were achieved in 79% of the studies ($n = 11$).

Trial Reporting

All 10 studies explained whether analysis was conducted at the individual or cluster level (see Table 2). Eight studies

were identified as cluster RCTs in the title (80%), and one was identified as a cluster RCT in the abstract (Mozley et al., 2007). In the remaining study, the unit of randomization was acknowledged as occurring by cluster in the methods section (Eklund, Gunnarsson, Sandlund, & Leufstadius, 2014). Five studies reported an ICC (range = 0.04-0.37) used to calculate the sample size (50%), and one reported the calculated design effect, for which an ICC must have been assumed (Eysen et al., 2013). All studies explicitly listed the number of clusters randomized and reported an average cluster size or sufficient information to calculate one. Most studies described a baseline comparison of individuals (90%); however, only five described the baseline comparison of clusters (50%). Nine reported on loss to follow-up of clusters or made it clear no clusters were lost (90%); eight reported on loss to follow-up of individuals from between the study arms (80%).

Four protocol reports were reviewed, of which three identified as cluster RCTs in the title. Three protocols provided ICCs (range = 0.01-0.05), and one provided a design effect (Tokolahi, Hocking, Kersten, & Vandal, 2014). As these were protocols, there were no data to report when describing baseline comparisons; all four described inclusion and exclusion criteria for the clusters; however, prior to recruitment and randomization, comparison between the study arms was not feasible. Three of the studies reported inclusion and exclusion criteria for the participants, with three explicitly stating that baseline demographics of individuals would be collected and reported. Similarly, reporting on loss to follow-up is limited by lack of data to report, and none of the

Table 1. Characteristics of Studies Included in Review.

Study	First author	Journal title	Article characteristics		Sample characteristics			Trial design characteristics			
			Publication year	Endorse CONSORT	Location	Clinical problem	Intervention (vs. No intervention unless otherwise stated)	Cluster units	Number of clusters	Cluster sizes	Statistician involvement
Full studies in alphabetical order (except where protocol and finding reports are grouped and had different first authors)											
1	Bundy	<i>Australian Occupational Therapy Journal</i>	2011	✓	Australia	Public health Children	Play materials in schools and parent/teacher education	Schools	12 (6:6)	18,4 ^a	Yes
2	Engelen Dopp (Protocol)	<i>Preventative Medicine BMC Geriatrics</i>	2013 2011	✓ ✓	The Netherlands	Dementia Older persons	Combined implementation of COTID training vs. standard COTID training	Nursing homes; hospitals; mental health services	45 (17 control; 28 intervention)	2,10 ^a	Yes
3	Dopp Eklund	<i>Clinical Rehabilitation Scandinavian Journal of Occupational Therapy</i>	2014 2014	✓ ✓	Sweden	Mental health Adults	Enriched day program vs. standard day program	Psychiatric day centers	8 (4:4)	54 ^a	No
4	Eyssen	<i>Disability and Rehabilitation BMC Paediatrics</i>	2013 2007	✓ ✓	The Netherlands Canada	Multiple sclerosis Adults Cerebral palsy Children	Client-centered vs. traditional occupational therapy Child-focused vs. task-focused occupational therapy	Hospital; Rehab centers Therapist	13 (7 control; 6 intervention) 79	19,9 ^a 1,77	Yes Yes
5	Law	<i>Developmental Medicine and Child Neurology</i>	2011	✓							
6	Mozley	<i>Aging and Mental Health</i>	2007	X	The United Kingdom	Depression Older people	Care home activity project	Residential care homes	8 (4:4)	17,88 ^a	No
7	Sackley (protocol)	<i>BMC Neurology</i>	2012	✓	The United Kingdom	Stroke Adults	Occupational therapy vs. TAU (no occupational therapy)	Stroke-related disability care homes	228 (114:114)	4,6 ^a 4=median	Yes
8	Sackley	<i>The BMJ Stroke</i>	2015 2006	✓ ✓	The United Kingdom New Zealand	Stroke Adults Stroke Adults	Occupational therapy vs. TAU (no occupational therapy) COPM to structure goal-setting vs. traditional goal-setting	Stroke-related disability care homes Inpatient rehabilitation centers	12 (6:6) 4 (2:2)	9,8 ^a 10,25 ^a 9-11,5= medians	Yes No
9	Taylor	<i>Clinical Rehabilitation</i>	2012	✓			Enriched day program	Care homes	16 (8:8)	13,13 ^a	Yes
10	Wenborn	<i>International Journal of Geriatric Psychiatry</i>	2013	X	The United Kingdom	Dementia Adults					
Protocols											
1	Barzel	<i>Trials</i>	2013	✓	Germany	Stroke Adults	HOMECIMT vs. TAU	Practice centers	48 (24:24)	3	Yes
2	Killiaspy	<i>BMC Psychiatry</i>	2013	✓	The United Kingdom	Mental health Adults	GetREAL staff training	Mental Health Rehabilitation Centers	40 (20:20)	12	Yes
3	McCluskey	<i>International Journal of Stroke</i>	2013	✓	Australia	Stroke Adults	Out and About Program vs. written clinical guidelines	Community teams	20 (10:10)	15	Yes
4	Tokolahi	<i>BMC Psychology</i>	2014	✓	New Zealand	Public health Children	Kia Piki te Hauora	Schools	14 (7:7)	11	Yes

Note. COTID = Community occupational therapy in Dementia; TAU = treatment as usual; COPM = Canadian occupational performance measure; HOMECIMT = Modified home-based constraint-induced movement therapy; GetREAL = Staff training to increase service user engagement.

^aAverage cluster sizes calculated based on n/number of clusters.

Table 2. Conduct and Reporting of Recommendations Within and Between Trials.

Studies	Full studies in alphabetical order (except where protocol and finding reports are grouped and had different first authors)										Protocols					
	Bundy et al. (2011)	Döpp et al. (2011)	Döpp et al. (2014)	Eklund, Gunnarsson, Sandlund, and Leufstaadus (2014)	Eyssen, et al. (2013)	Law et al. (2007)	Law et al. (2011)	Mozley, et al. (2007)	Sackley, et al. (2006)	Sackley, et al. (2012)	Taylor et al. (2012)	Wenborn et al. (2013)	Barzel et al. (2013)	Killaspy et al. (2013)	McCluskey et al. (2013)	Tokolahti, Hocking, Kersten, and Vandal (2014)
Trial conduct recommendations																
Justification of use of clustered design	✓	✓	✓	—	✓	—	—	—	✓	✓	✓	✓	✓	✓	✓	✓
At least four clusters per intervention group	✓	✓	✓	—	✓	✓	✓	✓	✓	—	—	✓	✓	✓	✓	✓
Allows for clustering in sample size calculation	✓	✓	✓	—	✓	✓	✓	✓	—	—	—	✓	✓	✓	✓	✓
Uses matching or stratification	—	Strat	Strat	Strat	Strat	Strat	Strat	Match	Strat	Strat	—	Match	Strat	—	Minim	Strat
Allows for clustering in analysis	✓	✓	✓	—	✓	—	—	—	✓	✓	✓	✓	✓	✓	✓	✓
Trial reporting recommendations																
Identified as a cluster RCT in the title	✓	✓	✓	—	✓	—	—	—	✓	✓	✓	✓	✓	✓	—	✓
Includes estimate of ICC	0.04	0.05	0.05	—	DE	0.1	—	—	—	0.37	—	0.05	0.05	0.04	0.01	DE
Lists number of clusters randomized	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Describes baseline comparison of clusters	✓	✓	✓	—	P	—	—	✓	✓	✓	—	—	—	—	—	—
Describes baseline comparison of individuals	✓	✓	✓	—	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Lists average cluster size	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Explains whether analysis conducted at cluster or individual level	Ind	Clu	Clu	Ind	Ind	Ind	Ind	Ind	Ind	Ind	Ind	Ind	Ind	Clu	Clu	Ind
Reports on loss to follow-up of clusters	✓	✓	✓	✓	✓	—	✓	✓	✓	✓	✓	✓	—	—	—	—
Reports on loss to follow-up of individuals within clusters	—	✓	✓	—	✓	—	—	✓	✓	✓	✓	✓	—	—	—	—

Note. X = recommendation not satisfied; ✓ = recommendation satisfied; ICC = intercluster correlation coefficient; DE = design effect; P = recommendation partially satisfied; Strat = stratified; Match = matching; RCTs = randomized control trials; Ind = individual outcomes; Clu = cluster outcomes; NA = not applicable.

protocols made explicit statements about how loss of clusters might be reported. One did identify that recruitment had been increased to accommodate probable loss to follow-up of clusters (Killaspy et al., 2013). None of the protocols explicitly stated that loss to follow-up of individuals would be reported; however, three of the protocols made explicit statements about using ITT analysis, and two planned to investigate missing data.

Trial Risk of Bias

Eight of the 14 studies reported use of appropriate random sequence generation (57%; see Table 3). Allocation concealment was conducted in a manner suggesting low risk of bias in 11 studies (79%); one revealed allocation to assist with recruitment (Taylor et al., 2012); the remaining two did not provide sufficient information to determine whether allocation concealment had occurred.

Blinding of participants and personnel was the most frequently rated item for potential risk of bias, with 12 studies judged to be at high risk of unblinding because participants were not blinded to allocation (86%). Outcome assessors were blinded to participant and cluster allocation in nine studies (65%), of which four planned or reported an assessment of whether or not blinding had been upheld. In two studies, the outcome assessors were unblinded (14%), and the remaining three studies did not provide sufficient information to determine whether or not blinding of outcome assessors occurred.

Incomplete data were reported to be managed in such a way as to minimize bias in seven of the studies (50%). Two of the studies were rated as having potentially high risk of bias regarding incomplete data (14%) due to a high proportion of missing data, and variability and lack of clarity in how missing data were imputed. Thirty-six percent of the studies (five) did not provide sufficient information for a decision to be made. All of the studies identified a primary outcome to be measured, and these were all reported in the 10 findings reports, resulting in the risk of bias due to selective reporting being rated *low* for all 14 studies.

Moderators of Trial Quality and Reporting

All of the studies post-dated the 2004 CONSORT extension, and only two articles appeared in the same journal; therefore, these factors were not appropriate to explore as potential moderators of quality. We included 10 primary reports in nine journals, of which eight endorsed the CONSORT statement and seven had a statistician as a co-author (see Table 4). All three quality items appeared to be positively associated with studies involving a statistician compared with those without. Adherence to one of the quality items (accounting for clustering in the analysis) appeared to be associated with studies that endorsed the CONSORT statement. These findings could not be confirmed through statistical analysis due to insufficient sample size.

Discussion

This review was the first to evaluate the quality and reporting of cluster RCTs evaluating occupational therapy interventions. Occupational therapy researchers have been slower to adopt this design, with trials using the cluster RCT design being conducted and reported only within the last decade and mostly in the last 5 years. All the studies reviewed post-dated the CONSORT extension, suggesting this may have provided some necessary guidance and direction for researchers.

Several interventions were not overtly labeled *occupational therapy*, which creates challenges for combining and reviewing intervention quality and assembling sufficient evidence of quality to justify occupational therapy interventions. In all but one of the studies where a justification for cluster randomization was provided, the reason for selecting a cluster RCT design was to prevent contamination. Clear justification for the use of a cluster RCT is important for defending methodological choices and for informing prospective researchers about important considerations of this design. No studies in this review were identified as using a cluster RCT design but reporting a standard RCT design, although it is possible the search strategy may not have identified such studies.

Describing the baseline comparison between clusters is important for identifying any chance imbalances that could potentially bias the findings (Eldridge et al., 2004); half of the 10 full studies reviewed provided this information. This omission may be because of high compliance with the use of stratification or matching to minimize the likelihood of chance imbalances at baseline.

The quality of the studies reviewed was reasonable, with adherence to quality indicators ranging from 79% to 86%. This review found more than three quarters of studies accounted for clustering in the sample size calculation (79%), which was superior to cluster RCTs conducted over a similar period in other health research (adherence range = 36%-65%; Diaz-Ordaz et al., 2013). However, lack of adherence was observed for reporting an ICC or DE, which is required to determine a sample size that accounts for clustering. Half of the full trials reported an ICC or DE, although all of the protocols did. In one protocol, the ICC was based on a pilot study (McCluskey et al., 2013); however, for the remainder, an evidence-based justification for how the ICC had been established was not described. Of the five studies that did not report an ICC or DE, three were pilot or feasibility trials, in which one of the aims was to calculate an ICC for a full trial (Mozley et al., 2007; Sackley et al., 2006; Taylor et al., 2012). Future researchers may benefit from basing their sample size calculations on similar ICCs to those reported in these studies if the context is similar to their population of interest. Although there was a wide spread of ICCs (0.01-0.37), the most common ICCs were 0.04 and 0.05. Overall, these findings suggest occupational therapy researchers have grasped the importance of

Table 3. Risk of Bias Within and Across Studies.

Studies	Random sequence generation	Allocation concealment	Blinding of participants and personnel	Blinding of outcome assessment	Incomplete outcome data	Selective reporting
Full studies in alphabetical order (except where protocol and finding reports are grouped and had different first authors)						
Bundy et al. (2011), Engelen et al. (2013)	?	+	-	-	+	+
Dopp et al. (2011), Döpp et al. (2014)	?	?	-	+	-	+
Eklund, Gunnarsson, Sandlund, and Leufstadius (2014)	-	+	+	+	?	+
Eyssen et al. (2013)	+	+	+	+	?	+
Law et al. (2007), Law et al. (2011)	+	+	-	+	-	+
Mozley et al. (2007)	+	+	-	-	?	+
Sackley et al. (2006)	+	+	-	+	?	+
Sackley et al. (2012), Sackley et al. (2015)	+	+	-	+	+	+
Taylor et al. (2012)	+	-	-	+	+	+
Wenborn et al. (2013)	?	+	-	?	?	+
Protocols						
Barzel et al. (2013)	?	+	-	+	+	+
Killaspy et al. (2013)	?	?	-	+	+	+
McCluskey, et al. (2013)	+	+	+	+	+	+
Tokolahi, Hocking, Kersten, and Vandal (2014)	+	+	-	+	+	+

Note. ? = risk of bias unclear; + = low risk of bias; - = high risk of bias.

Table 4. Number (Percentage) of Studies Adhering to Quality Criteria by Key Characteristics.

Criterion	Total (n = 10)	Statistician involvement		CONSORT endorsed	
		Yes (n = 7)	No (n = 3)	Yes (n = 8)	No (n = 2)
Allows for clustering in sample size calculation	7 (70)	6 (86)	1 (33)	5 (63)	2 (100)
Allows for clustering in analysis	7 (70)	6 (86)	1 (33)	6 (75)	1 (50)
Recruitment bias as determined by level of allocation concealment					
Low	8 (80)	6 (86)	2 (66)	6 (75)	2 (100)
High	1 (10)	0 (0)	1 (33)	1 (13)	0 (0)
Unclear	1 (10)	1 (14)	0 (0)	1 (13)	0 (0)

ensuring their studies are sufficiently powered for statistical analysis and are largely following recommended guidelines to achieve this.

Accounting for the clustering in the analysis itself is a priority area for improving future research, with 79% adherence to this quality criterion falling at the lower range of comparable studies, where adherence ranged between 78% and 88% (Diaz-Ordaz et al., 2013). Most of the trials accounting for clustering in the analysis reported the use of a mixed effects modeling with the cluster as a random effect. One used tests based on adjusted standard errors (Sackley et al., 2006). This is consistent with the techniques described in a systematic review of cluster RCTs in stroke (Sutton, Watkins, & Dey, 2013).

All studies reported an average cluster size or sufficient information to calculate the average. However, this does not always translate to sufficient information being available to calculate the variation in cluster sizes if these are not fixed. Cluster size variation is important to consider in relation to sample size as it can significantly affect the statistical power of data analyzed in cluster RCTs (Eldridge et al., 2006). Future studies can improve reporting in this area by stating the range of the cluster sizes and/or providing a coefficient of variation.

Potential risk of bias in the randomization process was identified in almost half the studies reviewed (43%), largely due to insufficient information being provided. This may reflect limitations in word allowances or decisions made in the editorial process.

One of the greatest challenges for complex interventions, such as occupational therapy, is the difficulty of blinding participants and personnel who are actively engaged in the intervention (Medical Research Council, 2008). The two studies that overcame this risk either blinded participants to the type of occupational therapy received, client-centered or traditional (Eyssen et al., 2013), or involved only unblinding the cluster guardians who introduced the experimental intervention to therapists as part of implementing best practice guidelines (McCluskey et al., 2013). Participant blinding will continue to challenge researchers of occupational therapy interventions and should be acknowledged as a potential bias when the study design cannot overcome this.

Reporting on loss to follow-up of clusters was reasonable; reporting on loss of individuals between study arms was lower, and only one study provided additional information about the loss of individuals from within clusters (Sackley

et al., 2015). This added information is valuable for enabling the reader to understand any patterns or potential bias that may have emerged, for example, if the loss of individuals was evenly spread across the clusters in one arm of the study but was all from the same cluster in the other arm.

It is worth noting that although the recommendations arising from this review are reported independently, on many occasions, they are co-dependent. One of the quality recommendations—having at least four clusters per study arm—was not satisfied in one study due to the way randomization was conducted. Eklund et al. (2014) reported using drawing of lots—a randomization procedure with low risk of bias—after combining the eight clusters into two groups and randomizing the groups. Although the authors acknowledged this was not “strict randomisation,” in this review, the more significant consequence was on the effective sample size, which became only two clusters (one per study arm; p. 274). Our exploration of potential moderators suggests that statistician involvement is a stronger indicator of good study conduct and reporting than whether or not the journal endorsed the CONSORT statement. This supports the need for ongoing statistician involvement to ensure the quality of cluster RCTs.

Strengths and Limitations

We used rigorous searching and data extraction procedures to conduct a comprehensive evaluation of cluster RCTs evaluating occupational therapy interventions. However, trials not reported as using a cluster RCT design or as occupational therapy may have been missed.

Assessment of quality was based solely on information provided in published articles, raising the possibility that omissions are an artifact of restrictions in journal space. This may become less of a concern with increasing use of open-access, online journals. Although open-access journals are still limited in the field of occupational therapy, many of the articles included were not from occupational therapy-specific journals. Finally, the small sample size prevented statistical analysis of potential moderators.

Conclusion

Quality of cluster RCTs of occupational therapy interventions is comparable with those from other areas of health

research and needs improvement. Particular focus should be on identifying and justifying the use of a cluster RCT design and accounting for the clustering in the sample size and analysis. Where possible, involving a statistician is likely to improve the quality and reporting of such trials. Increased reporting of ICCs would improve the credibility of published research and aid researchers in estimating appropriate ICCs for future trials. It is also important for researchers to report a comparison of clusters at baseline and provide more detailed information regarding loss to follow-up from within clusters. To enable systematic reviews and meta-analysis of cluster RCTs evaluating occupational therapy interventions in the future, more detailed reporting is required and more interventions must be identified as sitting within the scope of occupational therapy.

Acknowledgment

The authors thank Margie Olds for her independent review of the authors' protocol.

Authors' Note

The findings of this review have been presented at the Asia Pacific Occupational Therapy Conference, September 2015, Rotorua, New Zealand.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The open-access submission of this review was supported by the Centre for Person-Centred Research at Auckland University of Technology (AUT), New Zealand.

References

- Barbui, C., & Cipriani, A. (2011). Cluster randomised trials. *Epidemiology and Psychiatric Sciences, 20*, 307-309. doi:10.1017/s2045796011000515
- Barzel, A., Ketels, G., Tetzlaff, B., Kruger, H., Haevernich, K., Daubmann, A., . . . Scherer, M. (2013). Enhancing activities of daily living of chronic stroke patients in primary health care by modified constraint-induced movement therapy (HOMECIMT): Study protocol for a cluster randomized controlled trial. *Trials, 14*(1), Article 334. doi:10.1186/1745-6215-14-334
- Bennett, S., Hoffmann, T., McCluskey, A., Coghlan, N., & Tooth, L. (2013). Systematic reviews informing occupational therapy. *American Journal of Occupational Therapy, 67*, 345-354. doi:10.5014/ajot.2013.005819
- Bennett, S., McKenna, K., McCluskey, A., Tooth, L., Hoffmann, T., & Strong, J. (2007). Evidence for occupational therapy interventions: Effectiveness research indexed in the OTseeker database. *British Journal of Occupational Therapy, 70*, 426-430.
- Bundy, A. C., Naughton, G., Tranter, P., Wyver, S., Baur, L., Schiller, W., . . . Brentnall, J. (2011). The Sydney playground project: Popping the bubblewrap—Unleashing the power of play: A cluster randomized controlled trial of a primary school playground-based intervention aiming to increase children's physical activity and social skills. *BMC Public Health, 11*, Article 680. doi:10.1186/1471-2458-11-680
- Campbell, M., Elbourne, D. R., & Altman, D. G. (2004). CONSORT statement: Extension to cluster randomized trials. *British Medical Journal, 328*, 702-708. doi:10.1136/bmj.e5661
- Campbell, M., Thomson, S., Ramsay, C. R., MacLennan, G. S., & Grimshaw, J. M. (2004). Sample size calculator for cluster randomized trials. *Computers in Biology and Medicine, 34*, 113-125. doi:10.1016/s0010-4825(03)00039-8
- Creek, J. (2003). *Occupational therapy defined as a complex intervention*. London, England: College of Occupational Therapists.
- Diaz-Ordaz, K., Froud, R., Sheehan, B., & Eldridge, S. (2013). A systematic review of cluster randomised trials in residential facilities for older people suggests how to improve quality. *BMC Medical Research Methodology, 13*(127). doi:10.1186/1471-2288-13-127
- Döpp, C. M., Graff, M. J., Teerenstra, S., Adang, E., Nijhuis-van der Sanden, R. W., Olderikkert, M. G., & Vernooij-Dassen, M. J. (2011). A new combined strategy to implement a community occupational therapy intervention: Designing a cluster randomized controlled trial. *BMC Geriatrics, 11*, Article 13. doi:10.1186/1471-2318-11-13
- Döpp, C. M., Graff, M. J., Teerenstra, S., Olde Rikkert, M. G., Nijhuis-van der Sanden, M. W., & Vernooij-Dassen, M. J. (2014). Effectiveness of a training package for implementing a community-based occupational therapy program in dementia: A cluster randomized controlled trial. *Clinical Rehabilitation, 29*, 974-986. doi:10.1177/0269215514564699
- Eklund, M., Gunnarsson, B. A., Sandlund, M., & Leufstadius, C. (2014). Effectiveness of an intervention to improve day centre services for people with psychiatric disabilities. *Australian Occupational Therapy Journal, 61*, 268-275. doi:10.1111/1440-1630.12120
- Eldridge, S., Ashby, D., Feder, G. S., Rudnicka, A. R., & Ukoumunne, O. C. (2004). Lessons for cluster randomized trials in the twenty-first century: A systematic review of trials in primary care. *Clinical Trials, 1*(1), 80-90. doi:10.1191/1740774504cn006rr
- Eldridge, S., Ashby, D., & Kerry, S. (2006). Sample size for cluster randomized trials: Effect of coefficient of variation of cluster size and analysis method. *International Journal of Epidemiology, 35*, 1292-1300. doi:10.1093/ije/dy1129
- Eldridge, S., & Kerry, S. (2012). *A practical guide to cluster randomised trials in health services research: Statistics in practice*. Cornwall, UK: John Wiley.
- Engelen, L., Bundy, A. C., Naughton, G., Simpson, J. M., Bauman, A., Ragen, J., . . . van der Ploeg, H. P. (2013). Increasing physical activity in young primary school children—It's child's play: A cluster randomised controlled trial. *Preventive Medicine, 56*, 319-325. doi:10.1016/j.ypmed.2013.02.007
- Eyssen, I. C., Steultjens, M. P., de Groot, V., Steultjens, E. M., Knol, D. L., Polman, C. H., & Dekker, J. (2013). A cluster randomised controlled trial on the efficacy of client-centred occupational therapy in multiple sclerosis: Good process, poor outcome. *Disability and Rehabilitation, 35*, 1636-1646. doi:10.3109/09638288.2012.748845
- Higgins, J.P., Altman, D.G., Gotzsche, P.C., Juni, P., Moher, D., Oxman, A.D., . . . Cochrane Statistical Methods Group. (2011).

- The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *British Medical Journal*, 343, Article d5928. doi:10.1136/bmj.d5928
- IBM. (2011). *IBM SPSS missing values 20*. Retrieved from http://public.dhe.ibm.com/software/analytics/spss/documentation/statistics/20.0/en/client/Manuals/IBM_SPSS_Missing_Values.pdf
- Isaakidis, P., & Ioannidis, J. P. A. (2003). Evaluation of cluster randomized controlled trials in Sub-Saharan Africa. *American Journal of Epidemiology*, 158, 921-926. doi:10.1093/aje/kwg232
- Ivers, N. M., Taljaard, M., Dixon, S., Bennett, C., McRae, A., Taleban, J., . . . Donner, A. (2011). Impact of CONSORT extension for cluster randomised trials on quality of reporting and study methodology: Review of random sample of 300 trials, 2000-8. *British Medical Journal*, 343, Article d5886. doi:10.1136/bmj.d5886
- Killaspy, H., Cook, S., Mundy, T., Craig, T., Holloway, F., Leavey, G., . . . King, M. (2013). Study protocol: Cluster randomised controlled trial to assess the clinical and cost effectiveness of a staff training intervention in inpatient mental health rehabilitation units in increasing service users' engagement in activities. *BMC Psychiatry*, 13(216), 1-8. doi:10.1186/1471-244X-13-216
- Kim, S. Y., Yoo, E. Y., Jung, M. Y., Park, S. H., & Park, J. H. (2012). A systematic review of the effects of occupational therapy for persons with dementia: A meta-analysis of randomized controlled trials. *NeuroRehabilitation*, 31(2), 107-115. doi:10.3233/NRE-2012-0779
- Law, M., Darrach, J., Pollock, N., Rosenbaum, P., Russell, D. J., Walter, S. D., . . . Wright, V. (2007). Focus on function—A randomized controlled trial comparing two rehabilitation interventions for young children with cerebral palsy. *BMC Pediatrics*, 7(31), 1-12. doi:10.1186/1471-2431-7-31
- Law, M., Darrach, J., Pollock, N., Wilson, B., Russell, D. J., Walter, S. D., . . . Galuppi, B. (2011). Focus on function: A cluster, randomized controlled trial comparing child- versus context-focused intervention for young children with cerebral palsy. *Developmental Medicine & Child Neurology*, 53, 621-629. doi:10.1111/j.1469-8749.2011.03962.x
- McCluskey, A., Ada, L., Middleton, S., Kelly, P. J., Goodall, S., Grimshaw, J. M., . . . Karageorge, A. (2013). Improving quality of life by increasing outings after stroke: Study protocol for the out-and-about trial. *International Journal of Stroke*, 8, 54-58. doi:10.1111/j.1747-4949.2012.00966.x
- Medical Research Council. (2008). *Developing and evaluating complex interventions: New guidance*. London, England: Author. Retrieved from www.mrc.ac.uk/complexintervention-guidance
- Moberg-Mogren, E., & Nelson, D. L. (2006). Evaluating the quality of reporting occupational therapy randomized controlled trials by expanding the CONSORT criteria. *American Journal of Occupational Therapy*, 60, 226-235. doi:10.5014/ajot.60.2.226
- Mozley, C. G., Schneider, J., Cordingley, L., Molineux, M., Duggan, S., Hart, C., . . . Cruickshank, A. (2007). The care home activity project: Does introducing an occupational therapy programme reduce depression in care homes? *Aging & Mental Health*, 11, 99-107. doi:10.1080/13607860600637810
- Norton-Mabus, J. C., & Nelson, D. L. (2008). Reporting of randomized controlled trials in occupational therapy and speech therapy: Evaluation using an expansion of the CONSORT statement. *OTJR: Occupation, Participation and Health*, 28, 64-71.
- Sackley, C., Burton, C. R., Herron-Marx, S., Lett, K., Mant, J., Roalfe, A. K., . . . Feltham, M. G. (2012). A cluster randomised controlled trial of an occupational therapy intervention for residents with stroke living in UK care homes (OTCH): Study protocol. *BMC Neurology*, 12, Article 52. doi:10.1186/1471-2377-12-52
- Sackley, C., Wade, D. T., Mant, D., Atkinson, J. C., Yudkin, P., Cardoso, K., . . . Reel, K. (2006). Cluster randomized pilot controlled trial of an occupational therapy intervention for residents with stroke in UK care homes. *Stroke*, 37, 2336-2341. doi:10.1161/01.STR.0000237124.20596.92
- Sackley, C., Walker, M. F., Burton, C. R., Watkins, C. L., Mant, J., Roalfe, A. K., . . . OTCH Trial Investigators. (2015). An occupational therapy intervention for residents with stroke related disabilities in UK care homes (OTCH): Cluster randomised controlled trial. *British Medical Journal*, 350, Article h468. doi:10.1136/bmj.h468
- Sutton, C. J., Watkins, C. L., & Dey, P. (2013). Illustrating problems faced by stroke researchers: A review of cluster-randomized controlled trials. *International Journal of Stroke*, 8, 566-574. doi:10.1111/j.1747-4949.2012.00915.x
- Taylor, W. J., Brown, M., William, L., McPherson, K. M., Reed, K., Dean, S. G., & Weatherall, M. (2012). A pilot cluster randomized controlled trial of structured goal-setting following stroke. *Clinical Rehabilitation*, 26, 327-338. doi:10.1177/0269215511419384
- Teerenstra, S., Eldridge, S., Graff, M., de Hoop, E., & Borm, G. F. (2012). A simple sample size formula for analysis of covariance in cluster randomized trials. *Statistics in Medicine*, 31, 2169-2178. doi:10.1002/sim.5352
- Tokolahi, E., Hocking, C., Kersten, P., & Vandal, A. C. (2014). Cluster-randomised controlled trial of an occupational therapy group intervention for children designed to promote emotional wellbeing: Study protocol. *BMC Psychology*, 2(1), Article 16. doi:10.1186/2050-7283-2-16
- Wenborn, J., Challis, D., Head, J., Miranda-Castillo, C., Popham, C., Thakur, R., . . . Orrell, M. (2013). Providing activity for people with dementia in care homes: A cluster randomised controlled trial. *International Journal of Geriatric Psychiatry*, 28, 1296-1304. doi:10.1002/gps.3960