The randomized controlled trial (RCT) has become the most widely accepted research design for testing the efficacy of health care interventions. The purposes of this paper are to (a) introduce the essential elements of RCTs, and (b) discuss some of the special problems faced by occupational therapists in conducting and interpreting RCTs. The CONSORT (Consolidated Standards of Reporting Trials) statement is recommended as an introduction to the components of RCT quality. Problems pertinent to the study of the efficacy of occupational therapy and related interventions include the importance of: theory, background, and rationale; treatment fidelity; theory-based outcomes; management of non-masked (non-blinded) interventionists and participants; and multiplicity of statistical analyses. This paper can help practitioners judge the quality of an RCT, and it can help the profession work toward the development of a cadre of qualified researchers who can adapt the well-established methods of RCTs to the study of occupation-based interventions.

The goal of a randomized controlled trial (also called randomized clinical trial and RCT) is to evaluate an intervention in terms of its health benefit and safety for a specific population (Meinart, 1986). An RCT may address either of the following questions: (a) Is an intervention efficacious and safe (compared to no intervention)? or (b) Is an intervention relatively efficacious and safe (compared to an alternative intervention) (Friedman, Furberg, & DeMets, 1998, p. 2)? Currently, RCTs are considered the gold standard for testing the efficacy of medical interventions, because extensive research indicates that RCTs provide the best possible quantitative evidence of efficacy and effectiveness (Altman et al., 2002). Each year scientists across the world conduct thousands of RCTs (Olkin, 1995). The methods for RCTs were originally developed primarily for the evaluation of medical interventions, particularly drug therapies (Hill, 1962; Spilker, 1991).

The populations studied in RCTs have most commonly been defined in terms of specific diagnostic categories, and the health benefits investigated have been physiological. Recently, however, RCTs have been used to address other kinds of health issues with direct implications for occupational therapy, such as falls risk in older persons (Cumming et al., 1999); fear of falling (Tennstedt et al., 1998); functional implications of assistive technology (Mann, Ottenbacher, Fraas, Tomit, & Granger, 1999); and strategies used by caregivers of persons with dementia (Gitlin, Corcoran, Winter, Boyce, & Hauck, 2001).

The purposes of this paper are to (a) introduce the main features of RCTs, and (b) discuss some of the special problems faced by occupational therapists in conducting and interpreting RCTs. In the 2000 Eleanor Clarke Slagle lecture, Holm (2000) argued that some research methodologies are inherently superior to others in their capacities to provide evidence for practice, and she stated that the best...
support for evidenced-based occupational therapy practice will consist of systematic reviews of multiple, well-designed RCTs. Holm challenged the profession to develop practitioners who know how to appraise evidence based on research. To accomplish this goal, the profession must foster increased knowledge of RCTs, including the guidelines to be used in judging quality. In addition, the profession must develop a cadre of qualified researchers who can adapt the well-established methods of RCTs to the study of occupation-based interventions. As Holm pointed out, the profession of occupational therapy has only begun to evaluate its many interventions.

**Essential Elements of RCTs**

An RCT is defined as “a prospective study comparing the effect and value of intervention(s) against a control in human beings” (Friedman et al., 1998, p. 2). Key elements include random assignment of participants to interventions; systematic controls against bias; and outcomes reflecting clinically meaningful differences.

Several texts are recommended to the serious scholar desiring to understand the many complexities of RCTs (Chow & Liu, 1998; Everitt & Pickles, 1999; Friedman et al., 1998; Meinart, 1986; Spilker, 1991). In addition, the Cochrane Library (2002) is an excellent resource providing regularly updated databases for RCT methods and results. Recently, McKenna, Bennett, Hoffmann, McCluskey, Strong, and Tooth (2003) have constructed a Web-based database (OTseeker.com) containing abstracts of systematic reviews and RCTs relevant to occupational therapy. To provide the reader of this article with a brief yet authoritative introduction to RCTs, we have chosen the CONSORT (Consolidated Standards of Reporting Trials) statement (Altman et al., 2002). The purpose of the CONSORT statement is to improve the reporting of RCTs. In the mid-1990s, an international group of clinical researchers, statisticians, epidemiologists, and biomedically oriented journal editors developed and published the first version of CONSORT (Begg et al., 1996). CONSORT has been supported by a growing number of journals and editorial groups. In 2001, Moher et al. published a revision of CONSORT, based on methodological research concerning the original version. CONSORT provides not only a way to evaluate the reporting of RCTs but also a guide in the planning of RCTs.

Figure 1 depicts the four main phases of an RCT: enrollment; intervention allocation; follow-up; and analysis (Altman et al., 2002; Moher, Schulz, & Altman, 2001).

![CONSORT Diagram Showing the Flow of Participants Through Each Stage of a Randomized Controlled Trial](http://ajot.aota.org/pdfaccess.ashx?url=/data/journals/ajot/930153/ on 04/06/2018 Terms of Use: http://AOTA.org/terms)
First, potential research subjects are assessed for eligibility and willingness to participate. At the conclusion of enrollment, each participating subject is randomly assigned to one of the interventions. Pictured in Figure 1 is a two-arm trial involving the comparison of two interventions; additional arms are possible. For example, subjects could be assigned to one of three interventions: a novel intervention, a conventional intervention, or a control intervention. Next, subjects experience the interventions in carefully planned and controlled ways. Primary and secondary (less important) outcomes are measured at planned points during the trial. Primary outcome measures should reflect clinically meaningful variables. The statistical analysis is planned in advance of the trial, and determines whether the outcomes vary depending on which intervention was received. Often, the planned analysis involves adjustments reflecting the characteristics of the subjects prior to receiving the intervention (baseline characteristics). The logic of an RCT is as follows: if those allocated to the various interventions do not differ from each other prior to the intervention (as ensured by randomization and examination of baseline characteristics), if all subjects are treated in the same way except for the interventions, and if they differ in some meaningful way after the intervention (as estimated by the statistical analysis), the only logical conclusion is the probable superiority of one of the interventions.

Figure 1 reflects the importance of monitoring the numbers of persons who are ineligible for the study, who do not receive the intended interventions, who are lost to follow-up, who discontinue interventions, and who are excluded from the analysis. The reader of an RCT needs to know these numbers to make a decision as to the study’s validity (Egger, Juni, & Bartlett, 2001). Internal validity (the fairness of the comparison) could be threatened if significantly more subjects are lost to follow-up in one intervention than in another. External validity (generalizability) could be threatened if an unrepresentative sample of participants complete the study.

Table 1 (Altman et al., 2002) summarizes CONSORT criteria for reporting RCTs. The reader is urged to consult the CONSORT Web site (http://www.consort-statement.org) for a glossary of terms and for a researched-based explanation of why each item is important. For example, CONSORT’s emphasis on the processes of randomization is justified by Schulz, Chalmers, Hayes, and Altman (1995), who found that bias can be introduced in the assignment of participants to interventions. The CONSORT recommendation of a flowchart diagram for describing the numbers of participants enrolled in the study and the numbers of participants lost to the study at specific points is also justified by research conducted by Schulz et al. (1995).

Special Problems of Research Design
Presented by Complex Interventions

RCTs were developed and perfected through thousands of drug trials. However, the investigation of occupational therapy efficacy typically involves complexities that are not present in drug trials. For example, in a drug trial a placebo can be used in such a way that research subjects as well as researchers do not know if a particular subject is receiving the active agent or the placebo. This is not possible in occupational therapy research, where both the practitioner and the research participant must be aware of what is happening to the participant. It is important to point out that other professions must deal with similar kinds of complexities. As with occupational therapy, the fields of nursing, general rehabilitation, social work, and psychotherapy often involve interventions characterized by ongoing interactions between practitioners and active participants. These fields are also interested in theory-guided interventions involving multiple behaviors (Nigg, Allegrante, & Ory, 2002). In the rest of this paper, we will focus on CONSORT items that pose special problems to RCTs conducted by occupational therapists and others interested in multifaceted interventions and complex outcomes.

Theory, Background, and Rationale

CONSORT Item 2 (see Table 1) presents a special challenge to research in occupational therapy and others fields in which the research base is not yet broadly established. Presentation of scientific background and rationale is especially complex if the research question has received little or no prior study and if the intervention involves a series of interactions between those receiving and administering interventions. Such is often the case when studying occupational therapy interventions. Furthermore, occupational therapy involves the active participation of those receiving therapy, and their choices affect every aspect of therapy (Moyers, 1999). In contrast, drug studies need not be concerned about subjects’ recommendations concerning chemical composition.

Nigg et al. (2002) argued that an explication of the theoretical basis of an intervention is necessary for a full understanding of why change may or may not occur. Theory ties a particular RCT to an extensive matrix of knowledge and potential knowledge. Inevitably, a successful, theory-based RCT will lead to a series of future studies, whereby the mechanisms of action underlying the intervention and the longer-term outcomes of the intervention will be examined. Over time, the theory will be adjusted based on new data, or alternative theories will be proposed that purport to explain the data more comprehensively and succinctly.
Science is theory, buttressed by controlled studies and appealing for controlled studies (Kerlinger, 1986, pp. 8–12).

Writing with the complexities of family-based interventions in mind, Gitlin et al. (2000) recommended several theories for consideration in understanding why and how interventions work. These include stress process models; motivational theories; behavior change theories; and personal control theory. Two theoretical frameworks relevant to occupational therapy research have been put forward by Bandura (1997) and Lawton (1980). Bandura suggested that therapists can have positive influences on clients’ and patients’ beliefs about their self-efficacy, which can in turn lead to healthy patterns of living. This theory may be useful in examining teaching-learning processes in occupational therapy (e.g., Thomas, 1993). Lawton proposed that successful aging be conceptualized as a dynamic relationship between potentially revised environments and the changing abilities of the person. This theory may be useful in studying environmental modifications and assistive technology for aging persons.

We believe that occupational therapy researchers have special responsibilities in terms of theory, particularly theory specific to occupational therapy. The reader of an RCT investigating occupational therapy needs to know the connection between the intervention and the theoretical foundation.

---

**Table 1. CONSORT Essential Items in Reporting Randomized Controlled Trials (Altman et al., 2001)**

<table>
<thead>
<tr>
<th>PAPER SECTION</th>
<th>Item</th>
<th>Description</th>
<th>Reported page #</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>TITLE and ABSTRACT</strong></td>
<td>1</td>
<td>How participants were allocated to interventions (e.g., “random allocation,” “randomized,” or “randomly assigned”).</td>
<td></td>
</tr>
<tr>
<td><strong>INTRODUCTION</strong></td>
<td>2</td>
<td>Scientific background and explanation of rationale.</td>
<td></td>
</tr>
<tr>
<td><strong>METHOD</strong></td>
<td>3</td>
<td>Eligibility criteria for participants and the settings and locations where the data were collected.</td>
<td></td>
</tr>
<tr>
<td>Participants</td>
<td>4</td>
<td>Precise details of the interventions intended for each group and how and when they were actually administered.</td>
<td></td>
</tr>
<tr>
<td>Interventions</td>
<td>5</td>
<td>Specific objectives and hypotheses.</td>
<td></td>
</tr>
<tr>
<td>Objectives</td>
<td>6</td>
<td>Clearly defined primary and secondary outcome measures and, when applicable, any methods used to enhance the quality of measurements (e.g., multiple observations, training of assessors).</td>
<td></td>
</tr>
<tr>
<td>Outcomes</td>
<td>7</td>
<td>How sample size was determined and, when applicable, explanation of any interim analyses and stopping rules.</td>
<td></td>
</tr>
<tr>
<td>Sample size</td>
<td>8</td>
<td>Method used to generate the random allocation sequence, including details of any restriction (e.g., blocking, stratification).</td>
<td></td>
</tr>
<tr>
<td>Randomization: Sequence generation</td>
<td>9</td>
<td>Method used to implement the random allocation sequence (e.g., numbered containers or central telephone), clarifying whether the sequence was concealed until interventions were assigned.</td>
<td></td>
</tr>
<tr>
<td>Randomization: Allocation Concealment</td>
<td>10</td>
<td>Who generated the allocation sequence, who enrolled participants, and who assigned participants to their groups.</td>
<td></td>
</tr>
<tr>
<td>Randomization: Implementation</td>
<td>11</td>
<td>Whether or not participants, those administering the interventions, and those assessing the outcomes were blinded to group assignment. When relevant, how the success of blinding was evaluated.</td>
<td></td>
</tr>
<tr>
<td>Blinding (masking)</td>
<td>12</td>
<td>Statistical methods used to compare groups for primary outcome(s); Methods for additional analyses, such as subgroup analyses and adjusted analyses.</td>
<td></td>
</tr>
<tr>
<td>Statistical methods</td>
<td>13</td>
<td>Flow of participants through each stage (a diagram is strongly recommended). Specifically, for each group report the numbers of participants randomly assigned, receiving intended treatment, completing the study protocol, and analyzed for the primary outcome. Describe protocol deviations from study as planned, together with reasons.</td>
<td></td>
</tr>
<tr>
<td>Participant flow</td>
<td>14</td>
<td>Dates defining the periods of recruitment and follow-up.</td>
<td></td>
</tr>
<tr>
<td>Recruitment</td>
<td>15</td>
<td>Baseline demographic and clinical characteristics of each group.</td>
<td></td>
</tr>
<tr>
<td>Baseline data</td>
<td>16</td>
<td>Number of participants (denominator) in each group included in each analysis and whether the analysis was by “intention-to-treat.” State the results in absolute numbers when feasible (e.g., 10/20, not 50%).</td>
<td></td>
</tr>
<tr>
<td>Numbers analyzed</td>
<td>17</td>
<td>For each primary and secondary outcome, a summary of results for each group, and the estimated effect size and its precision (e.g., 95% confidence interval).</td>
<td></td>
</tr>
<tr>
<td>Outcomes and Estimation</td>
<td>18</td>
<td>Address multiplicity by reporting any other analyses performed, including subgroup analyses and adjusted analyses, indicating those prespecified and those exploratory.</td>
<td></td>
</tr>
<tr>
<td>Ancillary analyses</td>
<td>19</td>
<td>All important adverse events or side effects in each intervention group.</td>
<td></td>
</tr>
<tr>
<td>Adverse events</td>
<td>20</td>
<td>Interpretation of the results, taking into account study hypotheses, sources of potential bias or imprecision, and the dangers associated with multiplicity of analyses and outcomes.</td>
<td></td>
</tr>
<tr>
<td>Interpretation</td>
<td>21</td>
<td>Generalizability (external validity) of the trial findings.</td>
<td></td>
</tr>
<tr>
<td>Generalizability</td>
<td>22</td>
<td>General interpretation of the results in the context of current evidence.</td>
<td></td>
</tr>
</tbody>
</table>

---

Note: this column refers to the page number in the text of the RCT under review by CONSORT criteria.
dations of the profession of occupational therapy. Why should the intervention be called occupational therapy? The professional literature of occupational therapy provides many options. These include references to documents of the American Occupational Therapy Association, such as The Guide to Occupational Therapy Practice (Moyers, 1999), or specific theorists, such as Kielhofner (1995) or Nelson (1996). The ideal occupational therapy RCT would test a well-articulated model of practice solidly based in the use of therapeutic occupation, with interdisciplinary theoretical contributions.

**Interventions and Treatment Fidelity**

Item 4 of CONSORT calls for describing precise details of the interventions planned and actually carried out. An occupational therapy intervention usually involves a therapeutic relationship that is developed over time and that consists of many interactions. It is challenging yet essential to document that the intervention is implemented as intended. If the intervention is not implemented as planned, the results of the study can be biased. In one scenario, a truly effective intervention could be found to be not effective because it was not implemented as intended. In another scenario, a truly ineffective intervention could be found to be effective, because an alternative effective intervention was implemented in its place. Either scenario has negative consequences for researchers and clinicians who use this information. For these reasons, measures of treatment fidelity or implementation, which assess the degree to which the manipulation of the independent variable occurred as planned, have become essential components of RCTs (Moncher & Prinz, 1991).

In the critique of psychotherapy trials, Lichstein, Kiel, and Grieve (1994) identified a number of threats to treatment fidelity, strategies to enhance treatment fidelity, and methods of measuring treatment fidelity. In addition, they suggested that three components of treatment fidelity—delivery, receipt, and enactment—should be evaluated. Each of the three components will be discussed in more detail below. An example of these concepts used in an occupational therapy RCT is presented in Table 2.

**Treatment Delivery.** Measures of treatment delivery assess the degree to which therapists deliver an intervention to participants as intended. Examples of failure to document treatment delivery include vaguely defined occupational therapy interventions, the lack of a clear, detailed protocol for therapists to follow, inadequate training of therapists, unwillingness or inability of therapists to follow a protocol, and low levels of motivation by therapists involved in the study. In these situations, occupational therapists may omit aspects of the intended intervention or they may add unplanned intervention components to supplement the protocol. These alterations to the intervention protocol could deplete or enhance treatment efficacy. In either case, it becomes impossible to determine which specific aspect of the intervention is responsible for the obtained results (Lichstein et al., 1994).

There are several strategies that can enhance documentation of treatment delivery. A clear, detailed protocol or manual describing the occupational therapy intervention is extremely important. It should include not only the essential components of the intervention but also components that are not part of the intervention and may be part of the competing intervention. The intended intervention needs to be straightforward enough that a trained therapist can implement it on a consistent basis. Therapists’ training should include opportunities to observe or to practice or both the intervention prior to the study. The researcher should select therapists who are motivated and committed to following the intervention protocol, and also should make sure that there are no disincentives for therapists’ participation in the study. When possible, incentives for therapists’ participation should be offered.

<table>
<thead>
<tr>
<th>Fidelity Components</th>
<th>Enhancement Strategies</th>
<th>Measures</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Delivery</strong></td>
<td>Instructor training</td>
<td>Hours of instructor training</td>
</tr>
<tr>
<td></td>
<td>Use of published protocol</td>
<td>Assessment of course delivery</td>
</tr>
<tr>
<td></td>
<td>(Packer, Brink, &amp; Sauriol, 1995)</td>
<td>Videotapes of two groups</td>
</tr>
<tr>
<td><strong>Receipt</strong></td>
<td>Use of handouts</td>
<td>Quizzes on knowledge gains</td>
</tr>
<tr>
<td></td>
<td>Practice opportunities</td>
<td>Videotapes of two groups</td>
</tr>
<tr>
<td></td>
<td>Problem-solving discussions</td>
<td>Self-report of homework performance</td>
</tr>
<tr>
<td><strong>Enactment</strong></td>
<td>Practice opportunities</td>
<td>Energy Conservation Strategies Survey:</td>
</tr>
<tr>
<td></td>
<td>Homework assignments</td>
<td>– Type &amp; number of strategies used</td>
</tr>
<tr>
<td></td>
<td>Problem-solving discussions regarding what works &amp; what doesn’t work</td>
<td>– Effectiveness ratings</td>
</tr>
<tr>
<td></td>
<td>Verbal encouragement to perform homework</td>
<td>– Reasons for non-use</td>
</tr>
</tbody>
</table>

Various methods of measuring treatment delivery have been proposed. The method chosen should be tailored to the type of occupational therapy intervention and its delivery method. Specific assessments of treatment delivery can be developed to evaluate the degree that therapists adhere to the intervention protocol. Behaviors that are considered essential to the intervention need to be identified and rated by a neutral observer. For a group educational intervention, the observer might rate adherence to these items: “Therapist encouraged participant involvement in the session” or “Therapist summarized the six main conclusions to the session.” Participant report of treatment delivery may also be used. However, both instructor and participant reports are vulnerable to errors due to recall difficulty (Lichstein et al., 1994).

Videotaping or audiotaping an occupational therapy intervention is another way to document whether the intervention was delivered as intended. This strategy enables the investigator to analyze treatment delivery at a later time, more than one time, and by more than one person. For example, videotapes of two competing intervention approaches could be analyzed to determine whether a reviewer masked to the intervention approaches could differentiate between the two approaches. If the reviewer can identify accurately which intervention approach is being used, it indicates that these are unique approaches. If the reviewer is unable to identify differences, then there are problems in treatment delivery or the two approaches are not sufficiently different to be considered competing approaches. In some situations, the presence of a video camera, audiotape player, an observer, or all three, might affect the behavior of the participants and therapists, and affect the efficacy of the intervention. The best method of measuring treatment delivery is a pilot study.

Treatment Receipt. This aspect of treatment fidelity deals with the extent to which participants perceive, comprehend, and use the intervention as intended (i.e., participants understand the information presented and can demonstrate the desired behavior) (Lichstein et al., 1994). Failure in the area of treatment receipt may be due to shortcomings of therapists or participants, or from poor interaction—communication between the two. If a therapist delivers information using only one format (e.g., lecture) rather than multiple formats (e.g., handouts, audiovisual aids, demonstrations, discussions, practice opportunities, role-play, and homework assignments), there is high probability that many participants will not receive the intervention as intended and will not enact the behaviors that the investigator hopes to see. If participants have low motivation, inattentiveness, or cognitive deficits, they are unlikely to benefit from the intervention even if treatment delivery is optimal. In some cases, personality conflicts between a therapist and a participant may also impede treatment receipt.

The use of multiple formats described above to deliver the occupational therapy intervention often enhances treatment receipt. It is important to be responsive to the multiple ways that persons learn new information. Opportunities to discuss information help adult learners integrate new information with what they already know. Practice opportunities and role-play helps experiential learners to understand through doing (Turner, Clancy, McQuade, & Cardenas, 1990). Therapists’ summaries of the main points of an intervention increase participants’ memory of treatment content (Kazdin, 1979). A client-centered approach helps tailor the intervention to the unique needs and interests of each participant. In our increasingly diverse society, the occupational therapy intervention needs to be sensitive to culture and disability. Handouts may need to be in languages other than English for recent immigrants and may need to be enlarged for persons with visual impairments. Examples used need to be relevant to the environmental context of participants from diverse cultures and economic levels.

Measurement of treatment receipt is multifaceted. A common method is to use written quizzes to determine if participants understand the information received (e.g., Kivlahan, Marlatt, Fromme, Coppel, & Williams, 1990). Carey, Carey, and Meisler (1990) used an oral quiz (i.e., had participants describe the five problem-solving steps that they had learned in treatment). Some investigators evaluate participants’ understanding by asking them to role-play or demonstrate the desired behaviors (Turner et al., 1990). Videotapes or audiotapes can also be used to document treatment receipt as well as treatment delivery (i.e., some assessments may be used for more than one component of treatment fidelity).

Treatment Enactment. Ensuring treatment enactment is the third component of treatment fidelity. This component refers to the behavioral changes that a participant makes outside a therapy session as a result of an occupational therapy intervention. Threats to treatment enactment include not practicing the behavior, forgetting to do it, being unsure of the correct way to do it, experiencing a lack of success when doing it, lacking a suitable setting, and losing interest in the intervention (Lichstein, 1988). Other threats include (a) the participant’s belief that the intervention will not work, (b) the unavailability of assistance and cooperation of others, and (c) the participant’s habit of engaging in the desired behavior prior to the study.

Several strategies can enhance treatment enactment. One of the most common is providing opportunities for participants to practice the desired behaviors. Homework
assignments are an excellent way to foster practical applications of the concepts taught within their specific contexts. For example, Gitlin et al. (2001) asked caregivers of persons with dementia to practice environmental simplification and task analysis strategies; these practice sessions were followed up by discussion and additional problem-solving. For another example, Mathiowetz, Matuska, and Murphy (2001) asked persons with multiple sclerosis to practice energy conservation techniques and followed up the practice sessions with discussions of what did and did not work. When a strategy did not work, a problem-solving session ensued (as depicted in Table 2). Interventions spread out over several weeks are more conducive to strategies to enhance treatment enactment than brief interventions.

One method of measuring treatment enactment is through participants’ self-report. Using a questionnaire, participants can be asked how frequently they have performed the desired behaviors and asked to rate their effectiveness. If they have not performed the desired behaviors, they can be asked why. The latter information may be useful in refining the intervention over time. As an alternative to self-report, a trained reporter (e.g., a neutral observer or a household member) (Lichstein & Eakin, 1985) or a device (e.g., a pedometer recording amount of walking) can document treatment enactment. Lichstein and Hoelscher (1986) have argued that data provided by a trained observer is usually more accurate than self-report data. However, self-report may be all that is possible and is better than no assessment of treatment enactment.

When therapists read the report of an RCT, they should check for evidence that treatment fidelity was monitored. When treatment fidelity (i.e., treatment delivery, receipt, and enactment) has been documented, the reader can be confident that the hypothesized occupational therapy intervention in fact caused the outcomes. When planning an RCT, the researcher needs to include strategies for ensuring treatment fidelity. The outline provided in Table 2 is a useful guide to assist planning.

Table 3 illustrates these concepts. As an example, a 6-week, 2-hours-per-week, energy conservation course is conceptualized as the building blocks for achieving the long-term goals. Likewise, the researcher hypothesizes that the ultimate or long-term outcomes for a specific intervention might be. The assessments used to evaluate those outcomes are considered the proximal measurements. The researcher needs to think through how and why an intervention might achieve the ultimate outcome. What variables need to change before these variables can improve? First, participants in the course need to enact new behaviors beyond the physical presence of the interventionist (Burigo et al., 2001). However, occupational therapy interventions that do not depend on teaching-learning models or on remote participant follow-through might require different approaches to testing treatment fidelity. For example, an RCT testing the effects of modifying the living areas of long-term nursing residents would concentrate on delivery of environmental modifications that are individually meaningful and health-promoting; the residents would not have responsibility for receipt and enactment in the same way as in a teaching-learning intervention. Still, documentation of treatment (intervention) fidelity is essential in all occupational therapy RCTs, even though the specific components of the documentation depend on the particular nature of the study at hand.

Outcomes, Theory, and Mechanisms of Action

Item 6 of CONSORT specifies the need for clearly defined primary and secondary outcomes measures. Ideally, a study has a single, primary outcomes measure upon which the entire study is focused. However, it is often desirable to measure other relevant variables, particularly if they can shed light on the theory-based interrelationships among these variables, the intervention, and the primary outcome. For example, in an RCT investigating the effects of an occupational therapy home safety intervention on falls (the primary outcome) in older persons, it might be important to determine the relationships between fear of falling and the primary outcome.

When designing an RCT, it is important to conceptualize how and why an intervention works. This process is similar to the intervention planning process used by occupational therapists in practice. In collaboration with their clients, therapists identify the long-term and short-term goals for each client. In deciding on short-term goals, the therapist must consider what is needed sequentially for their client to achieve the long-term goals and what interventions will enable their client to get there (i.e., how and why will it happen?). Thus, achievement of the short-term goals becomes conceptually the building blocks for achieving the long-term goals. Likewise, the researcher hypothesizes what the ultimate or long-term outcomes for a specific intervention might be. The assessments used to evaluate those outcomes are considered the proximal measurements. The researcher needs to think through how and why an intervention might achieve the ultimate outcome. What variables need to change before the long-term goals are achieved? The assessments of these variables are considered proximal measurements because conceptually they need to change before the long-term goals can be achieved.

Table 3 illustrates these concepts. As an example, a 6-week, 2-hours-per-week, energy conservation course is hypothesized to decrease the impact of fatigue on the performance of everyday tasks and improve the quality of life for persons with multiple sclerosis. Hence, the table lists (a) impact of fatigue on everyday life and (b) quality of life as the distal variables. What needs to happen before these variables can improve? First, participants in the course need to implement the energy conservation strategies in their everyday lives. What variables need to change to make that happen? Self-efficacy theory (Bandura, 1982, 1997) predicts...
that participants must increase their knowledge of energy conservation strategies and practice the strategies. Knowledge and practice are predicted to lead to self-efficacy for performing the strategies; in turn, self-efficacy can be predicted to lead to implementation of energy conservation strategies in everyday life. Hence, knowledge of energy conservation strategies, practice of these strategies, and self-efficacy for performing these strategies are the proximal variables listed in Table 3. These proximal measurements explain how and why the energy conservation course is or is not effective in decreasing fatigue impact and improving the quality of life (distal measurements) of persons with multiple sclerosis. In addition, they provide support or lack of support for self-efficacy theory. Thus, the inclusion of proximal and distal measurements not only helps the researcher and the reader to understand what the outcomes are, but also helps them understand how and why the outcomes were or were not achieved. It should be noted that proximal measurements might also serve as measures of treatment fidelity as discussed earlier.

### Management of Non-Masked (Non-Blinded) Interventionists and Participants

Item 11 of CONSORT deals with masking (preferred here to “blinding”). Total masking is never possible if the occupational therapy intervention involves ongoing interaction between therapists and participants engaged in therapeutic occupations. Both the therapist and therapy recipient know that they are engaged in research and have the opportunity to show bias for or against the intervention. In addition, the therapy recipient knows something about the alternative intervention or control condition, because the process of informed consent involves a description of all conditions to which the person might be randomly assigned.

The occupational therapy researcher should employ various types of masking whenever possible and should compensate for a lack of masking by enhanced treatment fidelity, as described above. One important type of masking that is often possible involves the measurement of outcomes. It is highly desirable that a “wall” be erected between the persons conducting the outcomes assessments and the therapists administering the interventions. For example, it is often desirable to employ outsiders to conduct outcomes assessments, and they can be trained to have no interactions with the unit’s regular therapists (research assistants can serve as go-betweens for scheduling purposes). Another kind of desirable masking pertains to the special problem of interim analyses that can often be required by review boards for the protection of the rights of human subjects. A required interim analysis might lead to the earlier-than-planned cessation of an RCT if the hypothesis is clearly supported; if the alternative intervention or control condition is clearly superior; or if there is clearly no trend that will result in support of the superiority of either condition. [Note: These tests should be conducted at conservative levels of alpha to avoid false positives.] It is desirable if the person who makes decisions about the trial (usually the principal investigator) remains masked to the interim analyses. Decisions concerning stopping or prolonging the trial should not be made by research staff on the basis of interim results because of the possibility of bias, even though data and safety monitoring boards may stop the trial for ethical reasons. A final type of masking is essential for all trials and is not unique to occupational therapy research. Masking of the person who decides on the potential participant’s eligibility for inclusion in the study is a requirement for current RCTs. Items 8, 9, and 10 of CONSORT all deal with procedures to ensure that randomization to interventions has nothing to do with the possible biases of those who enroll the participants.

Given the impossibility of full masking in occupation-based RCTs, several strategies can be taken to mitigate limitations. Sometimes it is possible to randomly assign therapists (interventionists) to alternative interventions, and sometimes it is possible to keep therapists unaware of the alternative intervention or control. Random assignment of therapists is effective only in large studies in which many therapists are involved. If only a few therapists are involved, it might be possible that the ultimate difference between the interventions is due to the skill of the therapists, not to the inherent effectiveness of the intervention that ended with the better outcome scores. As for keeping therapists unaware of the alternative intervention, this can be difficult because of informal networks of communication even if therapists administer the interventions in scattered community settings, and it is unlikely to be accomplished in hospital...
settings where therapists can easily observe each other’s interventions (or interpretable signs of interventions).

Research participants’ perceptions can also lead to “bleeding” of interventions, whereby a patient envies what another patient is doing or what one’s roommate reports doing with interventionists. For ethical reasons, there are often limits on restricting information to research participants, and research participants always retain the right to discontinue participation in the research or to engage in other actions that they think are in their best interests.

When the occupational therapy researcher cannot ensure full masking, the solution is to enact procedures to help ensure intervention fidelity and outcomes validity. Each potential problem, as described above, should be analyzed in advance so that an action plan to address the problem can be drawn up. The procedures as actually implemented in the action plan should be measured; deviations from protocol and unexpected violations of internal validity must be reported so that readers of the report can make their own decisions concerning interpretation of the results (Item 20 of CONSORT).

Multiplicity of Statistical Analyses

Item 18 of CONSORT addresses the issue that multiple analyses of the same data create an increased risk of Type I error (false-positive findings). In elaborating and explaining CONSORT, Altman et al. (2002) stated that “Authors should especially resist the temptation to perform many subgroup analyses” (paragraph 3 of electronic link to Item 18). An example of a subgroup analysis would be a test of whether males and females respond differently in a falls prevention study. The special problem of occupational therapy research in complying with CONSORT on this issue relates once again to the fact that RCTs in our field are in their infancy. The complex interrelationships suggested by theory have seldom been tested before, so there is a great desire to test these relationships, especially theory-based relationships between proximal and distal outcomes across subgroups of interest. Indeed, it can be speculated that the greatest value of the first few RCTs in any area involves pointing the way toward future studies on increasingly carefully defined populations, interventions, and outcomes.

We recommend a balanced approach to this issue. The beginning researcher should focus most of his or her energies on a primary outcome, so that the overall sample size can be planned in advance in such a way as to avoid a Type II error (false-negative finding). Such an approach is consistent with CONSORT Items 7 and 18. Once the researcher has completed this first step, then secondary outcomes and subgroup analyses should be planned in depth. As Altman et al. stated, “Analyses that were prespecified in the trial protocol are much more reliable than those suggested by the data” (paragraph 3 of electronic link to Item 18). Alpha for the overall RCT should be partitioned, so that there is adequate power for testing the primary outcome. The researchers main interest in the subgroup analyses deals with effect sizes that can guide future trials, not with tests against alpha. Ottenbacher (1998) has reviewed various statistical procedures for dealing with multiplicity, and has recommended calculation of a percent error rate that lets the reader apply cautious interpretations without increasing the chances of a Type II error through a downward adjustment of alpha.

Conclusion

Tremendous advances have been made in the past decade in defining quality of RCTs. These advances have been made over the course of thousands of medical trials and increasing numbers of methodological studies. The CONSORT statement is a fruition of this past research. Methodological information remains uneven, however. Whereas several methodological studies have addressed the bias that occurs when those who enroll subjects can guess the randomization sequence (reviewed in Altman et al., 2002), no studies have been identified that deal with a typical problem of occupational therapy research: Should the same therapists administer both competing interventions, or should they specialize by intervention? We know that it would be better to keep therapists masked, but since that is impossible, what is the next best procedure? We call for methodological research that addresses the specific problems of RCTs in occupational therapy and related fields that do not have a well-developed research base and that deal with complex and prolonged interactions between therapists and participants.

The RCT has become the most widely accepted research design for testing the efficacy of health care interventions. However, other kinds of research design will continue to be important to answer other kinds of research questions and to pave the way for future RCTs. For example, randomization is impossible when the research question involves a comparison between two preexisting populations (e.g., a comparison among children with autism, children with mental retardation but not autism, and children with neither). For another example, the development of valid and reliable measures remains a priority for occupational therapy research and is a necessary prerequisite for measuring outcomes in RCTs.

This paper focused on RCTs issues that are especially problematic to occupational therapy research. These issues primarily reflect five CONSORT items dealing with (a) the importance of theory, background, and rationale; (b) treatment fidelity; (c) theory-based outcomes; (d) management

January/February 2004, Volume 58, Number 1
Acknowledgments

The authors thank Laura N. Gitlin, PhD, Thomas Jefferson University, and the Center for Outcomes Research and Education (CORE), based at the University of Illinois at Chicago and funded by the American Occupational Therapy Foundation and the American Occupational Therapy Association.

References


