CASE STUDIES, SINGLE-SUBJECT RESEARCH, AND N OF 1 RANDOMIZED TRIALS

Comparisons and Contrasts[1]

Catherine L. Backman,[2] MS, OT(C) and Susan R. Harris, PhD, PT

ABSTRACT Backman CL, Harris SR: Case studies, single-subject research, and N of 1 randomized trials: comparisons and contrasts. *Am J Phys Med Rehabil* 1999;78:170–176

Case studies, single-subject research designs, and N of 1 randomized clinical trials are methods of scientific inquiry applied to an individual or small group of individuals. A case study is a form of descriptive research that seeks to identify explanatory patterns for phenomena and generates hypotheses for future research. Single-subject research designs provide a quasi-experimental approach to investigating causal relationships between independent and dependent variables. They are characterized by repeated measures of an observable and clinically relevant target behavior throughout at least one pretreatment (baseline) and intervention phase. The N of 1 clinical trial is similar to the single-subject research design through its use of repeated measures over time but also borrows principles from the conduct of large, randomized controlled trials. Typically, the N of 1 trial compares a therapeutic procedure with placebo or compares two treatments by administering the two conditions in a predetermined random order. Neither the subject nor the clinician is aware of the treatment condition in any given period of time. All three approaches are relatively easy to integrate into clinical practice and are useful for documenting individualized outcomes and providing evidence in support of rehabilitation interventions.

KEY WORDS: Single-Subject Experiments, Case Reports, N of 1 Trials, Outcomes, Evidence-Based Practice

Single-subject research, case studies, and N of 1 randomized trials are all viable methods of scientific inquiry related to rehabilitation interventions. Compared with group experimental designs, all three are relatively easy to incorporate into usual clinical treatment, with the added benefit of providing a systematic method of documenting rehabilitation outcomes. Other sources provide detailed "how to" information on conducting single-subject research designs,[1,2] case studies,[3] and N of 1 trials.[4,5] The purposes of this article are to offer an overview of these methods, to introduce their major advantages and limitations, and to cite examples from the rehabilitation literature.

All three methods require that the investigator specify the problem to be studied. A clearly stated problem or research question guides the selection of the best method for investigation. A question that asks "what are the factors influencing compliance with recommended treatment?" is better suited to a case study than to single-subject research, because it is more exploratory in nature. Descriptive data gathered from interviews, observations, and other methods of eliciting information from an individual can generate possible explanations for enhancing compliance. On the other hand, a research question such as "what is the effect of a wrist splint on hand function in a woman with rheumatoid arthritis?" is well suited to single-subject research, because both the independent variable (a wrist splint) and the dependent variable (hand function) can be defined and a reliable method of measuring hand function can be established. In this case, the single-subject design will provide evidence for the effectiveness (or lack thereof) of the treatment. Not all clinically important problems lend themselves to investigation by one of the three methods described in this article. The following overview indicates some of the requirements for each method.

CASE STUDIES

Case studies are a type of descriptive research, involving a careful, preplanned observation of an

0894-9115/99/7802-0170$03.00/0
AMERICAN JOURNAL OF PHYSICAL MEDICINE & REHABILITATION

[1] From the Department of Health Care & Epidemiology (CLB), and the School of Rehabilitation Sciences, The University of British Columbia (CLB, SRH), Vancouver, British Columbia, Canada.

[2] All correspondence and requests for reprints should be addressed to: The University of British Columbia, School of Rehabilitation Sciences, T325–2211 Wesbrook Mall, Vancouver, British Columbia V6T 2B5, Canada.

individual or group.[6] A case study describes the subject in-depth and may include the use of observation, interviews, and other assessment techniques, resulting in both qualitative and quantitative data that can be used to explore and predict response to treatment or the natural history of a disease or condition.[3] Case studies are considered by some to be a specific nonexperimental, descriptive research design, whereas others use the term interchangeably with case reports, which is a description of clinical practice that does not involve research methodology. Step-by-step guidelines for writing and submitting a case report for publication have been produced by the American Physical Therapy Association.[3]

Typically, a case study consists of systematically recorded observations and narrative descriptions of an individual's characteristics and responses to treatment.[2,6] This elaborate documentation and description of behavior provides a basis for making plausible inferences and generating hypotheses for future research. A well-developed case study offers a rationale for the use of specific rehabilitation interventions or unique applications of treatment strategies and may offer suggestions for treatment modification or more efficient approaches. In keeping with its exploratory emphasis, the case study is most suitable for generating hypotheses, but it cannot test hypotheses. Because it neither manipulates nor attempts to control variables, a case study cannot demonstrate causal relationships between variables. Although a case study may include a report of objective findings, it does not provide very strong evidence in support of rehabilitation interventions because it lacks controlled comparisons (e.g., between treatment and no-treatment conditions), and its results cannot be generalized to others. Nevertheless, a case study may be the only realistic way of documenting responses to rehabilitation intervention for people with rare conditions or in unique clinical situations.

Greene and David[7] presented an example of enhancing the generalizability of case studies by repeating a structured case study design across multiple subjects at multiple sites. The main features of their approach were a conceptual framework that provides the structure to insure meaningful cross-site analysis, a sampling plan that ensures representativeness of the target population, procedures to guide each individual case study, and an analysis strategy that tests the limiting conditions imposed on the findings. Common to the analysis of both an individual case study and a multiple case study design is the identification of patterns as explanations or answers to the questions guiding the study. The investigator looks for relationships that make sense from among the observations and responses to interview questions. An explanatory pattern describes, for example, a sequence of actions or series of events, and the identification of the pattern allows the investigator to recognize similar patterns across events, subjects, or sites. Explaining this process is what lends credibility to the inferences made, and often the findings of a case study can be restated in the form of an empirically testable question.[7]

The case study approach was used to describe the effects of an aerobic exercise program for a 43-yr-old man with chronic multisystem impairments subsequent to a traumatic brain injury and multiple fractures 15 yr previously, heterotopic ossification of his hip, and a long history of smoking.[8] The authors reported the client's initial assessment findings and response to treatment during a 1-mo period. Reflecting on this case, they suggested that the prescribed exercise program produced an appropriate physiological response while accommodating the client's limitations in coordination and had sufficient appeal to enhance his motivation to continue exercising. These results offer hypotheses for testing in future research.

SINGLE-SUBJECT RESEARCH DESIGNS (SSRD)

Single-subject research designs provide a quasi-experimental approach to evaluating treatment effectiveness in a single subject or small group of subjects, in which subjects serve as their own controls.[1,2] There are several different types of SSRDs, but each of them includes a combination of at least a baseline phase and an intervention phase. They consist of systematic observation, measurement, graphing, and analysis of a carefully defined target behavior over time and are sometimes referred to as examples of time-series designs. The baseline phase consists of repeated measures of the dependent variable (target behavior or outcome) during a period of time before introducing the treatment or intervention plan. After the pattern of behavior has been established with several repeated measures (we suggest at least five data points), the intervention phase commences with introduction of the treatment. The target behavior continues to be measured repeatedly over time while the intervention is provided. Repeated measures help to improve the validity of the study, because a trend of behavior over several data points decreases the probability that the results were owing to chance.

Data are plotted on a graph and analyzed for changes in level, trend, and variability between the baseline and intervention phases (Fig. 1). Graphed data have the advantage of being easily understood by clients and clinicians alike, especially when the method of measuring the target behavior is clearly defined and clinically relevant. For example, if the client's goal is to improve mobility, say, propelling a manual wheelchair for 1 hr (long enough to go shopping), then the target behavior is minutes wheeled, and the progress is charted on the graph.

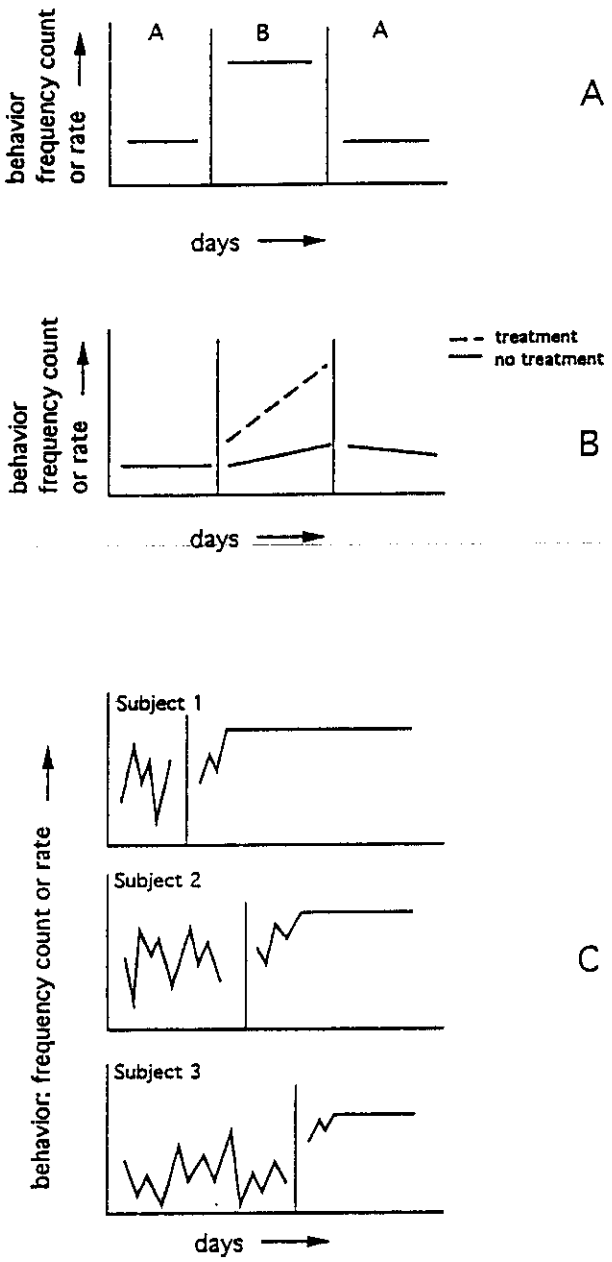

Figure 1. Examples of changes in level, trend, and variability in the visual analysis of single-subject research designs. *A*. Change in level in an ABA design. *B*. Change in trend in an alternating treatments design. C. Change in variability in a multiple baseline design across three subjects.

The introduction of a graded strengthening program could be the intervention, and minutes wheeled would continue to be plotted on the graph. The client and clinician can readily observe progress toward the goal and judge whether the intervention plan has an effect. However, inconsistency among raters relying primarily on visual analysis of graphed single-subject data has been demonstrated.[9]

Visual analysis of graphed data has been the foundation of the analysis of SSRDs, but recent studies supplement the analysis with semistatistical approaches to assist in interpreting the results and drawing conclusions. Examples of these techniques include the two standard deviation band method, the celeration line (or split-middle technique), and the C-statistic. Detailed explanations of these procedures can be found elsewhere,[1,2] but brief descriptions are provided here.

When considering any statistical approach, the underlying assumptions for the statistic must be considered. "Serial dependency" is an issue for any research design that uses repeated measures, because it violates the assumption of independence and renders statistics such as the *F* or *t* test inappropriate. Successive observations in a series tend to be correlated; that is, knowing the performance of a subject on one occasion often enables us to predict performance on a subsequent occasion. The extent to which this occurs can be assessed by examining autocorrelation or the correlation between data points separated by different time intervals within the series of observations.[10] If this correlation is significant, the data are considered to demonstrate serial dependency. The presence or absence of serial dependency dictates which statistics are appropriate for the analysis of single-subject data.

The two standard deviation band method assumes that data are normally distributed, so this procedure may not be appropriate for many SSRDs. Essentially, the standard deviation of the baseline data points is calculated, and a horizontal band the width of ±2 standard deviations is superimposed over the graphed data. If at least two consecutive intervention data points fall outside of this band, the intervention phase behavior is considered to be significantly different than the behavior exhibited during baseline (the likelihood of this event occurring by chance is less than 5 in 100).[2] The celeration line is a method of calculating a trend line for the baseline data, so 50% of the data points fall above the line and 50% fall below. Then, this line is extended through the intervention phase to provide a visual guide for predicting behavior over time. In effect, the celeration line tests the hypothesis that there is no difference in behavior across the two phases. If the intervention data points also fall such that 50% of them are above and 50% below the celeration line, then it may be concluded that there was no change in behavior resulting from introduction of the treatment. Finally, the C-statistic is a test designed for time series analysis; it can be used to evaluate a data set with as few as eight observations and is not affected by serial dependency. The C-statistic is a test that calculates the trend of the baseline data, then compares this with the trend of the baseline and intervention data combined. If the difference between the two calculations is significantly different, the behavior observed during the intervention phase is considered to be different than that observed during baseline.

Provided that there is no serial dependency of the

data points, conventional parametric statistics can be used to assist in data analysis, such as a *t* test to compare the mean value of the baseline phase with the mean value of the intervention phase. However, these procedures can also be problematic, because the mean performance level in a phase may not be the best representation of the target behavior.

Several books,[1,2] chapters,[11] and articles[12–14] exist to guide practitioners in selecting, evaluating, and conducting single-subject research. The most basic of SSRDs is the AB design, where A represents the baseline phase and B represents the intervention phase. Unless the target behavior changes dramatically between phases A and B, this design fails to provide convincing evidence of the cause-effect relationship between the intervention and the outcome. A withdrawal or ABA design helps to increase the confidence with which one can make a causal inference, especially if the target behavior reverts to levels initially observed in the first A phase. Yuen[15] used an ABAB design (an extension of the ABA design) to demonstrate the beneficial effect of an adaptation introduced to the workplace, which improved the task-specific productivity (making brackets) of a worker who was cortically blind. When the adaptation was withdrawn, productivity declined; when the adaptation was reinstated, productivity increased.

However, not all target behaviors are suited to a withdrawal design, because they may not be reversible. For example, rehabilitation programs aimed at increasing strength and endurance are rarely suited to a withdrawal design, because removal of the intervention is unlikely to immediately result in a loss of strength or endurance. A variation of the withdrawal design can be used to compare two different treatments, when the research question seeks to determine which of two treatments is more effective for the subject. This design, referred to as ABAC, consists of a baseline followed by Treatment B and a second baseline phase followed by Treatment C.

Another issue in rehabilitation research is controlling for the effects of natural recovery or healing or extraneous variables in the environment that are not easily controlled. The basic AB design may be appropriate, in that one would hypothesize that the rate of recovery should increase when intervention is introduced (thus, the slope toward improvement in the B phase should be a much steeper line than the slope established during baseline). A more effective way to control for extraneous variables is the multiple baseline design, in which the AB design is used across three or more subjects, each of whom is assigned a baseline of differing length. The analysis then compares trends across subjects as well as between the phases within each subject. This design has been used to investigate the effect of baclofen on spasticity in five subjects with spinal cord injury,[16] the effect of occupational therapy intervention on mealtime independence of four residents in a long-term care setting,[17] and the effects of caregiver promotion of self-care practice opportunities on the development of self-care skills in children with Down syndrome.[18]

Studying the effect of assistive devices or technology (such as orthoses, self-care aids, or computer adaptations) is particularly well suited to the use of an alternating treatments design. After the baseline phase, the treatment and no-treatment (or two different treatments) conditions are alternated, such that the target behavior is measured with and without the device. Therefore, two lines are plotted in the intervention phase and can be compared with each other as well as with the baseline data to assess the effect of the technology. Following the intervention phase with a second baseline phase helps to strengthen the design with respect to causal inference. The alternating treatments design has been used to investigate topics such as the effect of wrist splints on hand function in women with arthritis,[19] the use of tone-reducing ankle foot orthoses on the duration of standing balance in a 4-yr-old boy with cerebral palsy,[20] and the effect of body position on oxygen saturation in ventilated preterm infants.[21]

SSRDs provide a systematic approach to evaluating clinical change. They are ideally suited to rehabilitation research when individualized goals and treatment plans are important, because they provide objective evidence regarding the effect of treatment. However, they require clearly defined, observable target behavior that can be reliably measured, i.e., they are not suited to the evaluation of nondiscrete outcomes. SSRDs, as with other research designs, require attention to methodological requirements to ensure that causal inferences are made with confidence, and although there are many examples of SSRDs in the rehabilitation literature, there continue to be problems with adherence to basic rules that compromise the validity of findings.[12] Readers need to be aware also that the external validity of SSRDs is limited, because they are designed to measure outcomes that are functionally relevant to the individual being studied.

Critics of single-subject research have suggested that they lack generalizability. Although it is true that it would be inappropriate to generalize from a single study of a single subject, so too would it be inappropriate to generalize findings from a single-group comparison design using a nonrepresentative sample. Generalizability of group comparison designs depends on randomization, both in the selection of subjects and the assignment of subjects to treatment groups. Generalizability of single-subject research is established using replication. Hersen and Barlow[10] provide a compelling argument regarding the superiority of replication of single-subject studies compared with group designs to enhance generalizability of results. They discuss two types of replication procedures. Direct replication involves repeating the original study with a series of subjects possessing similar characteristics (e.g., same diag-

noses, problems, or impairments) and is analogous to the nonfactorial group design using a no-treatment control group. Direct replication allows the investigator to determine if the intervention that worked with one case works as well with other cases. Systematic replication involves exploring the effect of a previously successful experimental intervention (as indicated by direct replication) by using it in different settings, by different therapists, or with clients who have different characteristics than the original subjects. Systematic replication parallels a between-groups factorial design that incorporates different settings, therapists, or clients.

N OF 1 RANDOMIZED TRIALS

The N of 1 randomized, controlled trial (RCT) is, essentially, a single-subject research design. Although many of the guidelines for conducting SSRDs arise from research in the behavioral sciences, the N of 1 RCT has arisen from the medical literature and applies principles that are common to large RCTs to an individual instead. Most examples of N of 1 trials, then, tend to address the effectiveness of drugs, but they can certainly be applied to other types of discrete interventions that do not have a carryover effect. The N of 1 RCT can be used to help select optimal treatment for an individual by conducting a trial of two treatments, or of a treatment and a placebo, with that individual. To determine if an N of 1 RCT is appropriate, Sackett et al.[4] offer a list of questions to consider (Table 1). The response to each question should be "yes" before proceeding. These same authors provide easy-to-follow instructions for conducting and interpreting an N of 1 RCT.

As with SSRD, the N of 1 trial requires a specific target symptom or behavior that is both clinically relevant and can be measured or scored. For example, the subject's perception of pain on a 1 to 10 scale, the number of episodes of a specific symptom per day, or the distance walked without shortness of breath might be suitable outcomes for this type of trial. Typically, the subject would keep a daily or weekly diary of symptoms throughout the trial. The subject undergoes pairs of treatment periods, in which both the experimental treatment and placebo or alternative treatment are offered in turn (by random assignment). Pairs of treatment periods are continued until such time as the patient and clinician are satisfied that the two treatments are clearly different or clearly not different in effect (Fig. 2). The conduct of an N of 1 RCT investigating the effect of medications requires the assistance of a pharmacist to provide two identically appearing preparations in containers labeled for each phase of the trial. This reduces potential bias by ensuring that both clinician and patient are blinded as to the treatment (e.g., unaware of which is the active drug and which is the placebo).

The results of the N of 1 trial can be plotted on a graph and visually analyzed, although the paired *t* test (comparing the outcome measures from the two treatments) has been suggested as an appropriate statistical test.[4] The use of a paired *t* test has been criticized as being inappropriate, because the repeated measures may not be independent, which violates one of the assumptions underlying the statistic. However, Sackett and colleagues[4] analyzed 17 N of 1 trials and found no evidence of autocorrelation in any of them.

TABLE 1
Guidelines for N of 1 randomized trials

1. Is an N of 1 randomized trial indicated for this patient?
   a. Is the effectiveness of the treatment really in doubt?
   b. Will the treatment, if effective, be continued long term?
   c. Is the patient eager to collaborate in designing and carrying out an N of 1 RCT?
2. Is an N of 1 randomized trial feasible in this patient?
   a. Does the treatment have a rapid onset?
   b. Does the treatment cease to act soon after it is discontinued?
   c. Is an optimal treatment duration feasible?
   d. Can clinically relevant targets be measured?
   e. Can sensible criteria for stopping the trial be established?
   f. Should an unblinded run-in period be conducted?
3. Is an N of 1 trial feasible in my practice setting?
   a. Is there a pharmacist who can help me?
   b. Are strategies for the interpretation of the trial data in place?
4. Is the study ethical?

Reprint with permission, from Sackett DL, Haynes RB, Guyatt GH, Tugwell P: *Clinical Epidemiology: A Basic Science for Clinical Medicine*, ed #2. Boston, Little, Brown & Co., 1991.

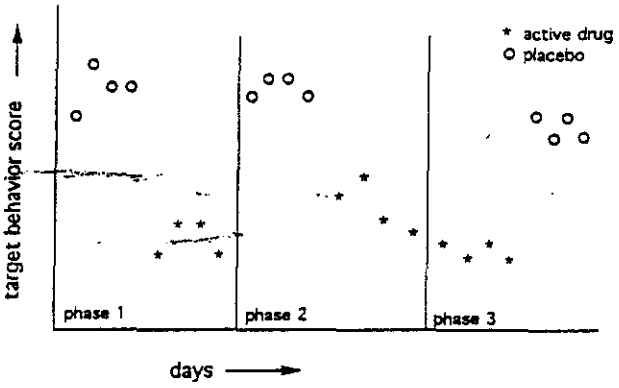

Figure 2. Results from a hypothetical N of 1 trial. In this trial, the desirable outcome is for the frequency of the target behavior to be as low as possible.

It is important to note that the treatment being evaluated must be expected to have a reasonably quick acting effect on the target outcome, and the effect must be reversible or cease when the placebo is introduced or the treatment discontinued.[4] The signs, symptoms, or target behavior must be something that is expected to occur on a regular or frequent basis. Otherwise, the causal relationship between the treatment and outcome cannot be adequately assessed with this design. For example, it would be reasonable to assess the effect of a beta agonist on asthma symptoms with an N of 1 trial, but the effect of beta blockers on reducing the incidence of myocardial infarction could not be evaluated using this design.

Although N of 1 trials require additional time and commitment on the part of clinicians and patients compared with usual treatments, they provide objective data that help to assess the outcome of treatment. The experience of Sackett and colleagues[4] with more than 50 N of 1 trials suggested that more than one-fourth of the trials resulted in major changes in long-term therapy that improved the quality of life of the individuals involved, results that would never have been achieved if the trials had not been conducted. Guyatt and colleagues[5] give an example an N of 1 clinical trial of theophylline compared with placebo in a 65-yr-old man with uncontrolled asthma.[5] After undergoing two pairs of treatment *v* placebo periods during which the patient recorded shortness of breath episodes and inhaler use, both the patient and his physician agreed there was clearly a reduction in symptoms and inhaler use during the second period of each treatment pair. When the double-blind code was broken, it was identified that the active drug was assigned to the second period during each treatment pair, thus it was determined that theophylline was having a beneficial effect for this man.

As with single-subject research, the generalizability of N of 1 trials is enhanced with replication. Zucker and colleagues[22] propose a method for combining the results of N of 1 trials to estimate population treatment effects, using a hierarchical Bayesian random effects model. In effect, this approach is a "meta-analysis" of several N of 1 trials. By applying this model to a series of 23 N of 1 clinical trials comparing amitriptyline with placebo for the treatment of fibromyalgia, they found that the population estimate of the effectiveness of amitriptyline was similar to both the direction and the level of confidence reported in published randomized, controlled trials of the efficacy of amitriptyline in patients with fibromyalgia.

SUMMARY

Individual or small group research as illustrated by case studies, single-subject research designs, and N of 1 clinical trials have an important role to play in evaluating rehabilitation interventions. The need to isolate treatment effects for an individual can rarely be addressed in between-group comparison designs but is readily demonstrated using these small-n approaches. SSRD and N of 1 RCTs have the power of repeated measures to support the inferences made regarding treatment effectiveness and can make use of clinically relevant methods of measuring target behaviors that have meaning to the individual. They are readily incorporated into the usual clinical care with only minor additional cost and, in most cases, are sufficiently flexible to accommodate the changing needs of the individuals studied. The case study is a source of hypotheses, whereas SSRD and N of 1 trials can test as well as generate hypotheses, although the generalizability of a single study is limited. Replication of studies with additional subjects, in other settings, or with different therapists will improve the generalizability of the findings. The documentation procedures developed for conducting these studies may be useful for routine documentation of client progress, identification of relevant performance indicators, data in support of quality assurance/quality improvement programs, and other methods to support evidence-based practice.

REFERENCES

1. Kazdin AE: *Single-Case Research Designs: Methods for Clinical and Applied Settings*. New York, Oxford University Press, 1982.
2. Ottenbacher KJ: *Evaluating Clinical Change: Strategies for Occupational and Physical Therapists*. Baltimore, Williams & Wilkins, 1986.
3. McEwen I (ed): *Writing Case Reports: A How-To Manual for Clinicians*. Alexandria, American Physical Therapy Association, 1996.
4. Sackett DL, Haynes RB, Guyatt GH, Tugwell P: Deciding on the best therapy, in *Clinical Epidemiology: A Basic Science for Clinical Medicine*, ed 2. Boston, Little, Brown & Co., 1991, pp 187–248.
5. Guyatt G, Sackett D, Taylor W, Chong J, Roberts R, Pugsley S: Determining optimal therapy: randomized trials in individual patients. *N Engl J Med* 1986;314:889–892.
6. Weiner IS, Eisen RG: Clinical research: the case study and single-subject designs. *J Allied Health* 1985;14:191–201.
7. Greene D, David JL: A research design for generalizing from multiple case studies. *Eval Prog Plann* 1984;7:73–85.
8. Kinney LaPier TL, Sirotnak N, Alexander K: Aerobic exercise for a patient with chronic multisystem impairments. *Phys Ther* 1998;78:417–424.
9. Harbst KB, Ottenbacher KJ, Harris SR: Interrater reliability of therapists' judgments of graphed data. *Phys Ther* 1991;71: 107–115.
10. Hersen M, Barlow DH: *Single Case Experiemental Designs: Strategies for Studying Behavior Change*. New York, Pergamon Press, 1976.
11. Portney LG, Watkins MP: Single-case experimental designs, in *Foundations of Clinical Research: Applications to Practice*. Norwalk, Appleton & Lange, 1993, pp 191–232.
12. Backman CL, Harris SR, Chisholm JM, Monette AD: Single-subject research in rehabilitation: a review of studies using AB, withdrawal, multiple baseline and alternating treatments designs. *Arch Phys Med Rehabil* 1997;78:1145–1153.
13. Gonnella C: Single subject experimental paradigm as a clinical decision tool. *Phys Ther* 1989;60:601–609.
14. Ottenbacher KJ: Clinically relevant designs for rehabilitation research: the idiographic model. *Am J Phys Med Rehabil* 1990;69:286–292.

15. Yuen HK: Improved productivity through purposeful use of additional template for a woman with cortical blindness. *Am J Occup Ther* 1993;47:105–110.

16. Hinderer S, Lehmann J, Price R, White O, DeLateur B, Deitz J: Spasticity in spinal cord injured persons: quantitative effects of baclofen and placebo treatments. *Am J Phys Med Rehabil* 1990;69:311–317.

17. Ott F, Readman T, Backman C: Mealtimes of the institutionalized elderly: a quality of life issue. *Can J Occup Ther* 1991;58: 7–16.

18. Kellegrew DH: Creating opportunities for occupation: an intervention to promote the self-care independence of young children with special needs. *Am J Occup Ther* 1998:52:457–465.

19. Backman CL, Deitz JC: Static wrist splint: its effect in three women with rheumatoid arthritis. *Arthritis Care Res* 1988;1: 151–160.

20. Harris SR, Riffle K: Effects of inhibitive ankle-foot orthoses on standing balance in a child with cerebral palsy: a single-subject design. *Phys Ther* 1986;66:663–667.

21. Bjornsen KF, Deitz JC, Blackburn S, Billingsley F, Garcia J, Hays R: The effect of body position on the oxygen saturation of ventilated preterm infants. *Pediatr Phys Ther* 1992;4:109–115.

22. Zucker DR, Schmid CH, McIntosh MW, D'Agostino RB, Selker HP, Lau J: Combinging single patient (N-of-1) trials to estimate population treatment effects and to evaluate individual patient responses to treatment. *J Clin Epidemiol* 1997;50:401–410.

CME SELF-ASSESSMENT EXAM—ANSWERS

American Journal of Physical Medicine & Rehabilitation

Vol. 78, No. 2, March/April 1999

CME Article Number Four: *T. Schulte, et al.*

1. C
2. A
3. A
4. A
5. C

CME Article Number Five: *R. Ruff, et al.*

1. C
2. C
3. D
4. D
5. B

CME Article Number Six: *K. Suzuki, et al.*

1. D
2. B
3. C
4. B
5. D