



# AOTA Evidence Briefs

## Multiple Sclerosis

*\*A product of the American Occupational Therapy Association's Evidence-Based Literature Review Project*

### MS #2

#### **Inpatient rehabilitation to benefit persons with multiple sclerosis**

Freeman, J. A., Langdon, D. W., Hobart, J. C., & Thompson, A. J. (1997). The impact of inpatient rehabilitation on progressive multiple sclerosis. *Annals of Neurology*, *42*, 236–244.

#### **Level IA1a**

Randomized controlled trial, 20 or more participants per condition, high internal validity, high external validity

#### **Clinical bottom line**

A short, multidisciplinary inpatient rehabilitation program that includes occupational therapy seems to reduce disability and overall *handicap*, but not *impairments*, in people with progressive multiple sclerosis.

In rehabilitation terminology, disability and handicap are conditions defined in relation to a root condition, impairment. Impairments are “abnormalities of body structure and appearance and with organ or system function, resulting from any cause;” disabilities are “the consequences of impairment in terms of functional performance and activity by the individual;” and handicaps are “the disadvantages experienced by the individual as a result of impairments and disabilities” (International Classification of Impairments, Disabilities, and Handicaps, 1980, p. 14) (conceptual basis used by authors for choosing which outcomes to measure).

#### **Sample**

The researchers recruited the study participants from patients in the multidisciplinary assessment clinic at the National Hospital for Neurology and Neurosurgery (London) who were subsequently admitted to an inpatient program. To be eligible, patients had to have a definitive clinical or laboratory diagnosis of multiple sclerosis, be in the secondary progressive phase of the disease (in which a gradual accumulation of wide-ranging, often complex disabilities occurs), and be considered appropriate for inpatient rehabilitation. Patients were excluded if they were experiencing a relapse, had experienced a relapse within the preceding month, were within 1 month of receiving steroids, required urgent admission, had other diseases that might interfere with outcomes, or were too cognitively impaired to give informed consent. Of 279 consecutive patients seen at the clinic over a 19-month period, 112 were admitted, and 70 met the criteria for inclusion in the study. Four participants dropped out before the second assessment. Thus the final sample was 66—22 men and 42 women, with an average age of 43.9 years.

#### **Procedures**

After they were selected for the study, the participants were sorted into three categories according to the severity of their disease—mild, moderate, or severe. They then were randomly assigned to a treatment group or a wait list (control) group. All participants were evaluated after assignment and again 6 weeks later.

#### **Outcomes**

The researchers were interested in three outcome areas: *impairment* or neurological status (as measured by Kurtzke's Functional System Scales and Kurtzke's Expanded Disability Status Scale); *disability*—specifically self-care, transfers, sphincter control (control of bowel and bladder), wheelchair locomotion, and walking (as measured

by the motor domain of the Functional Independence Measure [FIM<sup>SM</sup>]; and *handicap*—specifically physical independence, mobility, occupation, social integration, orientation, and economic self-sufficiency (as measured by the London Handicap Scale).

## Intervention

Medical, nursing, occupational therapy, and physical therapy personnel delivered multidisciplinary interventions to the treatment group for an average of 20 days (range 17 to 31). They used a patient-centered, functional, goal-setting approach, tailoring treatments to the individual participants' needs. The treatments included two 45-minute sessions of physiotherapy (the British term for physical therapy) and one session (presumably 45 minutes) of occupational therapy per day.

The control group was on a waiting list and received no intervention.

## Analyses

The researchers compared the treatment and control groups' **change scores** (*see Glossary*) for impairments, disability, and handicap, summarized in the Table. Although the overall measure of handicap showed **significant** (*see Glossary*) improvement, the **effect size** (*see Glossary*) was relatively small. FIM<sup>SM</sup> scores (except for walking) showed significant change.

## Results

For the impairment outcome, the study revealed **no significant** (*see Glossary*) differences between the groups.

For the disability outcome, the study revealed significant differences between the two groups on the overall score and on four of its five dimensions: self-care, transfers, sphincter control, and wheelchair locomotion. There was no significant difference for those who were walking before the study.

For the handicap outcome, there was a significant difference between the groups on the overall score, but there were no significant differences on the individual dimensions.

Significance and effect sizes (*r*) for outcome measures comparing the treatment and control groups for Freeman et al. (1997)

Outcome	Significance	Clinical effect ( <i>r</i> )	Size of effect
Impairment	Not significant	N/A	N/A
Disability (FIM)			
Self-care	Significant	N/A	N/A
Transfers	Significant	N/A	N/A
Sphincter control	Significant	N/A	N/A
Wheelchair locomotion	Significant	N/A	N/A
Walking	Not significant	N/A	N/A
Handicap			
Overall	Significant	.25	Small
Mobility	Not significant	.24	Small
Physical independence	Not significant	.16	Small
Occupation	Not significant	.14	Small
Social interaction	Not significant	.10	Small
Orientation	Not significant	.19	Small
Economic self-sufficiency	Not significant	.04	Negligible

N/A = *not available*

## Limitations

The sample size may have been too small to show a significant difference for the impairment and handicap outcomes (**sample size bias**) (*see Glossary*). Also, the improvement in the treatment group may have been due to the attention the members received rather than the treatment itself (**attention bias**) (*see Glossary*). Further, the evaluators were not **blinded** (*see Glossary*) to group status. Thus they may have unwittingly overestimated the treatment group's abilities or underestimated the control group's abilities. Moreover, the treatments were not well specified, so replication would be difficult (**treatment specification**) (*see Glossary*).

In addition, the researchers note that the generalizability of the findings is limited because the study was undertaken in a single specialist center and all the participants had moderate to severe disabilities.

## Reference

World Health Organization. International classification of impairments, disabilities and handicaps. Geneva: WHO, 1980.

## Glossary

**attention bias**—Also known as the Hawthorne effect, participants who receive some form of attention during treatment will often change their behavior, not because of the treatment per se, but because they are receiving attention. This bias is most frequently seen when the control group is wait listed or receives no treatment.

**blinded/blinding**—Blinding refers to the practice of keeping members of the research study unaware of which group a participant is assigned to (treatment or control) in the study. Single blinding usually refers to keeping study participants unaware of whether they are receiving the experimental or the sham treatment. Double blinding usually refers to keeping the participants and those who are administering the treatment unaware of who is receiving the experimental and who is receiving the sham treatments. In some cases, where it is impossible to blind those administering treatment, the individuals who are administering the outcome measures can be blinded to group status.

Studies in which blinding does not occur can have significant biases. When the participants know that they are receiving the experimental treatment, they often get better because they think they ought to (this is often referred to as the placebo effect). When researchers know that a participant is receiving the experimental treatment, they often subconsciously favor those participants when evaluating them on outcome measures. For instance, when timing a participant in the treatment group, researchers may unknowingly stop the watch a little faster or slower so the treatment participant seems to do better.

**change scores**—The scores created by subtracting the scores at baseline from the scores following intervention.

**effect sizes** (Cohen's  $r$ )—An effect size is a measure of clinical significance. It provides information about the magnitude of effect of the treatment. Although related to significance, it is not as influenced by the size of the sample. Therefore, it is possible to have an outcome on which the treatment had a large effect (e.g., the treatment group improved a lot more than the control group) and still have a nonsignificant result. If the results have a large effect but no significance, this means that this effect may be sample specific and not generalizable outside the study. There are many different types of effect sizes. What is reported here is Cohen's  $r$ . Cohen's  $r$  can be interpreted in a manner similar to a Pearson's correlation coefficient:

<b>Effect size <math>r</math></b>	<b>Size of the effect</b>
<0.99	Negligible
0.10 – 0.29	Small
0.30 – 0.49	Medium
>0.50	Large

Cohen, J. (1977). *Statistical power analysis for behavioral sciences*. New York: Academic Press.

**nonsignificant or no significance**—A statistical term that refers to study findings that are likely to be due to chance differences between the groups rather than to other factors (like the treatment of interest). A nonsignificant result is not generalizable outside the study. Like significance, a nonsignificant result does not indicate the clinical effect. Often studies will show nonsignificant results, yet the treatment group's mean will be better than the control group's. This is usually referred to as a trend in the right direction. Because significance is closely determined by sample size, nonsignificant results would often become significant if the sample size were increased.

**sample size bias**—Significance is strongly related to sample size. A study that has too small a sample will not show significance (a type II error), even when a treatment effect is present. Some research studies that are non significant demonstrate a trend toward the treatment having the desired effect, suggesting that a larger sample is needed to detect a significant treatment effect. On the other hand, too large a sample can prove just about anything.

**significance (or significant)**—A statistical term, this refers to the probability that the results obtained in the study are not due to chance, but to some other factor (such as the treatment of interest). A significant result is likely to be generalizable to populations outside the study.

Significance should not be confused with clinical effect. A study can be statistically significant without having a very large clinical effect on the sample. For example, a study that examines the effect of a treatment on a client's ability to walk may report that the participants in the treatment group were able to walk significantly longer distances than the control group. However, if you read the study you may find that the treatment group was able to walk, on average, 6 feet, whereas the control group was able to walk, on average, 5 feet. Although the outcome may be statistically significant, a clinician may not believe that a 1-foot increase will improve his or her client's function.

**treatment specification**—One of the goals of research is to ensure that any results can be replicated by others. In addition, clinician who are going to use researched protocols as part of their treatment need enough information about the treatment to implement it. Many treatment studies fail to clearly state the treatment intervention received by the experimental group, making replication and clinical application impossible.

■ Terminology used in this document is based on two systems of classification current at the time the evidence-based literature reviews were completed: *Uniform Terminology for Occupational Therapy Practice—Third Edition* (AOTA, 1994) and *International Classification of Functioning, Disability and Health (ICIDH-2)* (World Health Organization [WHO], 1999). More recently, the *Uniform Terminology* document was replaced by *Occupational Therapy Practice Framework: Domain and Process* (AOTA, 2002), and modifications to *ICIDH-2* were finalized in the *International Classification of Functioning, Disability and Health* (WHO, 2001).

This work is based on the evidence-based literature review completed by Nancy Baker, ScD, OTR, and Linda Tickle-Degnen, PhD, OTR/L, FAOTA.

For more information about the Evidence-Based Literature Review Project, contact the Practice Department at the American Occupational Therapy Association, 301-652-6611, x 2040.

