SPECIAL COMMUNICATION

Toward Improved Evidence Standards and Methods for Rehabilitation: Recommendations and Challenges

Mark V. Johnston, PhD, Marcel P. Dijkers, PhD

ABSTRACT. Johnston MV, Dijkers MP. Toward improved evidence standards and methods for rehabilitation: recommendations and challenges. Arch Phys Med Rehabil 2012;93 (8 Suppl 2):S185-99.

Interventions and programs for people with disability should be based on the best—the most discriminating and rigorous—methods of systematic review and knowledge translation possible. Extant systems for systematic review and practice recommendations have excellent features but severe difficulties are encountered when attempting to apply them to disability and rehabilitation. This article identifies issues in evidence synthesis and linked practice recommendations and describes both new and long-tested methods to address them. Evidence synthesis in disability and rehabilitation can be improved by: explicating criteria for evaluating nonrandomized evidence, including the regression discontinuity, interrupted time series, and single-subject designs, as well as state-of-the-art methods of analysis of observational studies; greater use of meta-analysis; considering effect size, direction of biases, and dose-response relationships; employing more discriminating methods of evaluating flaws in masking, considering also measurement reliability and objectivity; considering overall biases and conflicts of interest; increased attention to composition of review panels; and greater transparency in reporting of the bases of reviewers' judgments. Review methods need to be developed for assistive technology and for measurement procedures. Application to practice can be improved by attention to treatment alternatives, explicit evaluation of generalizability, synthesizing clinical experience as a source of evidence, and a focus on the best—rather than the ideally most-rigorous—evidence. Study outcomes should be measured and reviewed in terms meaningful to persons served. In sum, methods are available to improve evidence synthesis and the application of resulting knowledge. We recommend that these methods be employed.

Key Words: Chronic disease; Disabled persons; Evidence based practice; Practice guidelines as topic; Rehabilitation; Research design; Systematic review as topic.

DISCIPLINES CONCERNED with disability and rehabilitation (D&R) face a number of challenges in the synthesis of evidence and its application to practice, and there is much unease about the application of standard methods of systematic review (SR), largely developed within evidence-based medicine (EBM).[1] Much of this unease is related to the relative lack of randomized controlled trials (RCTs) in the field: some advocates of EBM do not accept any evidence weaker than that produced by RCTs, making it appear as if there is little or no evidence in D&R. While it is true that the relative paucity of RCTs limits knowledge of treatment efficacy, RCTs are far from the only useful method of obtaining evidence on treatment effectiveness. Conclusions that "the evidence is insufficient" do not help inform clinical decision making or policies. Methodologies are needed that are capable of identifying and synthesizing the current best evidence—not just the theoretically most rigorous evidence.[2,3]

There are now many published methods of systematic evidence review and associated methods of making practice recommendations,[4-14] as well as a large body of work on associated topics such as statistical inference, research design, and knowledge translation. While the literature provides a wealth of tools, it also raises the question: Which one should we use? Improved methods, insights, and standards are published every year. Which methods are up to date? The view espoused in this article is that evidence synthesis and practice recommendations for D&R should be based on the best, most discriminating, and best validated methodologic knowledge and evidence synthesis procedures that take into account the nature of D&R. Practicality, reliability, and expense of review also need to be considered. New or improved methods should be incorporated as they are published and proven to be sound, even if they are unfamiliar to us. The current article draws lessons from the great body of work in EBM, evidence-based practice (EBP),

List of Abbreviations

AAN	American Academy of Neurology
ACRM	American Congress of Rehabilitation Medicine
AT	assistive technology
D&R	disability and rehabilitation
EBM	evidence-based medicine
EBP	evidence-based practice
GRADE	Grading of Recommendations Assessment, Development and Evaluation
IES	Institute of Education Sciences
PEDro	Physiotherapy Evidence Database
RCT	randomized controlled trial
RDD	regression discontinuity design
SR	systematic review
SSD	single subject/case design
TSRD	time series research design

From the Department of Occupational Science and Technology, University of Wisconsin-Milwaukee, Milwaukee, WI (Johnston); and Department of Rehabilitation Medicine, Mount Sinai School of Medicine, New York, NY (Dijkers).

Presented to the AcademyHealth Annual Research Meeting, June 11, 2011, Seattle, WA. The points in the article have also been discussed in meetings of the American Congress of Rehabilitation Medicine's Clinical Practice Committee and the Task Force on Standards and Methods of Evidence of the National Center for Dissemination of Disability Research.

Supported by the National Institute on Disability and Rehabilitation Research (grant no. H133A060028) and intramural funds of the University of Wisconsin-Milwaukee.

No commercial party having a direct financial interest in the results of the research supporting this article has or will confer a benefit on the authors or on any organization with which the authors are associated.

Correspondence to Mark V. Johnston, PhD, Enderis 963, Dept of Occupational Science and Technology, University of Wisconsin-Milwaukee, PO Box 413, Milwaukee, WI 53201, e-mail: *johnsto@uwm.edu.* Reprints are not available from the author.

0003-9993/12/9308S-00969$36.00/0
doi:10.1016/j.apmr.2011.12.011

and research methodology in many fields, concentrating on newer insights, and applying them to D&R.

Given the limited number of RCTs, review methods in D&R need to be sensitive to non-RCT evidence if they are to identify best evidence.[2] Methods are needed that are sensitive to the particular characteristics of the behavioral, educational, and psychosocial interventions that characterize much of D&R. In this article, we use the terms behavioral and psychosocial as umbrella terms to describe therapies that work through the actions of the person. These nondrug and nonsurgical interventions include motor learning, physical training, education, cognitive-behavioral learning, and other interventions provided by physical therapists, speech/language pathologists, psychologists, and occupational therapists, and exclude assistive technology (AT) and environmental modifications. We also need to identify reasonable methods of interpolation or generalizing from the strong evidence that we do have.

This article identifies and describes leading-edge developments and methodologies that can improve evidence synthesis and its application in D&R. Because there is a lack of understanding and expertise in best practices in evidence synthesis and application in D&R, the article first provides a brief review of standard methods of synthesizing evidence, emphasizing widely respected methods which we consider to be best.[5] We then make recommendations or suggestions to improve evidence synthesis and the application of this evidence to clinical practice, based on the experience of systematic reviewers and guideline panels, newer evidence synthesis methods, such as Grading of Recommendations Assessment, Development and Evaluation (GRADE),[15-18] and recent methodologic work. We also reference newer methods of statistical analysis and research designs, which are commonly ignored in extant evidence grading publications but which are capable of providing increased accuracy and insight. In cases where the newer methods are well-developed and tested, we recommend that they be incorporated into the process of evidence synthesis and development of practice recommendations. In other cases, the needed direction and key considerations are stated without an exact, demonstrably optimal solution. Throughout, we present key points and references to assist future work on the matter, but due to space limitations, we provide limited discussion and support for some recommendations. There are many details that need to be worked out in future discussions.

Our focus is on methods for evaluating strength of evidence (or conversely, risk of bias) for research inferences and recommendations for clinical practice. Meta-analysis, while not our focus, is recommended when it is likely to improve the validity and robustness of results of SR (see Recommendation 2e). When we use the term clinical practice, we mean helpful interventions of many types, including interventions that occur in homes, workplaces, and other community settings. We use the terms treatment and intervention largely interchangeably, although the latter is somewhat broader.

Our recommendations and suggestions are presented for purposes of education, discussion, and debate. As a product of discussions in the Clinical Practice Committee of the American Congress of Rehabilitation Medicine (ACRM) and in the Task Force on Standards of Evidence and Methods of the National Center for Dissemination of Disability Research, this article is meant to serve as an authoritative report or white paper. The hope is to motivate the serious work, discussions, and consensus-building necessary to develop improved standards and methods that, while incorporating rigor, also improve the sensitivity of SRs and the quality of related practice recommendations.

EBP CONCEPTS

EBP involves more than the notion that published studies should be applied to clinical practice.[19] EBP is classically defined as the application of best evidence to clinical practice and its integration with client values and clinical expertise.[20] By clinical practice, we mean interventions of all kinds, whether applied in a clinic, in homes, or worksites in the community, including not only medical treatments but also behavioral, psychosocial, and educational programs and environmental modifications. To identify best evidence, one should do a complete search of the relevant literature and grade the quality or strength of the evidence. Resulting recommendations for practice should emphasize studies with the strongest research designs and the data most applicable to the questions at hand. Unlike traditional unstructured reviews, SRs employ preestablished criteria, based on tested methodologic knowledge, to reliably judge strength of evidence and to minimize bias.

Evidence reviews aim to inform recommendations for clinical practice. The research results applied to practice should be very sound and of very high quality, but a degree of judgment is virtually always required. Both overall quality of evidence and resulting recommendations are graded in terms of multiple levels rather than in traditional dichotomies (eg, rigorous vs not). While particular cut points for grading levels of evidence and recommendations are somewhat arbitrary (eg, level 1 vs level 2, or strong vs weak recommendations), expressing results in terms of levels provides more discriminating and transparent information than simplistic dichotomies or free-form expert opinion employed in the pre-EBP era. The core issue underlying levels is degree of certainty and freedom from bias: Is the finding unlikely, possibly, or likely to be altered or even reversed with new, better research? While synthesizing evidence can be complex, it is possible to develop reasonably direct procedures that even students can use to reliably rate research in most cases (eg, Physiotherapy Evidence Database [PEDro][21]), referring the more complex issues to experts. All professional evidence authorities—American Academy of Neurology (AAN),[5,14] the Cochrane Collaboration,[6] GRADE,[15-18] Agency for Healthcare Research and Quality, and the Institute of Medicine[3,10,11]—emphasize the need for objectivity and transparency in SRs and guideline formulation: biases associated with the personal, program, or professional interests of reviewers must be minimized.

A general recommendation is presented first, followed by recommendations oriented toward internal validity and then recommendations directed toward external validity and application to practice.

RECOMMENDATION 1: DEFINE OUTCOMES IN TERMS MEANINGFUL AND IMPORTANT TO THE PERSONS SERVED

Outcomes selected for SR and EBP should be those that patients and people with disability care about, such as health, activities and participation in everyday life, reduction in pain or distress, and quality of life. The outcomes chosen for review in evidence synthesis and used in subsequent practice recommendations should be those that are important to the individuals who experience the problem being addressed and are expected outcomes of interventions under consideration. Laboratory or other technical measures are valuable and often needed supplements to confirm and explain changes in the things patients care about, but they are proxy rather than primary outcome measures. Similarly, the views and values of disability spokespeople are valuable, but the results experienced by people