European Journal of Physical and Rehabilitation Medicine 2024 December;60(6):1100-11
DOI: 10.23736/S1973-9087.24.08713-6

SPECIAL ARTICLE

The role of single case experimental designs in evidence creation in rehabilitation

Wendy MACHALICEK [1] *, Douglas P. GROSS [2], Susan ARMIJO-OLIVO [3], Giorgio FERRIERO [4], Carlotte KIEKENS [5], Rachelle MARTIN [6], Margaret WALSHE [7], Stefano NEGRINI [5, 8], 5th Cochrane Rehabilitation Methodological Meeting participants ‡

[1]University of Oregon, Eugene, OR, USA; [2]University of Alberta, Edmonton, AB, Canada; [3]University of Applied Science, Osnabrück, Germany; [4]University of Insubria, Varese, Italy; [5]IRCCS Galeazzi Orthopedic Institute, Milan, Italy; [6]University of Otago, Dunedin, New Zealand; [7]Trinity College, Dublin, Ireland; [8]University "La Statale", Milan, Italy

‡ Members are listed at the end of the paper

*Corresponding author: Wendy Machalicek, Special Education and Clinical Sciences; College of Education, University of Oregon, Eugene, Oregon, USA. E-mail: wmachali@uoregon.edu

ABSTRACT

Randomized controlled trials (RCTs) are considered the gold standard of evidence guiding intervention selection in rehabilitation. However, conduct of sufficiently powered RCTs in rehabilitation can be expensive, pose ethical and attrition concerns when participants are assigned to ineffective treatment as usual conditions, and are infeasible with low-incidence populations. Single-case experimental designs (SCEDs), including N-of-1 RCTs are causal inference studies for small numbers of participants and not necessarily one participant as the name implies. These designs are increasingly used to evaluate the effectiveness of rehabilitation interventions in diverse clinical settings and employ design features including but not limited to randomization and each participant serving as their own control. These and other internal validity enhancements can increase the confidence in study results coming from these designs. This manuscript discusses the expanded application of SCEDs in rehabilitation contexts to answer everyday clinical rehabilitation research questions with emphasis on strategies to use: 1) to maximize internal validity of this family of designs; 2) improve utility, effectiveness, and acceptability of these designs for rehabilitation end-users (clinicians, policymakers, and participants); 3) build evidence bases in areas of rehabilitation where RCTs are uncommonly used. Primary considerations for situating SCEDs within the continuum of experimental designs include increasing internal validity within designs, improving transparency in conduct and reporting of these studies, and increasing access to advanced research methods training for rehabilitation professionals.

(*Cite this article as:* Machalicek W, Gross DP, Armijo-Olivo S, Ferriero G, Kiekens C, Martin R, *et al*.; 5th Cochrane Rehabilitation Methodological Meeting participants. The role of single case experimental designs in evidence creation in rehabilitation. Eur J Phys Rehabil Med 2024;60:1100-11. DOI: 10.23736/S1973-9087.24.08713-6)

KEY WORDS: Single-case studies; Rehabilitation; Research design.

Across several fields and disciplines in health care, including rehabilitation, researchers favor randomized controlled trials (RCTs) because well-controlled and sufficiently powered RCTs are argued to provide high levels of internal and statistical conclusion validity. However, RCTs can be financially prohibitive, treatment as usual conditions can pose ethical, recruitment and attrition concerns if the usual conditions are less effective than novel interventions, patients with the same condition can present differently, necessitating intervention adaptation and can be infeasible with low incidence populations and issues.[1] For these reasons, rehabilitation contexts favor adapted or

individualized intervention protocols rather than the "one size fits all" approach of RCTs.

Rehabilitation research requires a continuum of research approaches mapped onto patient populations, issues, contexts, and positionality of research questions within scholarly agendas. For research questions requiring experimental designs minimizing threats to internal validity and implementation with a small number of participants, rigorous SCEDs offer an attractive alternative. Researchers have noted this research approach's unique contribution to evidence creation and SCEDs are widely employed in many disciplines and fields including, but not limited to, education, clinical psychology, and healthcare, inclusive of rehabilitation.[2-5] Importantly, these designs can offer comparable internal validity to RCTs by using randomization strategies and blinding measurement, and they can contribute to external validity when replicating experiments across researchers, contexts, and participants.

Characteristics of single-case experimental design

SCEDs are experimental, causal inference time series studies focused on the individual as the unit of analysis. An expansive family of designs are available with common design features to pre-emptively address threats to internal validity, increase credibility, and improve external and social validity. SCEDs are well suited for answering research questions about effectiveness or efficacy of interventions, to build interventions or identify essential components or specific behavioral mechanisms, and to examine the comparative effectiveness of interventions or conditions in a study.[6] A recent paper by Yang, Armijo-Olivo and Gross comprehensively described characteristics of SCEDs;[7] thus, we refer readers to that paper to deepen knowledge on SCED characteristics and design logic.

Purpose of the present paper

Methods and analysis strategies used in SCED research have rapidly evolved over the past 20 years and there is renewed interest in the methodology and creation of design, analysis, and reporting standards.[8-12] Recent inclusion of SCED studies in meta-analyses defining evidence-based practice in education and related fields[13] has been fertile ground for methodologists. This unique time point encourages discussion of the design and reporting of SCEDs in rehabilitation and a need to situate their place in evidence creation, particularly in situations where RCTs are untenable.

This paper addresses key objectives regarding the position of SCED, including the use of N-of-1 RCTs within the scientific enterprise in rehabilitation. Attendees of the 5th Cochrane Rehabilitation Methodological Meeting September 2023 in Milan, Italy discussed four methodological objectives, which included clarifying: 1) when to apply SCEDs in rehabilitation contexts to answer applied clinical research questions; 2) strategies for maximizing the internal validity of SCEDs to produce rigorous studies and ultimately useful evidence in rehabilitation; 3) when and how SCEDs contribute to building the evidence in rehabilitation if RCTs are not yet available (*e.g*., understudied interventions, participants, and/or contexts where rigorous causal determination type research has been historically lacking); and 4) the conduct and reporting of SCED studies to improve usability for rehabilitation end-users (patients, research participants, clinicians, policymakers, and funders).

Objective 1: application of single-case experimental design in rehabilitation

Basic single-case experimental designs

Withdrawal/reversal designs

The most basic of SCEDs, built upon the non-experimental AB design where A is baseline and B is intervention, is the reversal or withdrawal design (ABAB).[14] These designs require reversible dependent variables (*e.g*. outcomes improve with preferred context or delivery of desired consequence, visual aids, or assistive technology such as increased functional hand use with use of standing frame). Figure 1 illustrates a hypothetical ABAB design with a single participant with multiple sclerosis where the first phase is baseline, the second phase is the first introduction of intervention (*i.e*., use of a digital planner with alarms), the third phase is a return to baseline (*i.e*., withdrawal of intervention), and the final phase is reintroduction of the intervention. The design provides opportunity for three demonstrations of the basic effect of intervention indicated by the 1, 2, and 3. The first demonstration occurred between the first two phases (Baseline-Intervention), the second demonstration occurred between the second and third phases (Intervention-Baseline) when the intervention is removed and participant response returns to initial baseline levels, and the third demonstration occurred between the third and final phases when the intervention is reintroduced. In Figure 1, the research question is, "Is there a functional relation between use of digital planner

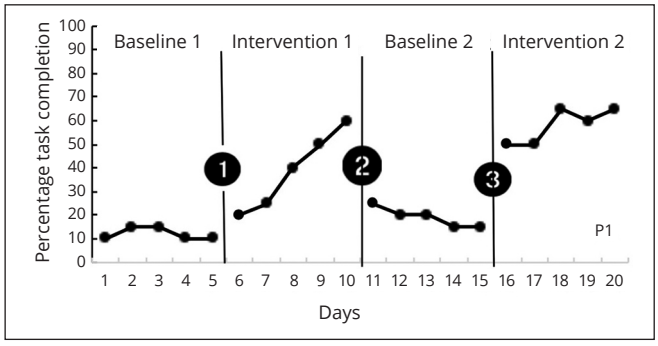

Figure 1.—Hypothetical ABAB withdrawal design.

with alarms and increased level of completion of domestic tasks at home?" The dependent variable for this question is percentage of task completion, the independent variable is the use of a digital planner with alarms. Although a common clinical concern of this design is removal of the intervention, this design aligns well with the use of accommodations and aids to determine if they add benefit to an existing intervention and provides a rigorous experimental design with single participants at a time.

Multiple-baseline designs (MBD) and multiple-probe designs

Figure 2 illustrates a hypothetical concurrent multiple-baseline design across three adult participants with Parkinson's disease. The research question for this example is, "Is there a functional relation between mobility training and improved level of mobility as evidenced by improved scores on TUG?". The independent variable in this hypothetical study is mobility training, and the dependent variable is mobility as measured via observation of participant completion of Timed Up and Go (TUG) Test Time in seconds.[15] Concurrent refers to baseline data collection beginning on the same day for each participant. In addition to introducing intervention to different participants in a time-lagged fashion, researchers can design phases across different settings and behaviors or skills if tiers are independent of one another. Non-concurrent MBD have participants start baseline at different points in time which could include start times with differences in days, weeks, months, or years.[16] Opinions differ on the comparative experimental control provided by concurrent and non-concurrent MBD;[17, 18] however, growing researcher opinion situates non-concurrent MBD as sufficient designs for examining whether a functional relation occurs in a study if effects are replicated across at least three cases at three different points in time.[17] A researcher might choose a non-concurrent MBD if they infrequently see a condition

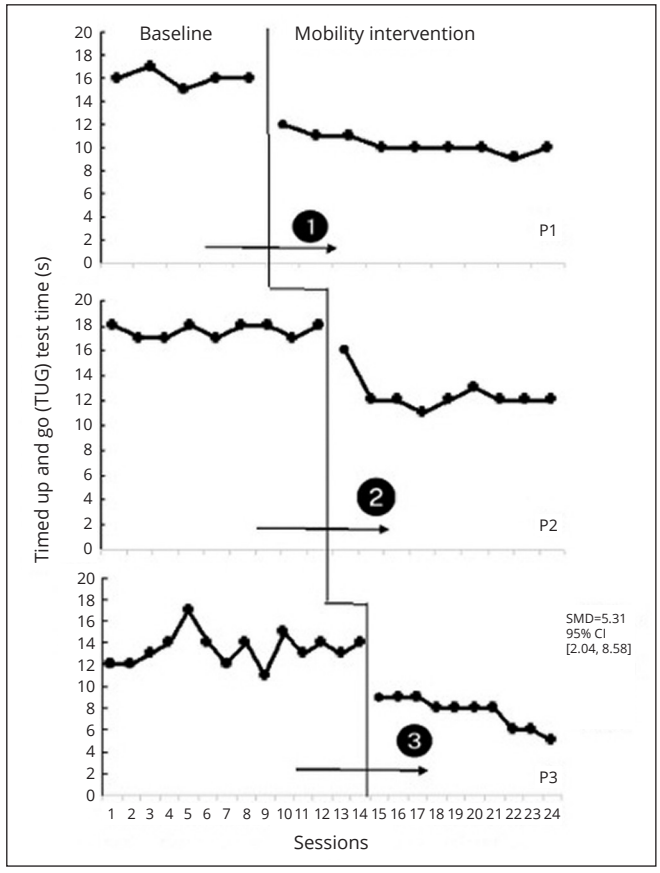

Figure 2.—Hypothetical concurrent multiple-baseline design.

within their clinical practice or face difficulty recruiting sufficient participants for a concurrent MBD. For example, a researcher examining the effectiveness of treadmill training on the steps taken and gait of girls with Rett syndrome, a rare neurogenetic condition, could implement the study across several months to align with the frequency with which patients with Rett syndrome attend clinics.

Multiple-probe designs (MPD) are best suited to dependent variables with predictable data patterns, in situations where repeated measurement is demoralizing or frustrating for participants or increases risk of testing as a threat to internal validity (*e.g.*, improvement via practice of a skill during baseline phase).[16] Figure 3 transforms the MBD illustrated in Figure 2 into a MPD across participants. The researcher performs repeated continuous data collection only for the first participant (P1) and then for the remaining participants only collects data intermittently at predefined time points. First, data is collected for each of the remaining participants (Participant P2, P3) at the beginning of baseline phase, again when participants above

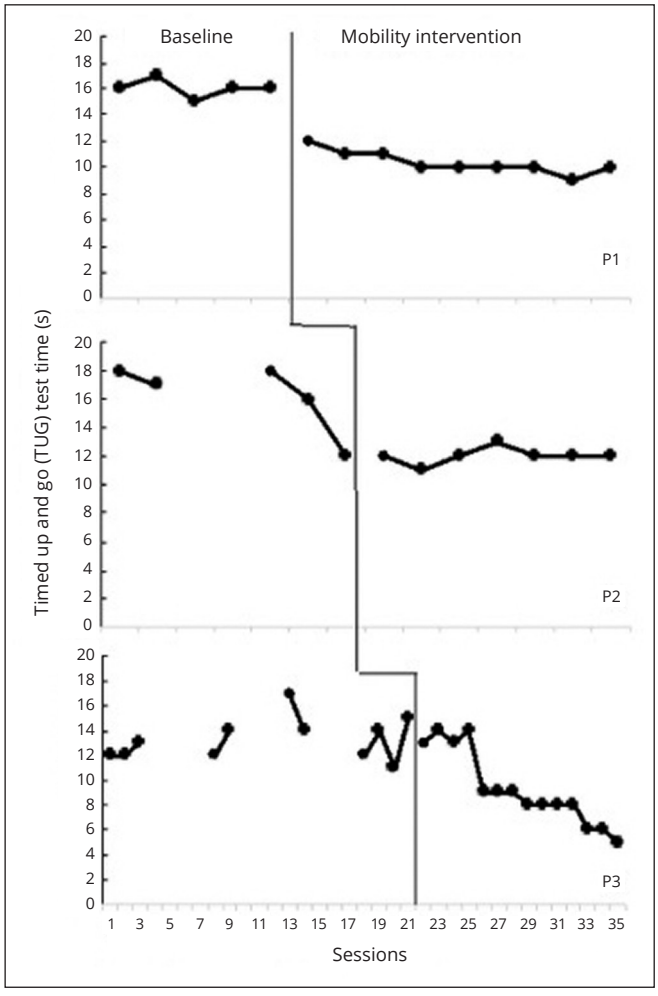

Figure 3.—Hypothetical multiple-probe design.

the present tier are beginning intervention, and finally just prior to entering intervention. This design provides less information about potential threats to internal validity arising from multi-treatment effects (*i.e.*, participant receives another intervention outside of the intervention study and their responses change during baseline) or maturation. MPD also require working knowledge of the median level of variability expected during baseline assessment since fewer data points are available to document the issue and demonstrate lack of threat to internal validity. Continuous data collection with a concurrent or non-concurrent MBD will better suit highly variable patterns of participant responding. A researcher might choose a MPD whenever they anticipate: 1) testing effects with repeated practice during baseline such as when examining the effectiveness of retrieval practice on the memory recall of participants with acquired brain injury; or 2) participant frustration during baseline that could result in demoralization and attrition. Both non-concurrent MBD and MPD readily apply to clinical settings where scholarly agendas map onto clients and issues as they arise in everyday practice.

Alternating treatments design

These designs employ rapid alternation between two or more intervention conditions and examine differential impact on the same dependent variable(s).[19, 20] Figure 4 illustrates a hypothetical alternating treatments design (ATD) with two intervention conditions of virtual reality enhanced rowing machine exercise or traditional rowing machine exercise applied to a single adolescent with cerebral palsy. The dependent variable is strokes per minute. ATD requires an easily and rapidly reversed dependent variable when intervention conditions change. Additionally, introducing the intervention condition should produce an immediate effect on the dependent variable. Participants must be able to discriminate between intervention conditions, and condition order is randomized with no more than two of the same intervention condition implemented in order. Researchers can opt to include a baseline phase and/or a best treatment phase and a control condition can be incorporated in case of lack of differentiated participant responses between intervention conditions (*e.g.*, random, rapid alternation between control condition, intervention condition A and intervention condition B). These designs offer a unique mechanism for including participants in the choice-making process regarding next intervention steps since they experience both intervention conditions and they and/or their caregivers can gauge an intervention's relative effectiveness and acceptability. Visual analysis differs for ATD; magnitude and degree of separation be-

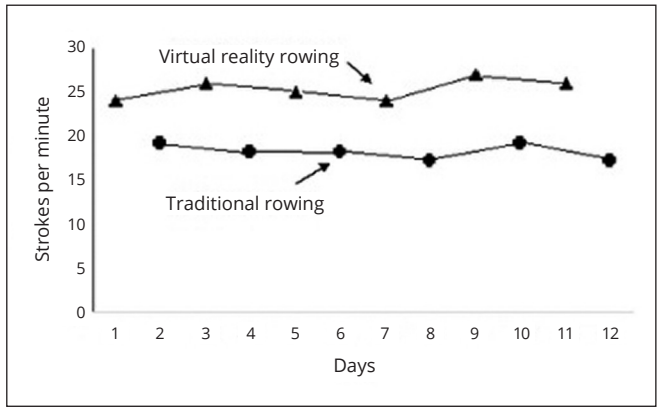

Figure 4.—Hypothetical alternating treatments design.

tween conditions as well as level, trend, and variability within conditions are core discriminations.

Other single-case design choices

Less commonly employed basic designs include the changing criterion design,[21] parallel treatments design,[22] and repeated acquisition design.[23] Methodological innovations exist to increase internal validity of these designs including range bound criteria, increased number of replications, varying the length of data prior to phase change, and use of an embedded reversal design where the participant returns to prior levels.[21-23]

N-of-1 randomized control trials

N-of-1 RCTs developed from SCEDs have been primarily employed to evaluate pharmaceutical or surgical interventions, to determine effectiveness of an intervention for an individual (non-experimental A-B design); or to compare two or more intervention conditions.[24, 25] The basic N-of-1 RCT consists of an AXYXY design (also referred to as an ABABA design or single-participant crossover trial) including a baseline or "run-in" phase followed by the introduction of an intervention or placebo "X" followed by withdrawal of intervention or placebo "X" and implementation of a different intervention "Y" or placebo or no intervention followed by subsequent withdrawal of intervention, placebo or no intervention "Y" (also called a washout phase in drug research) followed by reintroduction of X.[24] X and Y conditions are implemented in a randomly assigned order. Readers will note similarity between the XYXY component of the N-of-1 RCT and the ABAB design described earlier (Figure 1). By design the ABAB design ends on an intervention phase; researchers employing N-of-1 RCTs can choose to end on a phase of "best treatment." Depending on the intervention and researcher resources, these trials can be either single or double-blinded, but in behavioral, rehabilitation, or educational interventions, these trials often lack blinding due to impossibility of arranging necessary blinding. Researchers should look at innovative ways to reduce performance and detection bias as suggested in the literature.[1]

Hybrid designs

Basic SCEDs can be combined flexibly into a hybrid design allowing researchers to answer additional research questions and address design limitations inherent in primary design. Researchers can *a-priori* plan a hybrid design or include a secondary design to investigate serendipitous findings or address a loss of experimental control by building in additional demonstrations of the basic intervention effect. More commonly used hybrid designs include MBD with embedded reversal/withdrawal design, MBD with embedded ATD within the intervention condition, and withdrawal/ reversal design with embedded alternating of two intervention conditions or degree of intervention implementation (*i.e.*, ATD) within intervention phases.[26] For these specific combination designs, single and two-level hierarchical linear modeling are suggested statistical analysis approaches to supplement visual analysis.[27]

Objective 2: strategies to maximize internal validity in single-case experimental design

The most common threats in SCEDs include testing effects (during baseline assessment), the Hawthorne effect, maturation, history effects, and multi-treatment effects. SCED intervention researchers and methodologists have developed strategies to decrease and mitigate threats to internal validity and to increase transparency with particular emphasis on increasing the internal validity of studies *via* reduction of Type 1 errors associated with confirmation bias, and the use of protocol pre-registry to improve trustworthiness of the multitude of decisions a SCED researcher makes before and during the study.

Pre-registration of protocol

The quality of information reported in rehabilitation studies has been an ongoing concern motivating recent adoption of the Cochrane Rehabilitation RCTRACK project and GUIDE-Rehab projects.[28] The EQUATOR (Enhancing the Quality and Transparency of Health Research) Network also includes an N-of-1 checklist[29] (CONSORT Extension for N-of-1 Trials or CENT) for reporting quality and SCRIBE[30] provides another option for rehabilitation researchers designing and reporting SCED studies. However, these types of checklists do little to address concerns about publication bias contributing to overwhelming positive outcomes published and questionable research practices like dropping cases with small effect size.[31, 32] In addition, post hoc description of study procedures can lead to inaccurate or incomplete reporting, negatively impacting the trustworthiness of available research.

One potential solution is public third-party pre-registration of study methods inclusive of design and analysis. Pre-registration addresses publication bias occurring when only studies with positive outcomes are accepted, since meta-analyses can include unpublished, registered studies.

Additionally, editorial teams and reviewers can verify that researchers conducted research in the way reported and pre-registration provides opportunities for researchers to obtain feedback from other researchers on their methods which might reduce redundancy in research efforts. Pre-registration of SCED studies is a core upgrade in research design standards for single-case intervention research studies.[8] Options for pre-registration of rehabilitation studies include any of the primary registries listed by the World Health Organization's International Clinical Trials Registry Platform (ICTRP; 20198) and ClinicalTrials.gov. Single-case researchers have suggested that pre-registration of a SCED study should at minimum include: 1) basic descriptive information; 2) research questions; 3) participant characteristics; 4) thorough description of baseline condition characteristics and the independent variable/intervention condition; 5) dependent variables (*i.e.*, outcome measurement including timing of measurement); 6) hypotheses as relevant; and 7) phase change decision rules or number of data points to be collected in each phase.[33-35] Researchers should also report the validity and reliability of measurements used where available.

Response guided approaches and randomization strategies

Response-guided SCED allows researchers to obtain "steady state" behavior to demonstrate the presenting issue and document threats to internal validity (or absence thereof).[36] Researchers respond to observed variation in data patterns by obtaining more data points which has the effect of easing identification of a functional relation between the intervention and the dependent variables (and reduces Type 2 errors), but introduces bias (*i.e.*, increased probability of Type 1 errors). In addition, using a response-guided approach disallows statistical hypothesis testing using traditional parametric effect sizes and confidence intervals.[9]

In efforts to decrease bias introduced by response-guided approaches using visual analysis and increase statistical conclusion validity of SCEDs, Levin *et al.*[9, 37-40] have developed various randomization strategies for SCEDs that increase a study's internal validity and statistical conclusion validities. These strategies are summarized in Table I. Levin and Kratochwill provide a comprehensive description of these randomization strategies and their contributions to internal and statistical conclusion validities within clinical trials in medicine.[39] Open-source software to support planning and analysis of these designs is available in Microsoft Excel based ExPRT package[40] and SCRT-R software.[41]

Some randomization strategies like intervention start point randomization are best suited to instances where researchers can predict potential participant responses, specifically the window around median variability in baseline condition, and do not anticipate a problematic (*i.e.*, therapeutic) trend during baseline. Importantly, response-guided approaches to SCED often include various forms of randomization to boost internal validity but have less frequently capitalized on randomization strategies to calculate parametric effect sizes. N-of-1 RCTs require randomized selection of the intervention and blinding. However, researchers can incorporate the randomization strategy used in N-of-1 RCTs within a SCED, allowing for other forms of randomization and replication within the same study across participants, settings, or behaviors within the same participant. For those concerned about blinding assessors, participants and interventionists to rehabilitation interventions, researchers can build additional credibility

TABLE I.—*Procedural definition of randomization strategies and applicability to single-case experimental designs.*

Strategy	Definition	Design applicability
Case randomization	Random selection of the order in which cases enter intervention in a time-lagged, staggered manner	Multiple baseline designs; multiple probe designs
Within-case intervention-order randomization	Random order of presentation of experimental phases for each case (*e.g.*, random selection of participant encountering intervention 1 or intervention 2 first in multiple-phase crossover design)	AB or multiple-phase crossover designs; reversal/withdrawal design; alternating treatments designs
Between-case intervention randomization	Random selection of which cases encounter intervention condition, baseline condition or placebo control condition for each case with counterbalancing first condition across cases (*e.g.*, one half of the participants in a multiple-baseline design are randomly assigned to receive intervention 1 and the other half are randomly assigned to receive intervention 2)	AB or multiple-phase crossover designs; reversal/withdrawal design; alternating treatments designs
Intervention start-point randomization	Random selection of when a case moves from one experimental phase to another (*e.g.*, random selection of intervention phase start-point for each case)	All single-case designs

into the design by incorporating multiple types of randomization, affording a higher confidence that any observed intervention effect is due to intervention.

Masked visual analysis

One useful strategy for decreasing Type 1 error inherent in the response-guided approach of SCEDs is masked visual analysis (MVA).[42, 43] This approach allows single-case researchers flexibility to retain visual analysis as primary analysis of data patterns. By separating the intervention and analysis teams and blinding members of the analysis team to when the intervention began with each participant, MVA provides a unique approach for minimizing bias introduced by visual analysis. This approach allows for incorporation of randomization so statistical hypothesis tests can be used (and a P value derived). However, unlike randomization SCED, MVA allows researchers to incorporate randomization without sacrificing the logic of SCED, which often requires collecting more data in response to individual participant variation. Randomization allows for appropriate use of statistical hypothesis tests and is more highly powered to detect effects in this response-guided arrangement than a fixed phase length.[43]

The MVA approach provides a robust procedure guarding against the bias introduced with visual analysis of line graph data and allows for blinding. Similar procedures to blind statistical analysis have been encouraged when conducting RCTs to increase internal validity.[1] Limitations include potentially increased length of study, and the method requires a minimum of four participants or alteration of the randomization strategy to include a random selection of "no transition." If the researcher cannot assure predictable stability of data around the median data path during baseline, and transition states as the intervention is introduced, it can pose some difficulty to the *post-hoc* guesses of the analysis team due to high levels of variation in participant data. Procedural details are provided in published SCEDs using MVA procedures.[42]

Objective 3: contribution of single-case experimental design to building evidence in rehabilitation

The Oxford Centre for Evidence-Based Medicine currently considers either N-of-1 randomized trials or systematic reviews of randomized trials as Level 1 evidence for both treatment benefits and treatment harms related research questions.[44] We presume the level of evidence provided by SCED more broadly is often downgraded by a presumption these designs are used to informally assess participant response to an intervention and lack sufficient design elements such as randomization and blinding to counter threats to internal validity. However, SCEDs exist on a wide continuum and include rigorous designs capable of causal inference and reduced internal validity threats. When designed with rigor, implemented and analyzed with care, and transparently reported, SCEDs offer rehabilitation researchers an efficient methodology for intervention development, to examine effectiveness or efficacy of interventions, to identify effective interventions for populations with suboptimal outcomes, to incorporate participant choice of intervention to increase adherence and to build external validity through implementation and evaluation of interventions implemented in natural settings by natural change agents. A key point is that the unit of analysis of SCED studies is at the individual participant level which allows for the identification of variables that contribute to intervention response or non-response. These studies also allow for exploration of and documentation of the transition state (*e.g.* how long does it take to achieve intervention effect and what does the trajectory look like?; Does transition state differ at different dosages?) of participants responding following the introduction of the intervention. Both advantages assist in building efficacious interventions that use data to determine intervention components, their order of application if a packaged intervention and the duration of intervention.

Use of single-case experimental design to develop interventions

Often, researchers develop intervention packages based on clinical experiences and test these multicomponent interventions using a conventional group design study where the intervention package is compared to treatment as usual or placebo, which allows researchers to answer questions only about effectiveness of the packaged intervention when compared to the treatment as usual or placebo conditions and not the relative contributions of individual intervention components or treatment effects for individuals. Researchers can use SCED to dismantle and test the individual components (*e.g.*, dropout analysis like A-XY-Y-XY-Y).[45] Building interventions using an add-in approach with SCED/N-of-1 can build the necessary and sufficient intervention components or dosage to obtain the desired outcome (*e.g.*, add-in approach like A-X-XY-XYZ-XYXYZ).[45] Similarly, researchers can test the effects of parametric dosing of rehabilitation interventions (*e.g.*, incremental increases in dose of intervention to determine opti-

mal dose of intervention regarding timing of intervention, duration and intensity of intervention sessions including number of learning trials within a single session, different frequencies or intensities of physical movement, and different levels of planned participant engagement).[2] Used in these ways, SCEDs can inform an iterative approach to intervention development.

Additionally, researchers can employ a SCED/N-of-1 RCT to test preliminary effectiveness of the intervention or intervention components, thus building towards a fully powered RCT while working out measurement and protocol details. Krasny-Pacini suggested SCEDs on front or back end of RCTs as relevant options; the latter option allows the researcher to intervene with non-responding participants within the intervention group.[3]

Effectiveness trials

N-of-1 RCTs have been defined as effectiveness or comparative effectiveness trials.[5] Individual variation in participant presentation, progression of disability, and outcomes with intervention are present regardless of research methodology, but SCED can build in responsive modification and identification of what works for individuals, including those individuals for whom intervention outcomes are suboptimal. This experimental approach uniquely maps onto clinical decision making in rehabilitation for individual participants, making SCEDs attractive for developing optimal dosing of intervention components. With advances in placebo trials, N-of-1 trials, and randomization designs, researchers implementing SCEDs are not sacrificing experimental control or internal validity and the statistical conclusion validity of randomization SCEDs (*e.g.*, case and start point randomization) is equal to an RCT, albeit with fewer participants.

Building evidence across research teams and studies

Single-case designs allow for efficient (time and cost) development of effective interventions in translational and applied settings, which researchers can then deploy in multisite RCTs. Although SCED requires careful planning and intensive effort that can exceed efforts required to implement an underpowered RCT or observational study, the decreased recruitment and funding requirements can advance a scholarly agenda more quickly.

Much rehabilitation research has been conducted in high-income countries and focused primarily on white, educated, industrialized, rich, and democratic societies. The World Health Organization's Rehabilitation 2030 Initiative encourages development of comprehensive rehabilitation service delivery globally, expansion of research capacity, and creation of evidence through rigorous experimentation.[46, 47] Single-case research studies offer an efficient way to examine the effects of culturally adapted interventions within low- and middle-income countries and with racially, ethnically, and linguistically diverse participants, including those in rural and remote areas.[48]

Research synthesis single-case experimental design studies

Any rigorously designed SCED offers a promising trial of effectiveness, but further replication across researchers, participants and contexts is required to make valid decisions about evidence for interventions and to guide the development of policies regarding use of interventions.[49] Historically, in psychology and education, researchers used the external validity minimum guidelines of three research teams, five studies, and 20 participants,[50] and rehabilitation researchers recently suggested adoption of this approach.[7] Other researchers have discarded this approach in preference for formal meta-analysis of SCEDs due to problems of adjusting for null effects.[10] Researchers have suggested the meta-analysis of N-of-1 studies in situations with few RCTs to help understand variation in participant responses within RCTs.[24, 25] Several examples exist in the meta-analysis of N-of-1 studies, including examining interventions addressing fibromyalgia,[51] and reducing muscle stiffness for participants with non-dystropic myotonia.[52] Rehabilitation researchers will need to arrive at a consensus on how many study replications are necessary to count an intervention as an evidence-based approach or adopt meta-analysis guidelines for comparison of evidence of SCEDs and N-of-1 trials and RCTs in the same meta-analysis. The latter approach offers more flexibility in inclusion of well-designed SCEDs alongside conventional group design studies. Still, it requires the use of revised approaches adjusting for the small sample size of most SCED studies. Importantly, these determinations about what constitutes an evidence-based practice will likely vary depending on the condition and rehabilitation context as well as patient preference and cost-benefit ratio.[53]

Researchers have used multiple models for analyzing such combined research design evidence into a single effect size estimate, for example, through calculation of a design comparable effect size that is an altered (for SCED) standardized mean difference statistic (D-CES),[54] multilevel modeling,[55] and hierarchical linear modeling.[56] Applications of meta-analytic strategies for comparing con-

ventional group designs to SCEDs are emergent with open source software available.[55, 57] Researchers can plan meta-analysis using a freely available methods guide for effect size estimation and synthesis of single-case designs.[58]

Objective 4: conduct and reporting of single-case experimental design studies to improve utility for rehabilitation end-users

Single-case research designs have roots in applied behavior analysis. Since its inception as a field, they have upheld a standard of developing interventions and practices that address issues considered important by society and to ensure the social validity of interventions and practices.[59] Contemporary standards for SCEDs include these core traditions of selecting socially important outcomes, reporting social validity measures to ascertain participant perceptions regarding the intervention goals, procedures, and outcomes, and the definition of a causal relation as active manipulation of the independent variable that produces *clinically significant change* in the outcome variable when assessed using visual analysis. To determine the perceptions of participants, interventionists and other interested parties regarding the goals, procedures and outcomes of the intervention, questionnaires, rating scales, and interviews are used in SCED studies.[60] These measures have generally been administered once post-intervention, but best practices include administering in a pre-test post-test manner.[61] SCED researchers who ask participants about the goodness of fit between the planned study and their

TABLE II.—*Suggestions for conduct and reporting of single-case experimental design studies to improve utility for rehabilitation end-users.*

	Challenge	Suggestion(s)	Example
I.	Patients		
	Lengthy baseline assessment or random assignment increases attrition risk, participant frustration	Select designs aligned with clinical context and patient comfort	MBD study across 4 participants would require lengthy baseline for later tiers. Instead, select random start point MBD to communicate expected intervention start time, or consider another design (*e.g.*, MPD)
	Patients can struggle to make sense of peer-reviewed publications detailing intervention study findings if they can even access these documents	Clinical corners in journals, presentations at patient friendly conferences, meta-analysis including SCED studies and subsequent creation and dissemination of patient-friendly practice briefs summarizing and explaining findings through professional organizations, providers, and social media	Ease accessibility of correct interpretation of SCED studies and syntheses using simple language and visuals
	Patients have reasonable expectations around not delaying intervention for purpose of baseline assessment or randomization to no treatment condition	Participants serve as their own control in SCED studies; use designs reducing delay to intervention or allow participants to know when they will enter intervention	If intervention conditions have reversible effects, can use ATD. If we plan to stepwise increase or decrease a behavior, we can use changing criterion design; we can use range bound randomization designs
II.	Clinician scientists		
	Speed of clinical work misaligned with timeline of research endeavors.	Consider designs mapping onto clinical practice	Non-concurrent MBD across participants as they are consented into the study
	Data collection is time intensive	Consider permanent product data produced during clinical sessions; designs with less data collection where relevant	Build treatment fidelity data collection into clinical procedures. Use of MPD or ATD
	Rehabilitation by nature is multidisciplinary and comprehensive which makes it difficult to isolate independent variable	Develop conceptual logic model to identify all intervention components *a priori* and actively manipulate single independent variables at a time (can be a package) while holding steady all other core variables	Client with ABI might receive OT, PT, SLP, physician care. An intervention addressing PT would require minimally documenting other services and optimally holding constant other therapy session approach and dosage
	Graduate and post-graduate training in SCED design and analysis for rehabilitation professionals are limited	Intensive post-graduate professional development trainings tied to professional organization conferences	The International Collaborative Network for N-of-1 Trials and Single-Case Designs https://www.nof1sced.org
III.	Policy makers		
	Need for transparent reporting of study conduct and findings using field agreed-upon reporting guidelines	Peer-reviewed journals could require authors submitting SCED studies in rehabilitation to complete a reporting checklist	CONSORT N-of-1 Extension

treatment goals, lifestyle, and values before baseline assessment can pivot study goals, procedures and targeted treatment outcomes to incorporate participant voice. The ultimate goal of social validity assessment in SCED is to develop interventions producing clinically significant change using procedures and treatment targets that patients will find feasible, acceptable, and effective.[62] One way of determining whether the effects of the interventions of interest had clinically relevant results for individual patients is the use of the Global Rating Scale (GRS). Based on this anchor tool, the minimal important change can be calculated for the target outcome (dependent variable) based on the patient's perspective.[63] This type of analysis can be easily implemented in SCED.

Benefits of a response-guided approach to SCED include an individualized approach with outcome evaluation through data-based decision-making for each case. The unique contribution of SCED is the *continuous* gathering and assessment through visual analysis of an individual's outcome data to guide next steps in intervention. Using this approach, researchers gather frequent and repeated outcome data each time the intervention is implemented on small increments of progress towards the terminal outcome (*e.g.*, as illustrated in Figure 2, distance traveled during three times weekly 6-minute walk test for adults with Parkinson's Disease). By using proximal, repeated measures in addition to global standardized measures to measure distal outcomes, the clinician increases the opportunities to change the course of treatment to benefit the patient quickly and the patient benefits from seeing evidence of concrete progress.

As with any intervention research design, it can be challenging to make the results of SCED studies immediately useful for consumers; however, the publication of line graphs allows for consumer access to the raw data and individual participant response. Potential remedies to the challenges inherent in conducting SCED studies in rehabilitation contexts and reporting with replicable precision are summarized in Table II.

Concluding remarks

At the heart of applied rehabilitation research is a focus on developing interventions with a clinically meaningful impact on the functioning of individuals. While no intervention research design is without relative advantages and disadvantages, the SCEDs we have described model a rigorous and systematic approach to interventions utilizing ongoing progress monitoring and adaptation of intervention procedures to obtain optimal intervention outcomes. Rehabilitation researchers can use SCEDs to develop interventions, implement effectiveness trials, build and test effective interventions for understudied populations who experience less than optimal clinical outcomes, build scientific generality of evidence across research teams and contexts, and improve the external validity of studies by more often incorporating patient voice through social validity assessment.

References

1. Armijo-Olivo S, Mohamad N, Sobral de Oliveira-Souza AI, de Castro-Carletti EM, Ballenberger N, Fuentes J. Performance, detection, contamination, compliance, and cointervention biases in rehabilitation research: what are they and how can they affect the results of randomized controlled trials? Basic information for junior researchers and clinicians. Am J Phys Med Rehabil 2022;101:864–78.

2. Dallery J, Raiff BR. Optimizing behavioral health interventions with single-case designs: from development to dissemination. Transl Behav Med 2014;4:290–303.

3. Krasny-Pacini A. Single-case experimental designs for child neurological rehabilitation and developmental disability research. Dev Med Child Neurol 2023;65:611–24.

4. Krasny-Pacini A, Evans J. Single-case experimental designs to assess intervention effectiveness in rehabilitation: A practical guide. Ann Phys Rehabil Med 2018;61:164–79.

5. Selker HP, Cohen T, D'Agostino RB, Dere WH, Ghaemi SN, Honig PK, *et al.* A useful and sustainable role for N-of-1 trials in the healthcare ecosystem. Clin Pharmacol Ther 2022;112:224–32.

6. Kazdin AE. Single-case research designs. Third edition. Oxford University Press; 2021.

7. Yang L, Armijo-Olivo S, Gross DP. Single-case experimental design in rehabilitation: basic concepts, advantages, and challenges. Am J Phys Med Rehabil 2023;102:454–60.

8. Kratochwill TR, Horner RH, Levin JR, Machalicek W, Ferron J, Johnson A. Single-case intervention research design standards: additional proposed upgrades and future directions. J Sch Psychol 2023;97:192–216.

9. Kratochwill TR, Levin JR. Enhancing the scientific credibility of single-case intervention research: randomization to the rescue. Psychol Methods 2010;15:124–44.

10. Ledford JR, Lambert JM, Pustejovsky JE. Single-case-design research in special education: next-generation guidelines and considerations. Except Child 2023;89:379–96.

11. Ledford JR, Lambert JM, Barton EE, *et al.* The evidence base for interventions for individuals with ASD: A call to improve practice conceptualization and synthesis. Focus Autism Other Dev Disabl 2021;36:135–47.

12. Maggin DM, Barton E, Reichow B, *et al.* Commentary on the What Works Clearinghouse standards and procedures handbook (v. 4.1) for the review of single-case research. Remedial Spec Educ 2021;1–13.

13. What Works Clearinghouse. Procedures and standards handbook (version 5.0), 2022 [Internet]. Available from: https://ies.ed.gov/ncee/wwc/handbooks [cited 2024, Sep 19].

14. Tervo RC, Estrem TL, Bryson-Brockmann W, Symons FJ. Single-case experimental designs: applications in developmental-behavioral pediatrics. J Dev Behav Pediatr 2003;24:438–48.

15. Podsiadlo D, Richardson S. The timed "Up & Go": a test of basic functional mobility for frail elderly persons. J Am Geriatr Soc 1991;39:142–8.

16. Gast DL, Lloyd BP, Ledford JR. Multiple baseline and multiple probe

designs. In: Ledford JR, Gast DL, editors. Single Case Research Methodology Applications in Special Education and Behavioral Sciences. Third edition. Routledge; 2018.

17. Slocum TA, Pinkelman SE, Joslyn PR, Nichols B. Threats to internal validity in multiple-baseline design variations. Perspect Behav Sci 2022;45:619–38.

18. Horner RH, Machalicek W. Honoring uncontrolled events: commentary on Slocum et al. Perspect Behav Sci 2022;45:639–45.

19. Morin KL, Lindström ER, Kratochwill TR. Nonconcurrent multiple-baseline and multiple-probe designs in special education: A systematic review of current practice and future directions. Except Child 2024;90:126–47.

20. Skinner CH, McClurg V, Crewdson M. Alternating treatments designs: interpretation challenges and design solutions for validating and comparing interventions. Psychol Sch 2021;1–20.

21. Klein LA, Houlihan D, Vincent JL, Panahon CJ. Best Practices in Utilizing the Changing Criterion Design. Behav Anal Pract 2015;10:52–61.

22. Frampton SE, Guinness KE, Axe JB. The parallel treatments design: A systematic review. Behav Interv 2021;1–21.

23. Kirby MS, Spencer TD, Ferron J. How to Be RAD: Repeated Acquisition Design Features that Enhance Internal and External Validity. Perspect Behav Sci 2021;44:389–416.

24. Mirza RD, Punja S, Vohra S, Guyatt G. The history and development of N-of-1 trials. J R Soc Med 2017;110:330–40.

25. Punja S, Schmid CH, Hartling L, Urichuk L, Nikles CJ, Vohra S. To meta-analyze or not to meta-analyze? A combined meta-analysis of N-of-1 trial data with RCT data on amphetamines and methylphenidate for pediatric ADHD. J Clin Epidemiol 2016;76:76–81.

26. Tanious R, Onghena P. A systematic review of applied single-case research published between 2016 and 2018: study designs, randomization, data aspects, and data analysis. Behav Res Methods 2021;53:1371–84.

27. Moeyaert M, Akhmedjanova D. Effect size estimation for combined single-case experimental designs. Evid Based Commun Assess Interv 2020;14:28–51.

28. Negrini S, Chan L, Ferriero G, Frontera WR, Heinemann A. Current evidence from the randomized controlled trials rehabilitation checklist (RCTRACK) reporting guideline project. Am J Phys Med Rehabil 2021;100:2–4.

29. Vohra S, Shamseer L, Sampson M, Bukutu C, Schmid CH, Tate R, *et al.*; CENT Group. CONSORT extension for reporting N-of-1 trials (CENT) 2015 Statement. BMJ 2015;350:h1738.

30. Tate RL, Perdices M, Rosenkoetter U, Shadish W, Vohra S, Barlow DH, *et al.* The Single-Case Reporting Guideline In BEhavioural Interventions (SCRIBE) 2016 Statement. Phys Ther 2016;96:e1–10.

31. Shadish WR, Zelinsky NA, Vevea JL, Kratochwill TR. A survey of publication practices of single-case design researchers when treatments have small or large effects. J Appl Behav Anal 2016;49:656–73.

32. Tincani M, Travers J. Replication research, publication bias, and applied behavior analysis. Perspect Behav Sci 2019;42:59–75.

33. Cook BG, Fleming JI, Hart SA, Lane KL, Therrien WJ, van Dijk W, *et al.* A how-to guide for open-science practices in special education research. Remedial Spec Educ 2022a;43:270–80.

34. Cook BG, Johnson AH, Maggin DM. Open science and single-case design research. Remedial Spec Educ 2021;43:359–69.

35. Johnson AH, Cook BG. Preregistration in single-case design research. Except Child 2019;86:95–112.

36. Sidman M. Tactics of Scientific Research: Evaluating Experimental Data in Psychology. New York, NY: Basic Books; 1960.

37. Levin JR, Ferron JM, Gafurov BS. Additional comparisons of randomization-test procedures for single-case multiple-baseline designs: alternative effect types. J Sch Psychol 2017;63:13–34.

38. Levin JR, Ferron JM, Gafurov BS. An improved two independent-samples randomization test for single-case AB-type intervention designs: A 20-year journey. J Mod Appl Stat Methods 2019;18:1–20.

39. Levin JR, Kratochwill TR. Randomized single-case intervention designs and analyses for health sciences researchers: A versatile clinical trials companion. Ther Innov Regul Sci 2021;55:755–64.

40. Levin JR, Ferron JM, Gafurov BS. Novel randomization tests for two-sample multiple-baseline designs. J Educ Students Placed Risk 2022;27:353–66.

41. Bulté I, Onghena P. Randomization tests for multiple-baseline designs: an extension of the SCRT-R package. Behav Res Methods 2009;41:477–85.

42. Byun TM, Hitchcock ER, Ferron J. Masked visual analysis: minimizing Type I error in visually guided single-case design for communication disorders. J Speech Lang Hear Res 2017;60:1455–66.

43. Ferron JM, Joo SH, Levin JR. A Monte Carlo evaluation of masked visual analysis in response-guided versus fixed-criteria multiple-baseline designs. J Appl Behav Anal 2017;50:701–16.

44. OCEBM Levels of Evidence Working Group. The Oxford Levels of Evidence 2. Oxford Centre for Evidence-Based Medicine; 2024 [Internet]. Available from: https://www.cebm.ox.ac.uk/resources/levels-of-evidence/ocebm-levels-of-evidence [cited 2024, Sep 19].

45. Ward□Horner J, Sturmey P. Component analysis of behavior skills training in functional analysis. Behav Interv 2012;27:75–92.

46. World Health Organization Rehabilitation in Health Systems. 2017 [Internet]. Available from: https://apps.who.int/iris/handle/10665/254506 [cited 2024, Sep 19].

47. Henrich J, Heine SJ, Norenzayan A. The weirdest people in the world? Behav Brain Sci 2010;33:61–83, discussion 83–135.

48. Bice-Urbach BJ, Kratochwill TR. Teleconsultation: the use of technology to improve evidence-based practices in rural communities. J Sch Psychol 2016;56:27–43.

49. Kazdin AE. Single-case experimental designs: Characteristics, changes, and challenges. J Exp Anal Behav 2021;115:56–85.

50. Horner RH, Carr EG, Halle J. The use of single-subject research to identify evidence-based practice in special education. Except Child 2005;71:165–79.

51. Zucker DR, Ruthazer R, Schmid CH, Feuer JM, Fischer PA, Kieval RI, *et al.* Lessons learned combining N-of-1 trials to assess fibromyalgia therapies. J Rheumatol 2006;33:2069–77.

52. Stunnenberg BC, Raaphorst J, Groenewoud HM, Statland JM, Griggs RC, Woertman W, *et al.* Effect of Mexiletine on muscle stiffness in patients with nondystrophic myotonia evaluated using aggregated N-of-1 trials. JAMA 2018;320:2344–53.

53. Dijkers MP, Murphy SL, Krellman J. Evidence-based practice for rehabilitation professionals: concepts and controversies. Arch Phys Med Rehabil 2012;93(Suppl):S164–76.

54. Shadish WR, Hedges LV, Pustejovsky JE. Analysis and meta-analysis of single-case designs with a standardized mean difference statistic: a primer and applications. J Sch Psychol 2014;52:123–47.

55. Declercq L, Cools W, Beretvas SN, Moeyaert M, Ferron JM, Van den Noortgate W. MultiSCED: A tool for (meta-)analyzing single-case experimental data with multilevel modeling. Behav Res Methods 2020;52:177–92.

56. Moeyaert M, Yang P, Xu X. The power to explain variability in intervention effectiveness in single-case research using hierarchical linear modeling. Perspect Behav Sci 2021;45:13–35.

57. Pustejovsky JE, Chen M, Grekov P. Single-case effect size calculator (Version 0.7.2.9999) [Web application]; 2023 [Internet]. Available from: https://jepusto.shinyapps.io/SCD-effect-sizes/ [cited 2024, Sep 19].

58. Ferron JM, Kirby M, Lippen L. Effect Size Estimation and Synthesis of Single-Case Designs: A Methods Guide. Institute of Education Sciences; 2023 [Internet]. Available from: https://jepusto.github.io/SCD-Methods-Guide/ [cited 2024, Sep 19].

59. Baer DM, Wolf MM, Risley TR. Some current dimensions of applied behavior analysis. J Appl Behav Anal 1968;1:91–7.

60. Snodgrass MR, Chung MY, Kretzer JM. Rigorous assessment of social validity: A scoping review of a 40-year conversation. Remedial Spec Educ 2022;43:114–30.

61. Snodgrass MR, Chung MY, Meadan H, Halle JW. Social validity in single-case research: A systematic literature review of prevalence and application. Res Dev Disabil 2018;74:160–73.

62. Schwartz IS, Baer DM. Social validity assessments: is current practice state of the art? J Appl Behav Anal 1991;24:189–204.

63. Armijo-Olivo S, de Castro-Carletti EM, Calixtre LB, de Oliveira-Souza AI, Mohamad N, Fuentes J. Understanding Clinical Significance in Rehabilitation: A Primer for Researchers and Clinicians. Am J Phys Med Rehabil 2022;101:64–77.

Conflicts of interest
The authors certify that there is no conflict of interest with any financial organization regarding the material discussed in the manuscript.

Funding
This study was supported and funded by the Italian Ministry of Health - Ricerca Corrente. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Authors' contributions
Wendy Machalicek has given substantial contributions to the conception and the design of the manuscript, Douglas P. Gross, Susan Armijo, Giorgio Ferriero, Carlotte Kiekens, Rachelle Martin, Margaret Walshe, Stefano Negrini to the acquisition, analysis and interpretation of the data. All authors have participated in drafting and revisiting the manuscript critically. All authors read and approved the final version of the manuscript.

Group author members
Chiara ARIENTI (Milan, Italy); Irene BATTEL (Milan, Italy); Maria G. CERAVOLO (Ancona, Italy); Christopher COLVIN (Charlottesville, VA, USA); Claudio CORDANI (Milan, Italy); Pierre CÔTÉ (Toronto, Canada); Anne CUSICK (Sydney, Australia); Bernard DAN (Brussels, Belgium); Matteo J. DEL FURIA (Milan, Italy); Susanna EVERY-PALMER (Wellington, New Zealand); Peter FEYS (Hasselt, Belgium); Christoph GUTENBRUNNER (Hannover, Germany); Carsten B. JUHL (Copenhagen, Denmark); Stefano G. LAZZARINI (Milan, Italy); William M. LEVACK (Wellington, New Zealand); Federico MERLO (Milan, Italy); Thorsten MEYER-FEIL (Halle, Germany); Luca MIRANDA (Milan, Italy); Bianca MOSCONI (Milan, Italy); Randolph NUDO (Kansas City, KS, USA); Aydan ORAL (Istanbul, Türkiye); Cecile RØE (Oslo, Norway); Heather SHEARER (Toronto, ON, Canada); Jessica WONG (London, ON, Canada).

History
Article first published online: October 7, 2024. - Manuscript accepted: September 19, 2024. - Manuscript received: August 22, 2024.