

Big Evidence with Small Samples: A Methodological Perspective on N-of-1 (Single Case Experimental Designs) Based on Medical Studies

 Mehmet Tahir Huyut¹,  Orhan Aydın²

¹Department of Biostatistics and Medical Informatics, Erzincan Binali Yıldırım University Faculty of Medicine, Erzincan, Türkiye

²Department of Special Education, Division of Autism Spectrum Disorder, Ankara University Faculty of Educational Sciences, Ankara, Türkiye

Cite this article as: Huyut MT, Aydın O. Big evidence with small samples: a methodological perspective on N-of-1 (single case experimental designs) based on medical studies. *Arch Basic Clin Res*. 2026;8(3):288-298

ORCID IDs of the authors: M.T.H. 0000-0002-2564-991X, O.A. 0000-0002-9695-2414.

ABSTRACT

N-of-1 or single-case experimental designs (SCEDs) are gaining renewed importance as robust methodological approaches for assessing individualized treatment response in modern medicine. This overview aims to present the methodological foundation of SCEDs, its applications in medical research, and its position within evidence-based medicine from a disciplinary perspective. This study evaluates the advantages, limitations, and emerging synthetic approaches to the method through a review of the current literature and an examination of selected clinical N-of-1 applications. Findings indicate that SCEDs provide high internal validity, particularly when treatment response varies among individuals, and allow for strong statistical inferences even with small samples. However, methodological limitations, including limited external validity, measurement burden, and carry-over effects, underscore the importance of combining multiple N-of-1 trials using hierarchical models. In recent years, the literature has emphasized that the “large evidence with small samples” approach is preferable from both ethical and methodological perspectives, forming the basis for personalized treatment decisions. This review provides a comprehensive framework for applying SCEDs in medical research and suggests that multiple N-of-1 syntheses will be the primary driver of population-level inference. Consequently, SCEDs can be considered a powerful component of modern medicine because they prioritize individual response over the average patient in clinical decision-making.

Keywords: N-of-1 trials, single-case experimental designs, evidence-based medicine, personalized medicine, small-sample Inference, individual treatment response

INTRODUCTION

1. Epistemological Foundations: The Crisis of the “Average Patient” and the Need for Personalized Evidence

The gold standard of *modern medicine*, evidence-based medicine (EBM), primarily relies on randomized controlled trials (RCTs) that utilize a nomothetic approach, aiming to discover general laws applicable to populations.^{1,2} However, this “average patient” paradigm creates an epistemological gap when applied to the individual in clinical practice.³ While RCTs provide high external validity for populations, they often

mask inter-individual variability, leading to the “tyranny of the average”.^{4,5} In medical research, the tension between the nomothetic (population-based) and idiographic (individual-based) approaches has sparked significant criticism regarding the clinical utility of aggregate data.^{6,7}

Traditional EBM often struggles to address heterogeneous treatment effects, in which a medication that benefits the majority may be ineffective or even harmful to an individual.⁸⁻¹⁰ This limitation necessitates a shift toward a more nuanced understanding of evidence, one that prioritizes the unique



Corresponding author: Mehmet Tahir Huyut, **E-mail:** mehmettahirhuyut@gmail.com

Received: January 13, 2026

Revision Requested: February 9, 2026

Last Revision Received:

Accepted: July 6, 2026

Publication Date: October 28, 2026



Copyright© 2026 The Author(s). Published by Galenos Publishing House on behalf of Erzincan Binali Yıldırım University. This is an open access article under the Creative Commons AttributionNonCommercial 4.0 International (CC BY-NC 4.0) License.

biological and clinical context of the single patient.¹¹ The search for individual-specific evidence is not a rejection of EBM, but rather an evolution toward a more precise and person-centered methodology.¹²

2. The Bridging Role of SCEDs: Harmonizing Clinical Practice and Scientific Rigor

Single-case experimental designs (SCEDs), or N-of-1 trials, emerge as a critical bridge between the rigor of experimental research and the reality of clinical practice.^{13,14} Unlike anecdotal case reports, SCEDs utilize systematic manipulation, repeated measurements, and baseline comparisons to establish causal relationships within a single unit.^{15,16} This methodology allows the clinician to act as a researcher, transforming routine care into a disciplined scientific inquiry.¹⁷

By integrating the idiographic focus on the individual with the nomothetic requirement for rigorous control, SCEDs resolve the historical tension in medical knowledge acquisition.^{18,19} They provide a structured framework for assessing treatment efficacy in conditions in which traditional RCTs are impractical, such as rare diseases or highly personalized interventions.²⁰ Consequently, SCEDs function as a methodological cornerstone, ensuring that the evidence used in clinical decision-making is both scientifically robust and directly relevant to the patient being treated.²¹

MAIN POINTS

- N-of-1 (single-case experimental) designs provide a rigorous methodological framework for evaluating individualized treatment response. These designs offer high internal validity in situations where inter-individual heterogeneity limits the applicability of group-based evidence.
- Single-case experimental designs (SCEDs) act as a methodological bridge between nomothetic (population level) and idiographic (individual level) evidence. These approaches allow for causal inference at the patient level while remaining compatible with aggregate and hierarchical synthesis.
- Through repeated measurements and systematic manipulation of experimental phases, SCEDs enable the detection of clinically meaningful change at the individual level. This structure allows robust effect estimation even in small samples, supporting the concept of “big evidence from small samples.”
- Methodological challenges such as carry-over effects, autocorrelation, and limited external validity represent key analytical concerns in N-of-1 studies. These challenges can be effectively addressed through randomization, washout periods, time-series modeling, and Bayesian or meta-analytic aggregation of multiple single-case trials.
- SCEDs offer a practical and ethically sound approach for implementing personalized medicine in clinical research and practice. This is particularly relevant for chronic, heterogeneous, or treatment-resistant conditions, where individual response patterns are more informative than population averages.

3. The Role of SCEDs Research Today

SCEDs research has gained renewed importance in recent years, particularly in fields such as personalized medicine, pharmacogenetics, rehabilitation, behavioral medicine, and neuropsychology. The current research paradigm has moved beyond the classical RCT, which emphasizes average group effects, to understanding the variability of individual responses.^{4,22} In this context, SCEDs are considered a “self-randomized experiment”.²¹

In personalized medicine and pharmacogenetics, treatment response varies depending on genetic, biochemical, and environmental factors.²² This variability makes traditional RCT results based on group averages insufficient for clinical decision-making. The SCEDs approach allows direct testing of treatment efficacy in a single individual, thereby contributing to personalized treatment optimization. For example, the use of N-of-1 protocols is recommended to optimize drug responses in diseases with heterogeneous courses, such as chronic pain, hypertension, and major depression.²⁰⁻²² Such studies allow for the systematic testing of variables such as patient-specific doses, drug combinations, or treatment duration. Recently, N-of-1 studies have begun to be addressed not only at the individual level but also at the aggregate level. “Meta-analytic N-of-1 syntheses,” which combine studies with multiple SCEDs, have made it possible to produce generalizable findings at the group level by combining individual evidence.^{7,9,14,15,23} This approach is considered a methodological revolution, with the potential to simultaneously examine both individual efficacy and within-population variation. Particularly with the development of Bayesian hierarchical approaches in statistical modeling, aggregating studies with SCEDs has become both more powerful and more reliable.²⁴

Furthermore, the development of digital health technologies and artificial intelligence-based data analysis methods has made it possible to collect and analyze data at the individual level. Owing to mobile health applications, wearable sensors, and remote monitoring systems, the applicability of N-of-1 protocols in daily clinical practice has significantly increased.²⁵ Artificial intelligence algorithms can be used to automatically detect trends, predict efficacy, and generate personalized treatment recommendations from individual behavioral and physiological data streams.²⁶ These developments enable SCEDs to function not only as a research methodology but also as a clinical decision support system.

As a result, SCEDs have become a crucial component of the contemporary scientific paradigm, which is centered on individual heterogeneity. As a complement to classical group studies, SCEDs offer a new epistemological framework based on the principle of “evidence from the individual, not the population.” This approach is shaping the future of micro-level evidence-based practices in both clinical science and behavioral research.

4. Experimental Control Methods with SCEDs

The primary goal of SCEDs is to demonstrate the causal effect of an independent variable (treatment, medication, behavioral intervention, etc.) on a specific individual. To this end, the

participant serves as an experimental control, thus minimizing confounding effects arising from interindividual variation.¹⁶

Experimental control is achieved through systematically organized phases, ensuring a prediction-verification-replication pattern. The following section addresses design types in SCEDs. The following section addresses the use of randomization techniques in SCEDs, the control of carry-over bias and autocorrelation, and procedures for ensuring measurement reliability.

4.1. Types of SCEDs

The most commonly used methods for providing experimental control in SCEDs are interrupted time-series designs, which are structured around the presence or absence of an independent variable across specific phases. The main designs are as follows:

AB Design

It is the simplest form, comprising a baseline phase (A) and an intervention phase (B). However, it is limited in its ability to establish causality because it cannot be used to test effects with or without the intervention.^{15,16} Figure 1 shows the basic AB design, which is a quasi-experimental design. It provides only estimated information about the effect of the intervention.

Graphics features:

- X-axis: Sessions/Time/Measurement
- Y-axis: Dependent variable (e.g., pain score)
- Vertical line: Phase change/transitions (A→B)
- *Comment:* The intervention performed in Phase B resulted in a decrease in the patient's pain score.

However, this provides only a predictive *inference* regarding the intervention's effectiveness. To empirically support a strong hypothesis of the intervention's effectiveness, the intervention must be examined through the dimensions of "verification" and "replication."

- *Note:* Data lines and the forecast display line are used here for ease of understanding. They are not typically displayed in studies.

ABAB (Reversal/Withdrawal) Design

The sequence A (baseline), B (intervention), A (withdrawal of the intervention), B (re-intervention) is used to test the intervention effect. If the behavior or outcome variable

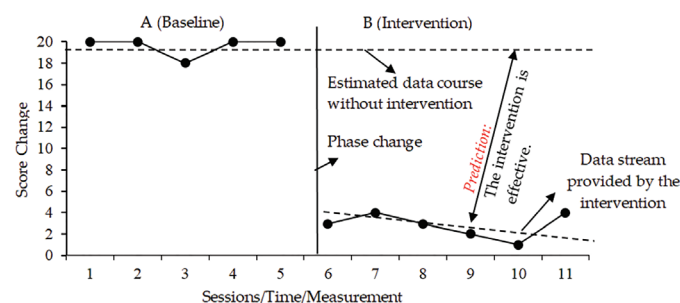


Figure 1. Basic AB design.

increases during the intervention period, returns to its previous level during the withdrawal period, and changes again during the re-intervention, this provides strong evidence for a causal relationship.^{16,27-29} Figure 2 illustrates the ABAB design and provides information on its use as an experimental design for reversible measures (e.g., blood sugar and blood pressure levels). It provides information about the intervention's effect through the dimensions of prediction, verification, and replication, and offers functional evidence of that effect.

Comment: Measured values systematically during the B phases. Since the values worsen in the A phases, where no intervention occurs and only measurements are taken, demonstrating the effectiveness of the intervention through the prediction-confirmation-replication triad establishes a strong causal relationship

Multiple Baseline/Probe Design

The intervention is initiated gradually across different behaviors, settings, or individuals. This provides a powerful alternative experimental control in situations where withdrawal is unethical (e.g., in medical or educational settings).¹⁹ Figure 3 illustrates the multiple-baseline design, which provides an experimental framework applicable to both reversible and non-reversible clinical outcomes. Reversible measures include physiological parameters (e.g., arterial blood pressure, blood glucose, heart rate), biochemical and hormonal indicators (e.g., thyroid hormone levels, inflammatory biomarkers, electrolytes), symptom-based assessments (e.g., pain intensity, anxiety or stress scores, fatigue levels), performance or functional measures (e.g., gait speed, cognitive reaction time, attentional performance), and daily-life indicators (e.g., sleep duration and quality, physical activity levels). The design is also suitable for non-reversible outcomes, such as axonal loss in multiple sclerosis, tumor size progression, manifestations of neurodegenerative diseases, and clinical worsening of congestive heart failure. In this design, the effect of the intervention is evaluated, compared with conditions where no intervention has been implemented, using the dimensions of prediction, confirmation, and replication to seek functional evidence. The key difference between the multiple probe and the multiple baseline designs is that, in the multiple probe design, measurements taken in Phase A are spaced. For example, if measurements are not taken in sessions 6-7 for the second and third graphs in Figure 2, and in session 9 for the third graph, this would constitute a variation of the multiple probe design.

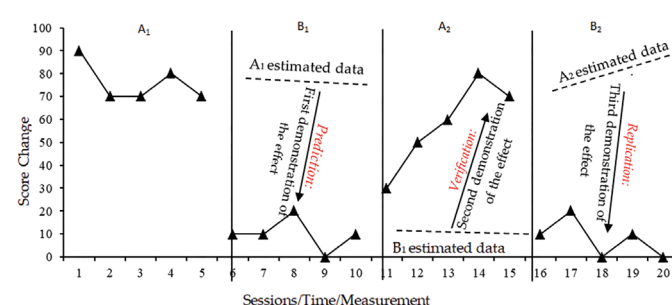


Figure 2. ABAB design.

Comment: In Figure 3, three participants are depicted in a stacked multiple-baseline design in which the intervention is introduced at different time points. In the upper panel, an improvement in the outcome is observed immediately following the onset of the intervention phase, whereas the remaining two participants continue to exhibit low, stable values, as reflected by ongoing A-phase measurements. This pattern provides preliminary prediction-level evidence regarding the potential effectiveness of the intervention. When the intervention is subsequently introduced to the participant in the middle panel, a similar improvement emerges, whereas the participant in the lower panel—who has not yet received the intervention—continues to display unfavorable trends. This provides confirmatory evidence supporting the effect of the intervention. Finally, when the intervention is implemented for the third participant in the lower panel, a consistent improvement is again observed, demonstrating replication of the effect. Collectively, this sequential pattern of prediction, confirmation, and replication provides strong experimental evidence for the intervention's impact on the dependent variable.

ABCBC or Multiple Intervention Reversal Designs

Systematic application of different variants of an intervention (e.g., doses or intensity levels) can be used for dose-response analysis.³⁰ Figure 4 illustrates the ABCBC design and provides information on the experimental design for reversible measurements. B and C are compared as separate interventions. While these interventions may have completely different treatment protocols, B and C may represent two variations in the delivery methods of an intervention. This provides information about which intervention is more ethical through prediction, validation, and replication, and supplies empirical evidence that one intervention is functionally more successful than the other.

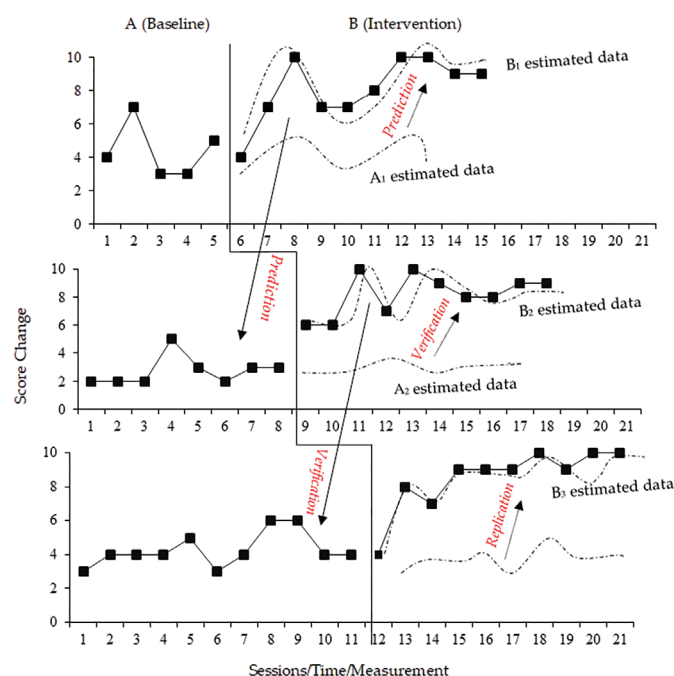


Figure 3. Multiple starting level design.

Alternating Treatment Design (ATD) or Multielement Design

A single-case experimental approach compares two or more interventions (e.g., Drug A and Drug B) through rapid, successive alternations. Randomized sequencing of intervention conditions minimizes carryover effects and allows for a direct comparison of treatment efficacy.⁹ Figure 5 illustrates the structure of the alternating treatment design. In an alternating treatment design, experimental control is established through the systematic divergence of data patterns across conditions as the interventions are presented rapidly in alternating order. At the outset, the researcher predicts that if the interventions truly differ in their effects, the data will begin to separate by condition at the initial measurement points. When such divergence emerges during data collection, the prediction is confirmed. The repeated alternation of conditions (B-C-B-C-...) provides replication by demonstrating the effect of each intervention multiple times across the measurement series. Together, these processes of prediction, confirmation, and replication yield strong experimental control, supporting the conclusion that observed changes in the dependent variable are attributable to the specific intervention delivered.^{16,28,29}

Comment: In Figure 5, interventions B and C are implemented to reduce the desired value. Although baseline values were high, the measured values after interventions B and C were lower than baseline. However, in the initial dataset, intervention C consistently produced lower values than intervention B, indicating greater effectiveness based on predictions in that dataset and confirmed by subsequent validation and replication.

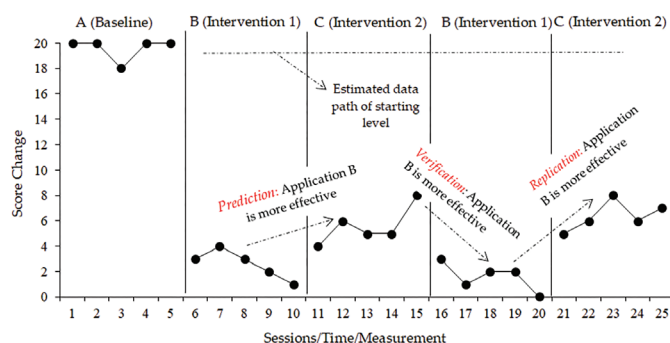


Figure 4. ABCBC design.

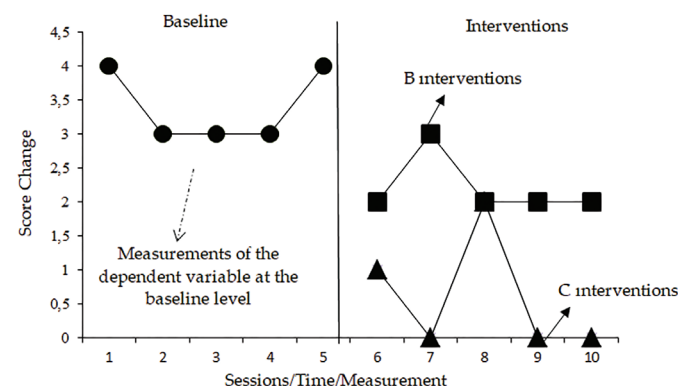


Figure 5. Design of rotational interventions.

4.2. Randomization, Washout, and Carry Over Effects

In modern N-of-1 protocols, randomization and washout periods play a critical role in improving internal validity. Randomizing the order of interventions minimizes expectancy and order effects.²² The washout period between interventions is the expected interval for the physiological or psychological effects of the previous treatment to subside. This is particularly important in pharmacological studies, to control for carry-over effects (the effect of the previous period carrying over into the next period).³¹

4.3. Trend and Autocorrelation Control

SCEDs data are discrete-time series, so issues such as trend (the natural direction of change) and autocorrelation (the dependency between successive observations) are common.^{32,33} This can lead to false-positive results or to an overestimation of the intervention effect.

Therefore, in the analysis:

- Autoregressive integrated moving average (ARIMA) models are used to account for trend and autocorrelation by modeling temporal dependence.³⁴

Alternatively, to determine whether the observed changes are random or not and to assess the statistical significance of the results, randomization tests³⁴ or Bayesian time series models³⁵ are preferred.

4.4. Measurement Reliability and Standardization

A strength of SCEDs is that repeated measurements on the same individual minimize measurement error. However, the following practices are recommended for reliable results:

- *Stable Conditions*: The measurement environment, time period and observer conditions are kept constant.
- *Standard Protocols*: Data collection and application procedures are carried out according to predetermined protocols.

- *Independent observers*: For reliability in observational data, the measurement of the dependent variable is made by at least two independent evaluators.^{29,30}

Application Examples of SCEDs Methods in Medical Research

The ten clinical studies presented in Table 1 were selected using purposive sampling rather than through systematic review. The primary inclusion criteria were: (1) the rigorous application of core SCED design principles (e.g., ABAB/reversal, multiple baseline, or alternating treatments), (2) representation across diverse medical specialties, including oncology, neurology, psychiatry, and chronic disease management and (3) the inclusion of both pharmacological and non-pharmacological interventions. Accordingly, ten illustrative examples of N-of-1 medical studies are summarized in Table 1, with each study described in terms of its objective, participant profile, selected design type, measurement framework (dependent and independent variables), applied intervention, and key outcomes.

These examples were chosen to demonstrate how different SCED architectures can generate high internal validity and support clinical decision-making across a wide range of medical contexts. Collectively, the studies illustrate the versatility of N-of-1 designs, showing their applicability to diverse conditions—from chronic pain and psychiatric disorders to rehabilitation and oncology—and their utility for purposes ranging from optimizing pharmacotherapy to evaluating individualized behavioral interventions. Many of these studies incorporate strategies to strengthen internal validity, such as randomization, washout periods, appropriate SCED selection, and time-series or autocorrelation-adjustment methods (e.g., ARIMA models, randomization tests, or Bayesian approaches).^{2,7} For clinical readers, the primary purpose of presenting these examples is to demonstrate how efficacy evidence can be generated at the individual patient level and how such evidence may inform personalized treatment decisions in real-world practice (Table1).

Table 1. The following tables contain brief, standardized summaries of each of the 10 sample clinical trials (purpose, design, measurements, intervention, outcome, and role of SCEDs in the study)

Example study	Study characteristics	Explanation
Effectiveness of transcranial direct current stimulation in chronic daily headaches: an experimental single case report with a novel protocol Rahimi et al., 2022 ³⁵	Objective	To evaluate the efficacy of tDCS for chronic daily headache at the individual level
	Participants	Adult female patient diagnosed with chronic headache
	Design	ABAB design
	Measures/dependent variables	Daily pain scores (VAS), pain intensity and diary records, frequency of analgesic use
	Intervention/independent variables	tDCS application (10-session protocol; separated by washout periods, 2 mA, 20 min, targeting dorsolateral prefrontal cortex)
	Outcomes	Pain intensity significantly decreased during tDCS periods, with reduced analgesic use
	Role of SCEDs	Allowed direct observation of neuromodulation effects at the individual level, guiding personalized treatment

Table 1. Continued

Example study	Study characteristics	Explanation
N-of-1 trials in cancer drug development. Cancer Discovery	Objective	To examine the effects of combined adaptive immunotherapy and chemotherapy on cancer progression at the individual level
	Participants	Adult male patient with advanced solid tumor
	Design	AB design
	Measures/dependent variables	Tumor size, immune cell profile (CD8 ⁺ , Treg), quality of life, clinical response, immune activation markers
	Intervention/independent variables	Immunotherapy + chemotherapy combination (adaptive dosing and cyclic chemotherapy regimen)
	Outcomes	Tumor regression, increased CD8 ⁺ proportion, and well-tolerated treatment course
	Role of SCEDs	Enabled individual response monitoring and dynamic dose optimization, illustrating a personalized oncology approach
N-of-1 trials comparing a non-steroidal anti-inflammatory drug with paracetamol in osteoarthritis	Objective	To compare the individual efficacy of NSAIDs versus paracetamol for osteoarthritis pain
	Participants	51 patients with osteoarthritis (each conducted as a separate N-of-1 trial)
	Design	Multiple crossover N-of-1 design (similar to alternating treatment design)
	Measures/dependent variables	Daily pain (VAS), morning stiffness, medication preference, pain scores, patient satisfaction
	Intervention/independent variables	NSAID versus paracetamol periods
	Outcomes	No significant difference for most patients, though some showed clear individual preference
	Role of SCEDs	Supported individualized treatment decisions without relying on group averages, providing a robust model for personalized medication selection in clinical practice
Efficacy and safety of amantadine as a treatment for apathy after brain injury: two single-case experimental design studies	Objective	To evaluate the individual effects of amantadine on post-brain injury apathy
	Participants	Two adults with traumatic brain injury
	Design	Multiple baseline across participants design
	Measures/dependent variables	Apathy Evaluation Scale (AES), neuropsychological performance tests
	Intervention/independent variables	Amantadine treatment (100-200 mg/day)
	Outcomes	Both cases showed reduced apathy scores and increased cognitive alertness
	Role of SCEDs	Enabled testing of drug efficacy at the individual level in rare and heterogeneous neurological cases
Personalized (N-of-1) trials for depression: a systematic review	Objective	To compare two different antidepressants in patients with major depressive disorder to identify the most suitable treatment at the individual level
	Participants	20 adults diagnosed with major depression, each followed as a separate N-of-1 trial
	Design	Multiple crossover N-of-1 design (similar to alternating treatments; e.g., Drug A → Washout → Drug B → Washout, repeated as needed)
	Measures/dependent variables	Depression severity (PHQ-9, HAM-D), side effect profile, daily mood tracking, functional assessments
	Intervention/independent variables	Type of antidepressant used
	Outcomes	Significant individual differences observed in 40% of patients; no difference at the group level
	Role of SCEDs	Guided personalized drug selection and informed clinical application of pharmacogenetic approaches

Table 1. Continued

Example study	Study characteristics	Explanation
Treating hypertension in children with N-of-1 trials	Objective	To personalize antihypertensive treatment by evaluating individual responses to three medications using repeated ABPM
	Participants	42 children with primary hypertension; 32 were included in the final analysis
	Design	N-of-1 randomized crossover design in which three antihypertensive medications—lisinopril, amlodipine, and hydrochlorothiazide—were prescribed in randomized order and individual treatment responses were compared
	Measures/dependent variables	Repeated 24-hour ambulatory blood pressure measurements (ABPM) and treatment-related adverse effects
	Intervention/independent variables	Three antihypertensive medications: lisinopril, amlodipine, and hydrochlorothiazide, administered in sequential treatment periods according to the N-of-1 protocol
	Outcomes	Individual treatment preferences varied: 16 preferred lisinopril, 8 amlodipine, and 4 hydrochlorothiazide; 4 remained uncontrolled
	Role of SCEDs	Enabled individualized comparison of treatment responses and identification of the preferred antihypertensive for each patient
Genetically-guided medical nutrition therapy in type 2 diabetes mellitus and pre-diabetes: a series of N-of-1 superiority trials	Objective	To examine how various diets or treatments affect individual glucose responses and develop personalized glycemic control strategies
	Participants	12 adults with type 2 diabetes, monitored using continuous glucose monitoring (CGM)
	Design	Alternating treatments design (e.g., low-carbohydrate diet vs. standard diet; periods randomized)
	Measures/dependent variables	Continuous glucose measurements (CGM), postprandial glucose response, daily mean glucose, glycemic variability, HbA1c (secondary)
	Intervention/independent variables	Two different diet/treatment protocols (e.g., low-carbohydrate vs. standard diet) or individualized lifestyle interventions
	Outcomes	Glucose responses varied widely among individuals; the “optimal” protocol differed for each patient
	Role of SCEDs	Enabled identification of personalized glycemic management strategies and showed that N-of-1 designs can generate individual glucose response profiles without requiring large sample sizes
Status of N-of-1 trials in chronic pain management: a narrative review	Objective	To evaluate the individual effects of mindfulness therapy on chronic pain
	Participants	Adults with chronic low back pain
	Design	ABAB design
	Measures/dependent variables	Pain intensity (VAS/NRS), attention to pain and emotion regulation scales, sleep quality, stress level
	Intervention/independent variables	8-week mindfulness program (mindfulness-based practices: breath awareness, body scan, emotion regulation)
	Outcomes	Pain intensity and stress decreased markedly during mindfulness periods, with functional improvements; effects returned during baseline periods.
	Role of SCEDs	Enabled precise individual assessment of intervention effects, guiding personalized treatment decisions and showing patient-specific impacts of non-pharmacological approaches.
An N-of-1 trial service in clinical practice: testing the effectiveness of stimulants for attention-deficit/hyperactivity disorder	Objective	To evaluate individual response and optimal dosage of stimulant medications in pediatric Attention-Deficit/Hyperactivity Disorder (ADHD)
	Participants	45 children (aged 6–12 years) clinically diagnosed with ADHD
	Design	Multiple crossover randomized double-blind placebo-controlled N-of-1 design (Repeated paired blocks: Drug A vs. Drug B vs. Placebo)
	Measures/dependent variables	Conners’ Rating Scales (teacher and parent daily behavior scores), ADHD symptom severity scores, and side effect logs
	Intervention/independent variables	Methylphenidate vs. dexamphetamine vs. placebo administered in randomized, double-blind treatment cycles
	Outcomes	71% of participants identified a superior medication and dose, optimizing long-term individual treatment
	Role of SCEDs	Enabled precise, patient-specific drug selection in pediatric psychiatry without relying on group averages

Table 1. Continued

Example study	Study characteristics	Explanation
Melatonin for chronic insomnia in Angelman syndrome: a randomized placebo-controlled N-of-1 trial series Braam et al., 2008 ⁴⁴	Objective	To evaluate the individual effects of melatonin on sleep duration and quality
	Participants	Adult with chronic insomnia
	Design	ABAB design (Melatonin → Placebo → Melatonin → Placebo)
	Measures/dependent variables	Sleep onset latency, total sleep time, sleep efficiency, subjective sleep quality (sleep diary, digital monitoring)
	Intervention/independent variables	3 mg melatonin vs. placebo (taken at bedtime)
	Outcomes	Melatonin periods shortened sleep onset latency and improved efficiency, while placebo periods showed clear deterioration
	Role of SCEDs	Effective for identifying individual responses in sleep disorders
SCEDs: Single-case experimental designs; tDCS: Transcranial direct current stimulation; NSAIDs: Non-steroidal anti-inflammatory drugs; Treg: Regulatory T cells; RCTs: Randomized controlled trials; HbA1c: HemoglobinA1c; ACE inhibitor: Angiotensin-converting enzyme inhibitor		

DISCUSSION

Overall Review—SCEDs: The Individualized Dimension of Evidence-Based Medicine

SCEDs (N-of-1) represent the individualized dimension of EBM. While the classic application of EBM provides decision support based on population averages, SCEDs personalize clinical choices by providing control at the individual level.^{2,9} Therefore, SCEDs stand out as a paradigm that pushes the boundaries of the “average patient” idea, producing first-order evidence of individual responses and preferences in medical decision-making.

Advantages of SCEDs

- *Superiority in Internal Validity:* SCEDs allow each participant to serve as their own control, enabling within-subject comparisons that minimize the influence of confounding variables and strengthen causal inference.^{14,16,28,29}
- *Clinical Applicability:* Single-case designs provide direct answers to the patient-level question of “which treatment is effective?”, thereby contributing to individualized clinical decision-making. This approach offers substantial value for treatment selection, dose optimization, and the assessment of treatment tolerability.⁴⁵
- *Ethical Advantages:* Particularly when alternative interventions are being evaluated, single-case designs help ensure that patients are not exposed to an ineffective treatment for extended periods. In most SCEDs, all participants receive one or more forms of the intervention over time, which enhances ethical acceptability.¹⁵
- *Resource Efficiency (Big Evidence from Small Samples):* Single-case designs have the potential to reduce the financial and time-related burdens associated with large RCTs. With proper planning and sufficient repetitions, even small samples can yield strong internal validity.

Limitations of SCEDs

External Validity Limitation: A result valid for a single case cannot automatically be generalized to the population;

therefore, hierarchical/aggregate analyses of multiple N-of-1 studies are required to transfer information from SCEDs to the population level.^{7,45}

Repeated Measurement Burden: Daily or more frequent measurements can increase patient burden, reduce patient compliance, and compromise data quality. The balance between measurement frequency and utility should be carefully established during the design phase.

Carry-over Effects: In pharmacological interventions, the possibility that effects from the previous period will spill over into the next period poses a threat. Washout times and randomized period orders should be controlled, and, if necessary, carryover effects should be addressed using appropriate statistical models.

Measurement Reliability and Objectivity: When subjective measures (e.g., VAS) are used, measurement error and placebo/expectancy effects can confound the results. Independent observers and objective measurements (biomarkers, actigraphy, CGM, etc.) should be included whenever possible.

Analytical Complexity: Constructs such as time-series dependence, autocorrelation, trend, and seasonality should be included in the analysis; otherwise, *P* values may be misleading.

Evidential Value and Hierarchy: The Position of the N-of-1

The placement of SCEDs within the evidence hierarchy is purpose-dependent. RCTs are essential for population-level generalizability, whereas SCEDs offer the strongest evidence for individualized treatment decisions.^{9,46} An optimal contemporary approach is to combine RCTs with multiple N-of-1 studies. Utilizing SCEDs for patient-specific decisions and integrating RCT and SCED evidence in meta-analyses will guide future evidence-based research, policy development, and clinical guidelines.

Statistical Methods and Analysis Recommendations (Guide for Researchers)

Below are principles that researchers can practically apply in the design-analysis process:

Design Phase

- *Design Selection*: Based on the research question, select the appropriate ABAB, multiple baseline/probe design, or alternating intervention designs.³⁴
- *Randomization*: Randomize the order of phase or intervention delivery to reduce the risk of a carryover effect.
- *Wash-Out Duration*: Based on the pharmacodynamics of the intervention; inadequate wash-out increases the risk of a carryover effect.
- *Measurement Plan*: Plan for both short-term (daily) and longer-term (weekly/monthly) measurements; minimize measurement burden.
- *Preregistration*: Registration reduces reporting bias and stabilizes the analysis plan.

Analysis Phase

- *Dependency Control*: Use AR(1), ARIMA, or mixed-effects models to assess autocorrelation between measurements. For example, ARIMA effectively models trends and autocorrelation.³³
- *Randomized Testing*: When stage orders are randomized, permutation-based tests offer robust and distribution-independent options.^{35,47}
- *Bayesian Approaches*: Suitable for expressing uncertainty in SCEDs, using prior knowledge, and combining multiple N-of-1 studies with a hierarchical model.²⁴
- *Effect Size and Clinical Significance*: Don't just look at the *P* value; Effect sizes (Performance-criteria based effect size/PCES³², Tau-U⁴⁸, % change, mean stage change⁴⁸, and patient-specific thresholds should be reported.
- *Meta-Analysis*: Hierarchical (mixed/Bayesian) models are preferred when combining multiple N-of-1 studies; this allows the individual response distribution and the population mean to be estimated together.^{7,33}

Ethical, Experimental Guidance and Clinical Decision Tool: A Practical Guide for Practitioners (Step-by-Step Guide)

The following steps provide a practical roadmap for researchers/clinical teams designing a study with SCEDs:

- *Clarify The Research Question*: Treatment assignment, dose optimization, or intervention comparison?
- Suitable candidates include patients with variable drug response or uncertain treatment preferences.
- *Choose and Justify The Design*: ABAB, multiple baseline, alternating interventions, or ABCBC; specify the ethical and practical pros and cons of each.
- *Create The Measurement Set*: Determine the primary and secondary dependent variables, measurement frequency, and objective measures (if possible).

- *Plan Wash-Out and Duration*: Base the wash-out period on pharmacokinetic properties, and ensure a sufficient number of replicates.
- *Randomization and Blinding*: Implement double-blind or at least blinded assessment if applicable.
- *Data Management and Quality Control*: Verify daily records, independent observer assessments, missing data strategies.
- *Predefine The Analysis Plan*: Check for autocorrelation, effect size, randomization test, or Bayesian analysis methods.
- *Confirmed Consent and Ethical Review*: The design, burden, and potential benefits/harms should be clearly explained to the patient.
- *Reporting*: Consolidated Standards of Reporting Trials (CONSORT) guidelines should be followed; transparency regarding design type, randomization, washout, number of measurements, missing data, and analysis decisions should be reported.^{14,36}

To facilitate the translation of SCEDs methodology into routine clinical practice, the following checklist provides a quick reference for clinicians to determine the suitability of an N-of-1 trial for a specific patient or clinical scenario.

When to Consider an N-of-1 (SCEDs) Trial?

- *High Inter-individual Variability*: When standard population-based treatments yield inconsistent results across different patients.
- *Chronic and Stable Conditions*: When the disease course is long-term and symptoms are relatively stable (e.g., hypertension, chronic pain, insomnia, or stable neurodegenerative symptoms).
- *Treatment Optimization*: When seeking the optimal dosage or comparing the efficacy/safety of two similar therapeutic options for a specific individual.
- *Rare Diseases*: When the rarity of the condition precludes the recruitment of a sufficient sample size for a traditional RCT.
- *Assessing Idiosyncratic Responses*: When a patient experiences unusual side effects or benefits that are not documented in large-scale trials.

When are SCEDs Inappropriate (Contraindications)?

- *Acute or Rapidly Progressing Conditions*: If the clinical status changes too rapidly to allow for stable baseline measurements (e.g., acute infections, life-threatening emergencies).
- *Curative Treatments*: If the intervention is intended to provide a permanent cure, as this prevents the "return to baseline" required for withdrawal designs (e.g., surgical procedures, definitive antibiotic courses).
- *Prolonged Washout Periods*: If a drug's pharmacological effects persist for an extended period, making crossover transitions impractical due to carry-over effects.

• *High Measurement/Burden of Care*: If the frequency of data collection or the complexity of the monitoring protocol significantly compromises the patient's quality of life or clinical safety.

Reporting and Publication Standards

Transparent reporting is crucial in single-subject studies. Reporting should include full disclosure of design type, period lengths, washout details, measurement frequency, missing data, and statistical methods. Supplementary materials (raw data, time-series graphs) prepared according to single-subject research standards (Kratochwill et al.³⁴) facilitate peer review.

Future Directions: Holistic Data Analysis, Digital Integration, and Systematic Evidence

While this overview provides a comprehensive methodological framework, it is not a systematic review. Future research should conduct systematic reviews and meta-analyses of existing N-of-1 trials within specific disease categories. Such efforts would be instrumental in synthesizing individual-level evidence into broader clinical guidelines, thereby enhancing the external validity of SCEDs' findings at the population level.

In this context, the future of SCEDs encompasses two parallel directions: (1) translating multiple N-of-1 studies into population-level inferences through hierarchical, Bayesian, or meta-analytic approaches; and (2) integrating real-time data collection and analysis via digital health and artificial intelligence applications. Together, these developments make the goal of "big evidence from small samples" achievable both methodologically and in practice.^{4,26,32}

CONCLUSION

A Clinical and Methodological Call

SCEDs are one of the most practical ways to implement the individualized dimension of EBM. The proposition that "the future of medical science lies in understanding the individual response, not the average patient" is increasingly being proven true not only rhetorically but also in practice.⁴⁷ However, this potential must be used responsibly: Without rigorous design, reliable measurements, appropriate statistical adjustments, and strategies for examining the data holistically, SCEDs can lead to misinterpretations or conclusions with low generalizability. We recommend that researchers design integrated studies that utilize the strong internal validity of single-subject studies while addressing the need for external validity through multiple N-of-1 syntheses and modern statistical approaches.

Footnotes

Author Contributions

Concept Design – M.T.H., Data Collection or Processing – M.T.H., Analysis or Interpretation – M.T.H., O.A., Literature Review – M.T.H., O.A., Writing, Reviewing and Editing – M.T.H.

Funding: The authors declare that this study has received no financial support.

Declaration of Interests: The authors declare that they have no known competing financial or personal relationships that could be viewed as influencing the work reported in this paper.

REFERENCES

1. Sackett DL, Rosenberg WM, Gray JA, Haynes RB, Richardson WS. Evidence based medicine: what it is and what it isn't. *BMJ*. 1996;312(7023):71-72. [\[CrossRef\]](#)
2. Guyatt GH, Keller JL, Jaeschke R, Rosenbloom D, Adachi JD, Newhouse MT. The N-of-1 randomized controlled trial: clinical usefulness. *Ann Intern Med*. 1990;112(4):293-299. [\[CrossRef\]](#)
3. Hamburg MA, Collins FS. The path to personalized medicine. *N Engl J Med*. 2010;363(4):301-304. [\[CrossRef\]](#)
4. Schork NJ. Personalized medicine: time for one-person trials. *Nature*. 2015;520(7549):609-611. [\[CrossRef\]](#)
5. Kent DM, Hayward RA. Limitations of applying summary results of clinical trials to individual patients: the need for risk stratification. *JAMA*. 2007;298(10):1209-1212. [\[CrossRef\]](#)
6. Mirnezami R, Nicholson J, Darzi A. Preparing for precision medicine. *N Engl J Med*. 2012;366(6):489-491. [\[CrossRef\]](#)
7. Zucker DR, Ruthazer R, Schmid CH, et al. Lessons learned combining N-of-1 trials to assess fibromyalgia therapies. *J Rheumatol*. 2006;33(10):2069-77. [\[CrossRef\]](#)
8. Allport GW. *Personality: a Psychological Interpretation*. New York: Henry Holt and Company; 1937. [\[CrossRef\]](#)
9. Aydin O. A description of missing data in single-case experimental designs studies and an evaluation of single imputation methods. *Behav Modif*. 2024;48(3):312-359. [\[CrossRef\]](#)
10. Rothwell PM. External validity of randomised controlled trials: "To whom do the results of this trial apply?". *The Lancet*. 2005;365(9453):82-93. [\[CrossRef\]](#)
11. Campbell DT, Stanley JC. *Experimental and quasi-experimental designs for research*. 1st ed. Boston, MA: Houghton Mifflin; 1963. [\[CrossRef\]](#)
12. Smith JA. Evaluating the contribution of interpretative phenomenological analysis. *Health Psychology Review*. 2010;5(1):9-27. [\[CrossRef\]](#)
13. Aydin O. Rise of single-case experimental designs: a historical overview of the necessity of single-case methodology. *Neuropsychological Rehabilitation*. 2024;34(3):301-334. [\[CrossRef\]](#)
14. Aydin O. Characteristics of missing data in single-case experimental designs: an investigation of published data. *Behav Modif*. 2023;48(2):182-215. [\[CrossRef\]](#)
15. Kazdin AE. *Single-case research designs: methods for clinical and applied settings*. 1st ed. New York, NY: Oxford University Press; 2011. [\[CrossRef\]](#)
16. Dallery J, Raiff BR. Optimizing behavioral health interventions with single-case designs: from development to dissemination. *Transl Behav Med. Policy Res*. 2014;4(3):290-303. [\[CrossRef\]](#)
17. Byiers BJ, Reichle J, Symons FJ. Single-subject experimental design for evidence-based practice. *Am J Speech Lang Pathol*. 2012;21(4):397-414. [\[CrossRef\]](#)

18. Kratochwill TR, Hitchcock JH, Horner RH, et al. Single-case intervention research design standards. *Remedial Spec Educ*. 2013;34(1):26-38. [\[CrossRef\]](#)
19. Kravitz RL, Duan N, Braslow J. Evidence-based medicine, heterogeneity of treatment effects, and the trouble with averages. *Milbank Q*. 2004;82(4):661-687. [\[CrossRef\]](#)
20. Parker RI, Vannest K. An improved effect size for single-case research: nonoverlap of all pairs. *Behav Ther*. 2009;40(4):357-367. [\[CrossRef\]](#)
21. Onghena P, Michiels B, Jamshidi L, Moeyaert M, Van den Noortgate W. One by one: accumulating evidence by using meta-analytical procedures for single-case experiments. *Brain Impairment*. 2018;19(1):33-58. [\[CrossRef\]](#)
22. Schmid C, Yang J. Bayesian models for N-of-1 trials. *Harv Data Sci Rev*. 2022;2022(SI3):10.1162/99608f92.3f1772ce. [\[CrossRef\]](#)
23. Dallery J, Cassidy RN, Raiff BR. Single-case experimental designs to evaluate novel technology-based health interventions. *J Med Internet Res*. 2013;15(2):e22. [\[CrossRef\]](#)
24. Lillie EO, Patay B, Diamant J, Issell B, Topol EJ, Schork NJ. The N-of-1 clinical trial: the ultimate strategy for individualizing medicine? *Per Med*. 2011;8(2):161-173. [\[CrossRef\]](#)
25. Barlow DH, Hersen M. *Single Case Experimental Designs: Strategies for Studying Behavior Change*. Oxford, UK: Pergamon Press; 1984. [\[CrossRef\]](#)
26. Tate RL, Perdices M. *Single-Case Experimental Designs for Clinical Research and Neurorehabilitation Settings: Planning, Conduct, Analysis, and Reporting*. 1st ed. New York, NY: Routledge; 2019. [\[CrossRef\]](#)
27. Tekin-Iftar E. *Single-Case Research in Education and Behavior Sciences*. 2nd ed. Ankara, Turkey: Anı Publishing; 2018. [\[CrossRef\]](#)
28. Kent DM, Rothwell PM, Ioannidis JP, Altman DG, Hayward RA. Assessing and reporting heterogeneity in treatment effects in clinical trials: a proposal. *Trials*. 2010;11:85. [\[CrossRef\]](#)
29. Araujo A, Julious S, Senn S. Understanding variation in sets of N-of-1 trials. *PLoS One*. 2016;11(12):e0167167. [\[CrossRef\]](#)
30. Aydin O, Tanious R. Performance criteria-based effect size (PCES) measurement of single-case experimental designs: a real-world data study. *J of App Behav Analysis*. 2022;55(3):891-918. [\[CrossRef\]](#)
31. Shadish WR, Cook TD, Campbell DT. *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*. 1st ed. Boston, MA: Houghton Mifflin; 2002. [\[CrossRef\]](#)
32. McKnight PE, McKnight KM, Sidani S, Figueredo AJ. *Missing data: a gentle introduction*. New York, NY: Guilford Press; 2000. [\[CrossRef\]](#)
33. Edgington ES. *Randomization tests*. 3rd ed. New York, NY: Marcel Dekker; 1995. [\[CrossRef\]](#)
34. Kratochwill TR, Horner RH, Levin JR, Machalicek W, Ferron J, Johnson A. Single-case design standards: an update and proposed upgrades. *J Sch Psychol*. 2021;89:91-105. [\[CrossRef\]](#)
35. Rahimi MD, Zargani H, Nikkhah K. Effectiveness of transcranial direct current stimulation in chronic daily headaches: an experimental single case report with a novel protocol. *Case Rep Neurol*. 2022;14(3):381-388. [\[CrossRef\]](#)
36. Gouda MA, Buschhorn L, Schneeweiss A, Wahida A, Subbiah V. N-of-1 trials in cancer drug development. *Cancer Discov*. 2023;13(6):1301-1309. [\[CrossRef\]](#)
37. March L, Irwig L, Schwarz J, Simpson J, Chock C, Brooks P. N-of-1 trials comparing a non-steroidal anti-inflammatory drug with paracetamol in osteoarthritis. *BMJ*. 1994;309(6961):1041-1044. [\[CrossRef\]](#)
38. Spauwen P, Ter Mors B, van Harten P, Domensino AF, Ponds R, van Heugten C. Efficacy and safety of amantadine as a treatment for apathy after brain injury: two single-case experimental design studies. *Neuropsychological Rehabilitation*. 2020;32(6):872-896. [\[CrossRef\]](#)
39. Kronish IM, Hampsey M, Falzon L, Konrad B, Davidson KW. Personalized (N-of-1) Trials for depression: a systematic review. *J Clin Psychopharmacol*. 2018;38(3):218-225. [\[CrossRef\]](#)
40. Samuel JP, Tyson JE, Green C, et al. Treating hypertension in children with N-of-1 trials. *Pediatrics*. 2019;143(4):e20181818. [\[CrossRef\]](#)
41. Gkouskou KK, Grammatikopoulou MG, Lazou E, Sanoudou D, Goulis DG, Eliopoulos AG. Genetically-guided medical nutrition therapy in type 2 diabetes mellitus and pre-diabetes: a series of N-of-1 superiority trials. *Front Nutr*. 2022;9:772243. [\[CrossRef\]](#)
42. He W, Cui Z, Chen Y, Wang F, Li F. Status of N-of-1 trials in chronic pain management: a narrative review. *Pain Ther*. 2021;10(2):1013-1028. [\[CrossRef\]](#)
43. Nikles CJ, Mitchell GK, Del Mar CB, Clavarino A, McNair N. An N-of-1 trial service in clinical practice: testing the effectiveness of stimulants for attention-deficit/hyperactivity disorder. *Pediatrics*. 2006;117(6):2040-2046. [\[CrossRef\]](#)
44. Braam W, Didden R, Smits MG, Curfs LMG. Melatonin for chronic insomnia in Angelman syndrome: a randomized placebo-controlled trial. *J Child Neurol*. 2008;23(6):649-654. [\[CrossRef\]](#)
45. Kravitz RL, Duan N, eds; DEcIDE Methods center N-of-1 guidance panel. Design and implementation of N-of-1 trials: a user's guide. AHRQ Publication No. 13(14)-EHC122-EF. Rockville, MD: Agency for healthcare research and quality; 2014. [\[CrossRef\]](#)
46. Shadish WR. Analysis and meta-analysis of single-case designs: an introduction. *Journal of School Psychology*. 2014;52(2):109-122. [\[CrossRef\]](#)
47. Heyvaert M, Onghena P. Randomization tests for single-case experiments: State of the art, state of the science, and state of the application. *Journal of Contextual Behavioral Science*. 2014;3(1):51-64. [\[CrossRef\]](#)
48. Manolov R, Solanas A, Sierra V. Extrapolating baseline trend in single-case data: problems and tentative solutions. *Behav Res Methods*. 2018;51(6):2847-2869. [\[CrossRef\]](#)