Chapter 2
The Application of the Single Subject Design

The single subject design is a family of designs that share fundamental concepts and methodologies. The basic components of a single subject design are similar to other research designs, which include the measurement of a variable of interest or outcome variable, and the effect of an intervention on this variable. In general, the researcher expects the intervention or treatment (i.e., the independent variable) to have an impact on the outcome (i.e., the dependent variable). Research conducted in the area of psychology and social sciences commonly refers to the dependent variable as the target behavior [1–3]. In contrast, researchers in the biomedical sciences commonly refer to the dependent variable as the outcome, or more specifically, the clinical impact as measured by laboratory values, intensity, number, or duration of a symptom, and so forth. The term "target behavior" can be limiting when applied to biomedical research, as biomedicine involves numerous types of outcomes, in which behavior is of one possibility. Thus, the terms "outcome" or "outcome of interest" will be used, as these are more accurate descriptors for dependent variables in biomedicine.

Choice and Measurement of Outcomes

The choice of the outcome must be driven by the study goals, and well-controlled measurements of the outcome are repeatedly conducted throughout the design. The outcome variable should include a descriptive name, a general definition, an elaboration of the outcome facets, and basic examples [1, 4]. In essence, the outcome variable should be operationally defined (i.e., observable, measurable and verifiable). Depending upon the study design and the research questions, outcomes within biomedicine might include systolic blood pressure readings, HbA1c levels, Beck Depression Inventory scores, and lymphocyte counts, among others. As previously mentioned, the operationally defined outcome is expected to change over the course of the study. The measurement of the outcomes can be obtained through methods such as observation, self-report, clinical assessment, and physiological measurement. When considering the methods of gathering data, the temporal frequency of recording the outcome is also of importance.

J.E. Janosky et al., *Single Subject Designs in Biomedicine*,
DOI 10.1007/978-90-481-2444-2_2, © Springer Science+Business Media B.V. 2009

Quantitative assessment of the outcome variable is critical. Examples of assessment include measures of frequency, interval, duration, and intensity. Frequency recording of a particular outcome occurs within a specified time-frame. Within each predetermined time-frame, the number of times an outcome occurs is measured. An example could include the number of times that self-monitored blood glucose values exceed 180. Also, interval measurement entails dividing blocks of time into smaller intervals, and then measuring the outcome during each interval. Duration recording is another assessment strategy that simply measures the length of time an outcome occurs, and finally intensity measurement involves the relative magnitude of an outcome [5]. In biomedicine, examples of outcome variables might include blood pressure readings of systolic and diastolic blood pressure taken through a physician via a standardized protocol, self-reported blood glucose levels taken at home via glucometer use, and electroencephalogram activity (EEG) evaluated by a radiologist. The measurement of these outcomes could be recorded at various intervals throughout a specified period of time, such as every day or once a week. Continuous assessment of an outcome is frequently used in single subject designs, since the overarching purpose is to analyze the effects of an intervention or treatment over time [2]. This approach allows for examination of outcome patterns between the time a treatment is implemented, withheld, or removed. As with any research design, measurements repeatedly gathered over time must be conducted under standardized protocols. Examples include procedures for collecting the data, description of the measurement tools, the research or environmental setting, and other salient features that may affect the outcome of the study. It is also important to recognize the limitations of the measurement devices that are employed. For instance, reactivity could occur with self-monitoring and observation of behavior [6]. In terms of biomedicine, examples could include diabetic patients altering their usual diets or those with hypertension increasing their level of exercise. Participant reactivity will likely vary depending on the level of social desirability of the measured outcome. Accurate measurement is crucial for a sound design; thus, potential issues must be carefully evaluated prior to the study.

Finally, in between-group studies, it is sufficient to have two observations occurring during pre- and post-interventions, as multiple patients are included in the analyses. In contrast, it is common practice to gather numerous or repeated measurements for each patient in a single subject design. Repeated measures over time permit the researcher to analyze patterns and stability of the dependent variable during the various phases of the design. This allows the researcher to generate inferences regarding sources of variability on the outcome over time, particularly when alternating experimental designs.

Choice and Application of Interventions

Another important issue for consideration in single subject designs is intervention selection. The choice of the intervention must be based on the goal or purpose of the research. Interventions should be implemented in a standardized manner, so as

to reduce the chance for outcome variability due to methodological effects (See Chapter 4 for more detail on this topic). Also, procedures used for measuring the outcome during the baseline phase should be identical to procedures employed during the intervention [5]. In terms of intervention implementation, the cardinal rule is to change one variable at a time when proceeding to subsequent phases, so that the effects of each intervention can be evaluated independently [1]. It is difficult to obtain accurate conclusions if more than one variable is altered, since intervention effects cannot be parsed apart from the outcome variable. Frequently a criterion for successful treatment is identified a priori to determine the effectiveness of the intervention. If progress is not achieved during the intervention phase, then the treatment may be altered, or an entirely new treatment may be implemented [7] ; hence there is flexibility in intervention selection of single subject designs. Examples of interventions may include pharmaceutical therapy for hypertension (e.g., ACE inhibitor) or insulin therapy for blood glucose control (e.g., dosage of insulin). A surgical procedure (e.g., prosectomy) could also serve as an intervention for patients with prostate cancer.

Nomenclature of Single Subject Designs

Single subject designs are denoted through the tabulation and identification of phases of research activity, where these phases of research activity include baseline measurement of an outcome (A), treatment or intervention (B), removal of a previous treatment or intervention (C), and so forth. Capital letters have traditionally been used to indicate specific phases of the conduct of the research.

Most frequently designated as the A-phase, or baseline phase, of a single subject study, the A denotes a measurement of the outcome in the absence of an intervention. Frequently these measurements are recorded prior to the introduction of the intervention, as this allows for the natural occurrence of the outcome or target dependent variable [1, 2]. This view has been referred to as the natural course or the natural history. Baseline data serve as a standard of current performance that can be compared to future changes in the outcome [1, 7]. More specifically, the baseline projections are criteria for the evaluation of interventions, which are a crucial aspect of single subject research. Although there is no special formula for determining the length of time for measurement, it is suggested the baseline be continued until it has stabilized [1]. A baseline is considered stable when there is little variability in subsequent measurements, including no trends or slopes. Although the number of measures can vary considerably, the typical number of measurements for the A-phase of a design has been 5–7 [8]. Figure 2.1 illustrates a hypothetical example of baseline data. Without the introduction of an intervention, blood glucose levels are plotted to reflect natural baseline levels of blood glucose, along with predicting future levels of the outcome. Specifically, future levels are denoted by the dashed lines, or the mean level of the plotted data points.

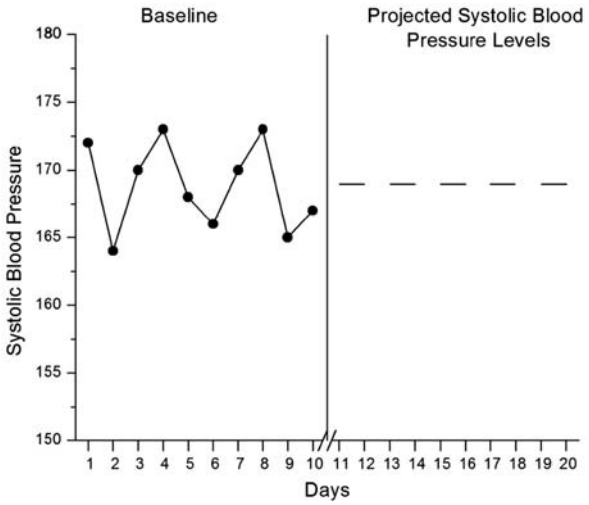

Fig. 2.1 Hypothetical baseline data of systolic blood pressure levels. Baseline data points (*solid lines*) are used to predict future systolic blood pressure levels (*dashed lines*)

Often referred to as the intervention or B-phase, the outcome is measured in order to determine the efficacy or effectiveness of the intervention. Data trends are examined between the A- and B-phases, particularly whether the outcome systematically increases or decreases over time [2]. If trends are similar between the baseline and intervention phases, then the utility of the intervention is questionable. However, excessive variability possibly due to unreliable or inaccurate measurements can interfere in analyzing and interpreting the data, consequently leading to muddled conclusions. If a patient is measuring at-home readings of blood glucose values with different meters, there is a possibility of negative impact on the outcome due to the conflicting measures of each meter. Additional treatments are denoted by subsequent letters, such as "B," "C," "D," and so forth. For example, a B-A-C design represents the implementation of an intervention "B," followed by withdrawal of intervention "A," and then the introduction of a new intervention "C."

The Family of Single Subject Designs: The Basic A-B Design

The A-B design forms the basis of the family of single subject designs [9]. Despite their simplicity, essentially all single subject designs are methodological variations of the A-B design. The single subject design has been considered an advanced form of a pretest-posttest design, as there are typically more frequent measurements of the outcome [2]. Following identification of the outcome, baseline data are gathered (A) and an intervention is implemented (B). The natural occurrence of the outcome is reflected in the A-phase, whereas outcome changes in the B- phase are attributed to the intervention [1]. In biomedicine, the features of an A-B design could include

adding nutritional agents to pharmaceuticals, including a chromium supplement as an additional control of blood glucose levels for an individual with diabetes. Data could be gathered on the outcome, such as blood glucose levels, prior to and following the introduction of the supplement. As an additional illustration, the effects of vitamin and mineral supplements, such as vitamin E, could be examined as a joint treatment to pharmaceuticals for an individual with high cholesterol.

Figure 2.2 displays an example of an A-B design from a study of a patient being treated by a family physician through care at a Federally-Qualified Health Center (FQHC). A pharmaceutical treatment was administered with the intention of lowering systolic blood pressure readings. Notice that baseline (A) stability was established following four measurements of systolic blood pressure. In the intervention phase (B), seven measurements of systolic blood pressure were gathered. Systolic blood pressure gradually decreased over time in the treatment condition. Thus, the treatment is assumed to have been responsible for the outcome changes.

The A-B design has the weakest internal validity of all the single subject design options. Multiple factors could potentially contribute to outcome changes during the intervention. Changes may occur due to practice effects, maturation, or random effects, for example [7] (See Chapter 3 for more details). Although the A-B design has many limitations, it has been shown to have some utility in settings where control-group analyses or repeated treatment withdrawals are not possible [1, 9, 10]. Nonetheless, extending the A-B design through incorporating additional elements is a better strategy for establishing evidence for a causal relationship between the intervention and observed outcomes.

There are several strengths and weaknesses of the A-B design. First, this design permits the researcher to analyze and compare an outcome variable before and

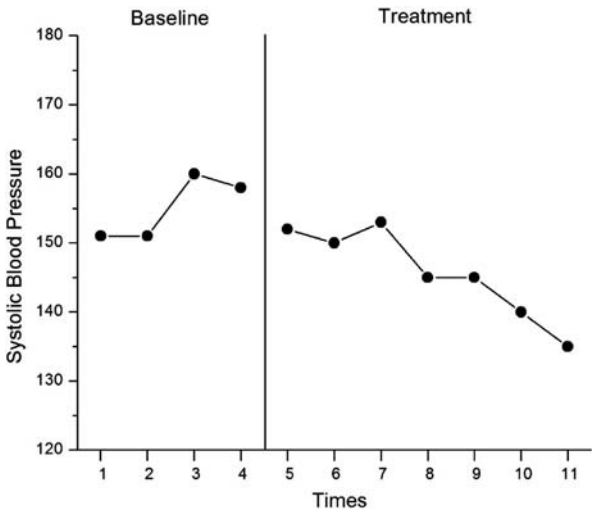

Fig. 2.2 Illustration of an A-B design targeting systolic blood pressure

during intervention, which affords greater reliability than an intervention-only design. The A-B design is simple and commonly used in clinical settings. However, a disadvantage of the design is that it cannot control many of the threats to internal validity, like maturation, history effects, testing effects, and instrumentation [11] (See Chapter 3 for more information). For example, a maturation effect may be responsible for outcome changes over time, rather than the intervention alone; that is, natural developmental changes in the life of the participant may have coincided with the treatment. The A-B design has utility in measuring the magnitude of outcome changes, despite being unable to solely attribute outcome variations to the intervention effects.

The A-B-A Design

The A-B-A design is more favorable than the A-B design because it adds potential control effects with a cessation of the intervention, or an intervention withdrawal. The intervention withdrawal occurs during one or more phases, in order to demonstrate that changes in the outcome only occur during the intervention [1, 2]. Intervention withdrawal increases the degree of certainty that changes in the outcome are attributed to the intervention; however, it should be noted that the influence of extraneous variables may never be entirely eliminated. In addition, although the A-B-A design is commonly referred to as a reversal design, this term may be misleading. A reversal design not only encompasses the withdrawal of an intervention, but often an attempt to revert the outcome variable to initial baseline levels [11, 12]. In single subject designs, it is not always plausible that an outcome will revert to its original levels following the withdrawal of an intervention. This would especially be the case for designs containing interventions with long-lasting effects, such as remediating of a disease (Fig. 2.3).

An example of an A-B-A design in clinical practice is illustrated through consideration of a pharmaceutical intervention for diabetes. A physician may be treating a patient for diabetes, with the expectation of controlling blood glucose, and is evaluating the effectiveness of a medication. The selected target outcome variable is hemoglobin A1C, recorded without any interventions during the baseline (A) phase. Several measurements are recorded until stable during the baseline and also under the same conditions (i.e., physician-gathered measurements, participant seated, etc.). During the intervention (B) phase, the medication is introduced and several hemoglobin A1C recordings are conducted over time. These recordings are also gathered under the same conditions of measurement. Next, the medication is discontinued (i.e., withdrawal of the intervention occurs), and once again, baseline blood glucose levels are recorded. Hemaglobin A1C measurements are examined across the A- and B-phases. In this hypothetical example, hemoglobin A1C returns to original baseline levels when the treatment is withdrawn in the second baseline phase. Since outcome levels in the B-phase are closer to desired levels, the medication intervention is assumed to have been responsible for the outcome effects.

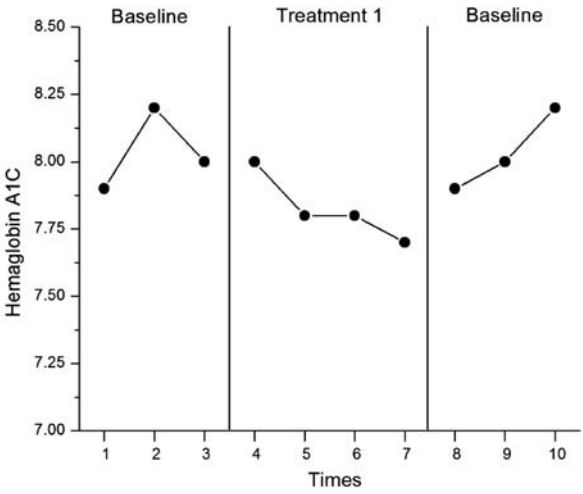

Fig. 2.3 A hypothetical example demonstrates how an A-B-A design can be used to study the effects of a medication intervention

Another issue in A-B-A designs is the timing of withdrawal of the treatment. Multiple factors are frequently involved in this decision-making process, such as time limitations, staff cooperation, and ethical considerations [1]. Intervention withdrawal is frequently necessary in order to attribute outcome improvement to intervention effects. There are ethical limitations concerning participants no longer receiving potentially beneficial interventions. This dilemma could be applied to the aforementioned example, in which a physician implements a medication for a patient with diabetes. Between phases of a single subject design, the physician may withdraw the beneficial medication and replace it with a placebo. Some researchers have argued that this is essential, whereas others have stated that it is unethical. However, the majority of researchers agree that once a study is terminated, patients should have access to beneficial interventions, regardless of whether they were withdrawn during the study. The issue regarding the appropriateness of withdrawal of interventions, successful and unsuccessful, is addressed in Chapter 5. Additional considerations when deciding the timing of intervention withdrawal include the efficacy or effectiveness of the intervention, cost of the intervention, availability of the medical system or the intervention, and other similar issues [1]. In essence, there are no steadfast rules in determining when to withdraw treatment.

As discussed previously, measurements must be obtained under standardized conditions. For example, standardized conditions entail measuring the outcome variable at the same time of day, using the same devices for recording or measurements, instructions, method of recording, and environmental conditions. Care must be taken because there is always potential for an extraneous variable, such as time of day, to impact the outcome measurement. For example, the blood glucose levels of a patient with diabetes may fluctuate depending on the time of day and whether

the measurements were fasting, postprandial, and so forth. Thus, deviating from the aforementioned conditions could result in spurious outcome effects [1]. If deviations from any conditions temporally coincide with the introduction of the intervention, a change in the outcome cannot be attributed solely to the intervention. It is possible that the alteration in the condition partially contributed to the outcome effects. In this case, it would be incumbent on the researcher to either re-evaluate the intervention using standard conditions, or evaluate the deviation that occurred before making conclusions.

Although the A-B-A design still contains some of the inherent flaws found in A-B designs, the withdrawal increases the ability to infer causality. Withdrawing an intervention may be used to determine whether or not the outcome returns to the level recorded at baseline. However, there are certain situations where conditions are irreversible, and the outcome is not expected to return to baseline [11]. For example, once treatments targeting social skills and reading are withdrawn, one cannot unlearn these skills. In addition, there are ethical issues in terminating the study after a baseline (A) phase, as patients are denied the full benefits of the intervention [1]. Following the study, the researcher should consider allowing patients access to various treatment options. Other potential problems associated with the A-B design include carryover effects of the multiple withdrawals and reinstatements of treatment interventions [1, 3]. Specifically, dependent variable changes in the final phase may not be similar to the initial baseline phase, in which the intervention had not yet been introduced.

A-B-A-B Design

Campbell and Stanley [9] refer to the A-B-A-B design as an equivalent time-samples design. This design corrects for some weaknesses of the A-B-A design, as the A-B-A-B design terminates on an intervention (B) phase. This extension is particularly useful in that effects can be analyzed between both B to A, and then A to B, which strengthens conclusions between the intervention and outcomes [1]. The previously discussed example of a pharmaceutical intervention for diabetes can be used to illustrate the extension in the A-B-A-B design. Initially, blood glucose measurements are gathered during the baseline phase (A), in which no medications are introduced. When the measurements are deemed stable, the medication is introduced during the intervention phase (B), and blood glucose measurements are again recorded. Next, the medication is withdrawn during the baseline phase (A), and then the medication is reintroduced for the final intervention phase (B).

B-A-B Single-Subject Design

The B-A-B design is commonly used to evaluate the methodological effectiveness of interventions. In this design, an intervention phase (B) is first introduced, then

withdrawn (A), and finally reinstated in the last phase (B). However, researchers have been known to implement a shortened baseline phase prior to the main B-A-B design [1]. Although the B-A-B design is more tenable than the A-B-A design, in that the intervention is implemented during the terminal phase, the absence of an initial baseline phase makes the A-B-A-B design more preferable [1]. There is added control to studies that include the collection of baseline data prior to the introduction of the intervention. The primary strength of the B-A-B design is that it can be implemented when a patient presents with a current treatment already in place and the investigator wants to clarify the effectiveness of the treatment, or determine the potential ramifications of non-treatment. Additionally, the B-A-B design serves as a precursor to a host of more complex designs, which involve alternating from an existing intervention (B) to a new intervention (C) hypothesized to be more effective. For example, in a B-A-B-A-C design, researchers can obtain information on the effectiveness of an intervention that is already in place and compare its effectiveness to non-treatment, as well as an alternative intervention. This methodology is useful when an existing treatment is insufficiently effective or accompanied by undesirable side effects.

Multiple Baseline Design

In the multiple baseline design, an intervention is introduced to different outcomes at various time periods. Visually, multiple baselines appear to be a series of A-B designs that are placed above one another. Although multiple baselines can include two or more baselines, studies most commonly analyze data over three or more baselines [1–2]. First, baseline data are simultaneously gathered on two or more baselines. The baseline data for each outcome reflect the current, naturally-occurring level without the intervention. Once baselines are stable for all outcomes, the researcher then applies an intervention on only one selected outcome, while baseline data continue to be recorded for the other outcomes. The simultaneous baseline measurements indicate whether changes only occur with the outcome specifically targeted by the intervention [1, 7, 13]. Figure 2.4 shows an example of a multiple baseline design.. Once criterion levels are reached in the first target outcome, the intervention is introduced to the second outcome. Consequently, once criterion levels are met in the intervention phase of the second outcome, the intervention is then introduced to the third outcome. Notice that each outcome only increases following introduction of the intervention.

Multiple baseline designs can include multiple baselines across participants, settings, and outcomes. A multiple baseline design across outcome could entail two or more target outcomes across the same treatment and in an identical setting. Also, a multiple baseline design across participants encompasses two or more individuals in the same setting, who receive the same intervention directed toward target outcomes. A criterion level is frequently established a priori for analyzing the success of an intervention. Intervention effects are demonstrated through achieving target

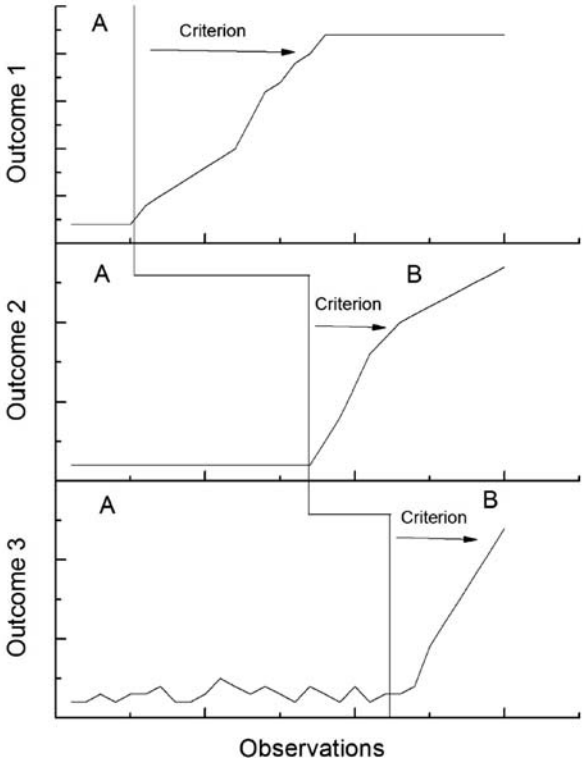

Fig. 2.4 Example graph of the multiple baseline design

criterion levels predetermined by the researcher. Thus, a changing criterion design can be included within multiple baseline designs (changing criterion designs will be further discussed in a later section) [7]. Researchers should select outcome variables that are somewhat independent from each other, as covariance can occur among target outcomes; however, completely unrelated outcomes may not respond to a single intervention [7, 14].

Multiple baseline designs are unique in that various design outcomes are tested as control conditions, and changes can be analyzed without implementing an intervention. When an intervention is applied to some outcomes and not others, an intervention and no-intervention condition can be used for comparison. Outcomes that are gathered simultaneously allow researchers to make inferences that baseline outcomes would continue to be stable if the intervention were not provided [7]. Baselines not yet receiving an intervention should be compared at the same time with those receiving the intervention, so as to determine potential intervention effects.

There are several advantages and disadvantages involved in multiple baseline designs. Situations can exist in which withdrawal or reversal designs are not appropriate. Carryover effects of the intervention may appear across phases, such as

with medication interventions, or withdrawing an intervention may pose risks to the participant [1]. Since ethical considerations are of utmost importance, multiple baseline designs, along with alternating treatment designs, can be very useful when withdrawals and reversals are inappropriate. Multiple baseline designs are also useful when more than one target outcome is in need of an intervention [7]. The aforementioned potential for covariance is an issue of concern in multiple baseline designs, as carryover effects can confound the results. Some researchers hold that multiple baseline designs are less efficient than withdrawal or reversal designs, as these contain more direct relationships between the intervention and target outcome [15]. Despite these challenges, multiple baseline designs are frequently used by researchers because they do not require reversals, and consequently, avoid some of the ethical issues inherent in other single subject designs.

Alternating Treatments Design

Also referred to as the multiple schedule design [16] and the multi-element design [17], the alternating treatments design evaluates the effects of two or more interventions on a single outcome. Two or more interventions are alternated rapidly, but not necessarily within a fixed period of time [1]. The term rapid might indicate that a participant receives alternating interventions each and every time he or she is tested, which might occur daily, weekly, or even monthly. Researchers do not analyze trends in improvement over time, since two or more interventions are alternating. Instead, for example, the researcher plots all the data points for Intervention A and compares it to trends in the data points for Intervention B. Also, although the term treatment is contained in the title of the design, this designation does not preclude other non-therapeutic interventions. Rather, any intervention can be implemented. It should also be noted that the alternating treatments design is commonly used in combination with other single subject designs, specifically when determining which of several treatments is most effective [7].

Figure 2.5 displays data from an alternating treatments design gathered from a study conducted on an African-American male (age 51) being treated by an internist. Specifically, three different insulin dosage regimens were employed in an alternating fashion, targeting hemaglobin A1C. Only two measurements were gathered during the baseline (A) phase; however, the researcher had data indicating baseline (A) levels could be established with only two measures. In most cases, it is recommended that baseline stability should be determined following multiple observations. As presented in Fig. 2.5, hemaglobin A1C levels decreased as treatment progressed. Levels continued to decrease with each new introduction of the three treatments. At the conclusion of the study, data points for each medication intervention are presented in separate plots. Trends are compared between each of the interventions. If there is greater improvement with one intervention relative to another intervention, it is inferred that this specific medication is more effective than the other medication.

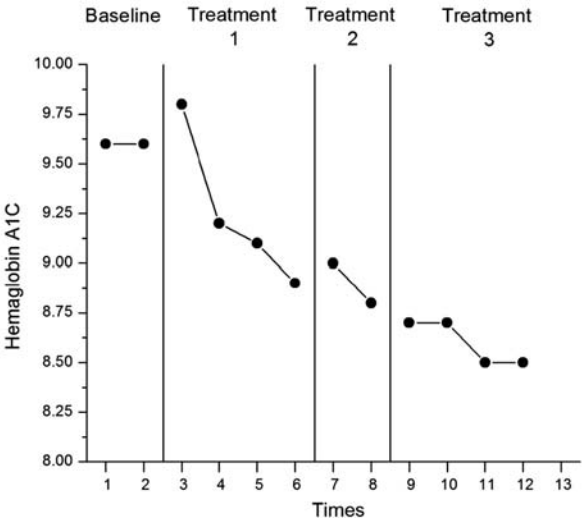

Fig. 2.5 An alternating treatments design is presented targeting hemoglobin A1C levels

The presentation order of the alternating interventions should not be systematic as in an A-B-A-B-A-B design, for example. The researcher should randomize the presentation order of interventions to control for sequential confounding (i.e., order effects or carryover effects), in which the introduction of one intervention influences a subsequent intervention [1]. Intervention order should be counterbalanced, so as to minimize carryover and order effects. For example, three interventions could be randomly presented in the following blocks: C-A-B, A-B-C, and B-C-A. Carryover effects can also be decreased by separating intervention sessions with a time interval and slowing down the timing of alternations [1]. Additionally, researchers should present each block of interventions for an equal number of times, as doing so strengthens experimental control and creates consistency within the experimental procedures [18].

Various types of alternating treatments designs exist, some of which do not incorporate baseline phases. Alternating treatments with no baselines are useful in that interventions can be immediately implemented. Nonetheless, it should be noted that although it is unnecessary to collect baseline data in the alternating treatments design, it is prudent to still gather baseline data if at all possible [1, 19]. Many researchers using this design include baseline data by replacing an intervention phase with a no-treatment phase, commonly referred to as the alternating treatments with a control condition design [18]. However, it should be cautioned that a no-treatment phase is not the equivalent of a pre-intervention baseline, and multiple treatment interference can occur when a no-treatment phase is used between various intervention phases [7, 15, 20]. Specifically, carryover effects may occur with interventions preceding the introduction of a baseline phase.

An additional variation of the alternating treatments design includes a baseline followed by an alternating treatments design. Although baseline stability is not a requirement of the alternating treatments design, the initial baseline should include an outcome that is stable. There are situations in which baseline stability is unnecessary for ethical purposes, such as with severe conditions that may benefit from immediate employment of the intervention [7]. Another situation not requiring baseline stability includes trends progressing in the opposite direction of the goal. In this case, baseline data collection can be discontinued and the intervention implemented. An alternating treatments design beginning with a baseline phase could also be altered to contain only the most effective intervention for the final phase. Eliminating less effective interventions can save time and money for the researcher. In addition, as discussed previously, for ethical reasons it is essential to continue effective interventions following study termination.

The alternating treatments design has many advantages and disadvantages. The design is very useful for researchers analyzing the effectiveness of several interventions. Also, the design progresses more rapidly due to the alternating intervention phases. If designs contain baseline phases, it is unnecessary for data to be stable prior to intervention implementation [7]. In addition, there are fewer ethical concerns when compared to other designs, since intervention withdrawal is unnecessary. Although counterbalancing can be employed to decrease order effects, multiple intervention interference is an issue of concern, as interventions are continually alternated [1]. Despite the potential for carryover or confounding effects in multiple treatments, interference can be minimized by implementing interventions that substantially differ from one another. Also, alternating treatment designs are not appropriate for targets that cannot be reversed, such as learning a skill. Intervention implementation is rapid; therefore, this design should not be used for interventions producing slow change over time. Although the alternating treatments design has several disadvantages, the application of this design can be quite useful in a wide-array of biomedical settings.

Changing Criterion Design

In changing criterion designs, intervention effects are demonstrated through achieving target criterion levels that are predetermined by the researcher, such as a specific blood pressure level. Within this design, the outcome must gradually improve over time, in order to meet specified criteria. Criteria are repeatedly altered throughout the intervention to reflect improvement in the outcome, and rewards can be implemented when criterion levels are met or surpassed. The purpose of contingencies is to facilitate the increase or decrease of the target outcome. Following baseline collection (A) in the changing criterion design, the intervention (B) is divided into subphases requiring target outcome progression toward the ultimate goal [21]. Similar to the basic A-B-A-B design and multiple baseline design, a baseline is used for comparative purposes. If the intervention is responsible for change, outcome levels

in each subphase should correspond with shifts in the specified criterion. However, fluctuating outcomes would likely reflect effects from extraneous variables that are inconsistent with desirable criterion levels.

An example of a changing criterion design involves increasing minutes of daily exercise. Initially the patient may engage in little to no exercise during the baseline phase. In this case, the specified criterion may be engagement in 15 minutes of exercise per day. If the criterion is met, the patient could earn reinforcements or rewards, such as setting aside time for an enjoyable activity or money for exercise-related item purchases. If the patient is consistently meeting the criterion for several consecutive days, the criterion could be increased to 20 minutes, 25 minutes, and so on. In essence, the goal is gradually increased as the target outcome both meets the criterion and is stabilized. The criterion is continually altered until the desired level is achieved.

Issues for consideration with changing criterion designs include phase length, magnitude of criterion changes, and number of phase or criterion changes [3, 7, 22]. In terms of phase length, subphase levels are used as baselines for subsequent phases; thus, it is essential that outcomes are stable before progressing to a new subphase. If the outcome is able to change rapidly, then shorter subphases can be implemented. Causal relationships cannot be concluded from intervention effects; however, the relationship between the intervention and outcome is strengthened when dependent variable levels remain close to the designated criterion during each subphase.

There are no stringent guidelines for determining the magnitude change in the criterion that should occur over subphases. If guidelines require only a small change in the outcome, then there may be ambiguity as to whether other extraneous factors, such as maturation or practice effects, were responsible for the changes [7]. Alternatively, criteria demanding large changes that are not reached may indicate the magnitude is too large. When deciding the initial criterion level, the lowest or highest baseline data point can be used for an approximate estimate. Other options for determining initial criterion levels include calculating a 10 or 15 percent increase or decrease of the mean baseline level [2]. Throughout the course of the study, larger criterion changes can be implemented with outcomes of greater variability, while more stable outcomes can use smaller criterion changes [22]. These criteria changes should improve the detection of correspondence between the outcome and the criteria.

In addition, the number of criterion changes included in a study should be considered. Although a minimum of two criterion shifts must be included in this design, multiple subphases are generally implemented [2]. It would be difficult to demonstrate intervention effects with only one criterion shift; however, an excessive number of criterion shifts may create ambiguity. The determination of number of criterion shifts is frequently contingent on the magnitude of criterion changes and length of phases [7]. For example, length of time available for the study could be an issue for consideration, along with outcome stability during subphases.

The changing criterion design does not require withdrawing or withholding an intervention to demonstrate relationships. Rather, the design can be adapted to include additional subphases containing reversals to a previous criterion level or baseline [7]. There are challenges in analyzing unidirectional changes over time during an intervention phase. Extraneous variables may be responsible for improvement of the target, rather than the intervention alone. In order to rule out threats to internal validity, such as practice effects, bidirectional changes should be evaluated. Intervention effects can be analyzed by increasing or decreasing the criterion and determining if the target outcome corresponds with those changes. Relationships between the intervention and outcome are further strengthened if changes occur in the direction of the specified criterion.

There are several advantages and disadvantages associated with the changing criterion design. The design is useful with target outcomes that can increase or decrease in a stepwise fashion, particularly when the terminal goal can only be reached over a long length of time [18], such as with increasing medication dosages or drug titration. Changing criterion designs are also appropriate when evaluating interventions containing contingent reinforcement or punishment, and when treatment withdrawal cannot occur. Despite the desirable characteristics of the changing criterion design, it has been employed less often than other single subject designs [2]. This may be partially explained by the restricted application for certain target outcomes; for example, it is recommended the outcome be contained in the patient′s repertoire (e.g., smoking, reading, eating, etc.) [2, 23]. The aforementioned advantages of the design could also be seen as disadvantages, in that interventions must present contingencies, and outcomes must change gradually. Although there are restrictions for the implementation of the changing criterion design, it offers researchers unique options that are not found in the other single subject designs.

Summary

The various methodological components inherent in the family of single subject designs offer a wide array of options and flexibility for researchers, as each single subject design contains strengths and limitations. Consideration of the research question is essential when creating and selecting a design, since certain designs may be more appropriate for the investigation of specific research questions. The research question also dictates target informative outcomes for measurement, the actual intervention, and potential ethical concerns, among other issues. Single subject designs can be particularly useful for events in which interventions are costly and for unique populations. As a whole, single subject designs allow researchers to implement procedures that may be less cumbersome than large N designs [11]. Findings can also be used for comparison with other single and between-subject designs. The family of single subject designs offers flexible options that can be beneficial within the field of biomedicine.

References

1. Barlow DH, Hersen M. *Single case experimental designs: Strategies for studying behavior change* (2nd ed.). New York: Pergamon Press, 1984.
2. Kazdin A. *Single-case research designs: Methods for clinical and applied settings*. New York: Oxford University Press, 1982.
3. Kratochwill TR (Ed.). *Single subject research: Strategies for evaluating change*. New York: Academic Press, 1978.
4. Hawkins RP. Developing a behavior code. In DP Harmann (Ed.), *Using observers to study behavior: New direction for methodology of social and behavioral science* (pp.21–35). San Francisco, CA: Jossey-Bass, 1982.
5. Brown-Chidsey R, Steege MW. *Response to intervention*: *Principles and strategies for effective practice*. New York: The Guilford Press, 2005.
6. Haynes SN, Wilson CC. *Behavioral assessment: Recent advances in methods, concepts, and applications*. San Francisco, CA: Jossey-Bass, 1979.
7. Richards SB, Taylor RL, Ramasamy R, Richards RY. *Single subject research: Applications in educational and clinical settings*. San Diego, CA: Singular Publishing Group, 1999.
8. Janosky JE, Al-Shboul QM, Pellitieri TR. Validation of the use of a nonparametric smoother for the examination of data from a single-subject design. *Behavior Modification.* 1995; 19(3): 307–324.
9. Campbell DT, Stanley JC. *Experimental and quasi-experimental designs for research*. Chicago, IL: Rand McNally, 1966.
10. Cook TD, Campbell DT (Eds.). *Quasi-experimentation: Design and analysis issues for field settings*. Boston, MA: Houghton Mifflin Company, 1979.
11. Krishef CH. *Fundamental approaches to single subject design and analysis*. Malabar, FL: Krieger Publishing Company, 1991.
12. Leitenburg H. The use of single-case methodology in psychotherapy research. *Journal of Abnormal Psychology.* 1973; 82(1): 87–101.
13. Baer DM, Wolf MM, Risley RR. Some current dimensions of applied behavior analysis. *Journal of Applied Behavior Analysis.* 1968; 1: 91–97.
14. Tawney JW, Gast DL. *Single subject research in special education*. Columbus, OH: Merrill, 1984.
15. Cooper JO, Heron TE, Heward WL. *Applied behavior analysis*. Columbus, OH: Merrill Publishing Company, 1987.
16. Hersen M, Barlow DH. Single case experimental design: Strategies for studying behavior change. New York: Pergamon, 1976.
17. Ullman J, Sulzer-Azaroff B. Multielement baseline design in educational research. In E Ramp, G Semb (Eds.), *Behavior analysis: Areas of research and application* (pp. 371–391). Englewood Cliffs, NJ: Prentice-Hall, 1975.
18. Alberto P, Troutman A. *Applied behavior analysis for teachers* (5th ed.). Columbus, OH: Merrill, 1999.
19. Neuman S. Alternating treatments designs. In S Neuman, S McCormick (Eds.), *Single subject experimental research: Applications for literacy* (pp. 64–83). Neward, DE: International Reading Association, 1995.
20. Barlow DH, Hayes SC. Alternating treatments design: One strategy for comparing the effects of two treatments in a single subject. *Journal of Applied Behavior Analysis.* 1979; 12: 199–210.
21. Poling, A, Methot L, LeSage M. *Fundamentals of behavior analytic research*. New York: Plenum Press, 1995.
22. Hartmann DP, Hall RV. The changing criterion design. *Journal of Applied Behavior Analysis.* 1976; 9(4): 527–532.
23. Hall RV. *Behavior modification: The measurement of behavior.* Lawrence, KS: H & H Enterprises, 1971.