HHS Public Access

Author manuscript

Ann Behav Med. Author manuscript; available in PMC 2015 December 01.

Published in final edited form as:

Ann Behav Med. 2014 December ; 48(3): 392–401. doi:10.1007/s12160-014-9618-7.

Approaches for Informing Optimal Dose of Behavioral Interventions

Corrine I. Voils, PhD,
Durham Veterans Affairs Medical Center, Durham, NC USA. Duke University Medical Center, Durham, NC USA

Heather A. King, PhD,
Durham Veterans Affairs Medical Center, Durham, NC USA

Matthew L. Maciejewski, PhD,
Durham Veterans Affairs Medical Center, Durham, NC USA. Duke University Medical Center, Durham, NC USA

Kelli D. Allen, PhD,
Durham Veterans Affairs Medical Center, Durham, NC USA. Duke University Medical Center, Durham, NC USA

William S. Yancy Jr., MD, MHS, and
Durham Veterans Affairs Medical Center, Durham, NC USA. Duke University Medical Center, Durham, NC USA

Jonathan A. Shaffer, PhD
Columbia Medical Center, New York City, NY USA

Corrine I. Voils: corrine.voils@dm.duke.edu

Abstract

Background—There is little guidance about to how select dose parameter values when designing behavioral interventions.

Purpose—The purpose of this study is to present approaches to inform intervention duration, frequency, and amount when (1) the investigator has no a priori expectation and is seeking a descriptive approach for identifying and narrowing the universe of dose values or (2) the investigator has an a priori expectation and is seeking validation of this expectation using an inferential approach.

Methods—Strengths and weaknesses of various approaches are described and illustrated with examples.

Correspondence to: Corrine I. Voils, corrine.voils@dm.duke.edu.

Authors' Statement of Conflict of Interest and Adherence to Ethical Standards Authors Voils, King, Allen, Yancy, and Shaffer declare that they have no conflict of interest. Dr. Maciejewski has received consultation funds from Daichi Sankyo and ResDAC at the University of Minnesota, and owns stock in Amgen due to his spouse's employment. All procedures, including the informed consent process, were conducted in accordance with the ethical standards of the responsible committee on human experimentation (institutional and national) and with the Helsinki Declaration of 1975, as revised in 2000.

Results—Descriptive approaches include retrospective analysis of data from randomized trials, assessment of perceived optimal dose via prospective surveys or interviews of key stakeholders, and assessment of target patient behavior via prospective, longitudinal, observational studies. Inferential approaches include nonrandomized, early-phase trials and randomized designs.

Conclusions—By utilizing these approaches, researchers may more efficiently apply resources to identify the optimal values of dose parameters for behavioral interventions.

Keywords

Intervention dose; Intervention design; Dose–response; Methodology

Introduction

Determining an optimal dose is among the most important decisions investigators must make when designing a behavioral intervention [1]. Behavioral intervention dose may be characterized by duration, frequency, and amount [1, 2]. Duration refers to the period of time over which participants are exposed to the intervention and may be measured in hours, weeks, months, or years. Frequency refers to how often contact is made over a specified period of time, such as 52 visits over 1 year (once weekly). Amount refers to the total length of each intervention contact and is typically measured in minutes or hours. These dose parameters collectively determine cumulative intervention dose. A behavioral intervention may be delivered at a fixed or variable interval, and the dose may be delivered as needed (tailored) or uniformly (nontailored) [1].

When people consider optimal dose in the context of pharmacotherapy, they generally think of the dose that maximizes improvements in psychological and/or physical outcomes (“efficacy”) while minimizing adverse side effects (“toxicity”). Adverse effects of behavioral interventions are rarely discussed or assessed but could include physical harm (e.g., patients are injured as the result of a physical activity intervention), dependence on the interventionist, disillusion with research or health care, or reluctance to seek additional treatment [3]. Assuming that toxicity is minimized or nonexistent in a behavioral intervention, an optimal dose can be defined as either the maximally efficacious dose that is not conditional on patient adherence (intended dose) or the maximally effective (actual or observed) dose that is conditional on adherence. Behavioral intervention designers need to consider patient burden and adherence to ensure that the effective dose is as close as possible to the efficacious dose. Behavioral intervention A that is highly efficacious may have lower effectiveness because excessive burden leads to low attendance or high withdrawal. On the other hand, behavioral intervention B with the modest efficacy may have the same effectiveness as intervention A if patient adherence is high. There is tension between the efficacious and effective dose because the theoretical efficacious dose of a behavioral intervention can be designed, but the effective/actual dose can only be known upon application.

Understandably but regrettably, the efficacious dose of a behavioral intervention is often informed (or constrained) by budgetary and timeline considerations. For a grant mechanism with a specified time and budget limit, an investigator may use a deductive process to

determine a feasible intervention dose given staff salary levels, time frame of the project, and the desired sample size. Ideally, behavioral intervention dose parameters should be determined from formative work prior to conducting a full-scale randomized controlled trial (RCT) [4, 5]. In so doing, resources could be conserved and investigators could be more confident about the unique contribution of dose versus other intervention features (e.g., content).

Several approaches are available to inform the optimal dose of a behavioral intervention; the strengths and weaknesses of which are described in the following sections. We consider these approaches in the context of two situations: (1) the investigator has no a priori expectation and thus is seeking a descriptive approach for identifying and narrowing the universe of dose values or (2) the investigator has an a priori expectation about the optimal dose and is seeking validation of this expectation via an inferential approach. We focus first on the following descriptive approaches: retrospective analysis of RCT data, assessment of perceived optimal dose via prospective surveys or interviews of key stakeholders, and assessment of target patient behavior via prospective, longitudinal, observational studies. Then, we focus on nonrandomized, early-phase trials, and randomized designs for drawing inferences about dosing parameter values. Advantages and disadvantages of these descriptive and inferential approaches are described in the following sections and are summarized in Table 1.

Descriptive Approaches to Identify and Narrow the Universe of Dose Values

When investigators have no a priori expectation about optimal values for one or more dose parameters (duration, frequency, amount), several descriptive approaches are available to identify initial values of dosing parameters that can be validated via hypothesis testing in subsequent studies.

Retrospective Analysis of Data from Completed RCT(s)—If dosing data are available from a completed, relevant RCT, retrospective data analyses may be performed among participants randomized to treatment to examine the possibility of a dose–response relationship and the dose parameters that drive this relationship. Intervention exposure (patient adherence to the intervention or effective dose) may be operationalized as total minutes of contact in the intervention (a function of duration, frequency, and amount) or as the number or proportion of contacts received (a function of duration and frequency) [1]. Retrospective analyses can also be conducted to examine whether intervention exposure varies by patient characteristics to inform targeting of subsequent interventions to specific patient subgroups. For example, in a telephone-delivered intervention to improve diet and physical activity, the number of completed telephone calls was not associated with baseline levels of dietary intake or physical activity [6]. If the number of completed calls had been inversely associated with baseline physical activity levels, then the investigator could consider offering more calls to patients with lower baseline physical activity levels than to those with higher levels, thereby reducing burden for, and conserving resources devoted to, patients with higher baseline physical activity levels.

Although retrospective analyses of a completed RCT are appealing and commonly conducted in some contexts, generalizing to other contexts may be problematic. Changes in the target population, setting, interventionist, other study features, or other intervention features may yield different results. Additionally, the optimal dose may not be represented in a single study.

To improve external validity of information learned from previously conducted RCTs, researchers may systematically review a larger number of published, and perhaps unpublished, studies that examine a common outcome in relation to a specific intervention. Previously conducted RCTs provide normative information (i.e., which dose is typically examined) and may contribute to formal analyses of the dose–response relationship via meta-regression. A common practice in systematic reviews is to examine the frequency or total number of intervention contacts. For example, number of treatment sessions was not associated with improvements in depression in a systematic review of psychological interventions to improve depression among patients with coronary heart disease [7]. Analyses could be expanded to examine more dose parameters, including main effects and interactive effects, when such data are available [1].

Formal systematic reviews have informed clinical guidelines for behavioral interventions. For example, the US Preventive Services Task Force obesity guideline indicates that high-frequency interventions are more effective than moderate or low-frequency interventions, leading to the recommendation that weight loss interventions be delivered over 12 to 26 sessions in the first year [8]. These reviews have also informed policy. For example, the Center for Medicare and Medicaid Services now reimburses intensive primary care obesity treatment, defined as one face-to-face visit every week for month 1, one face-to-face visit every other week for months 2–6, and, among patients who lose at least 3 kg in months 1–6, one face-to-face visit every month for months 7–12 [9].

An advantage of retrospective data analysis from a single or multiple completed RCTs is that efficacious and effective dose can be examined if intended dose and actual dose are reported. Differences between intended dose (i.e., expected duration, frequency, and amount) and actual dose can provide information on feasibility and patient adherence. Large deviations between intended and actual dose may suggest that one or more dose parameters must be attenuated to improve feasibility, intervention acceptability via patient adherence, or interventionist treatment fidelity.

One disadvantage to secondary data analysis, whether of a single RCT or multiple RCTs, is that causation in dose–response relationships cannot be inferred because randomization that was applied to allocate patients to treatment or control was not applied to allocate patients in the treatment arm(s) to receive more or less intervention. The actual intervention dose was, instead, driven by patient selection. For example, suppose in a behavioral intervention to reduce congestive heart failure symptoms that patients with greater intervention exposure experienced fewer symptoms. Such a result could be interpreted in two ways: either the intervention was effective in reducing symptoms or healthier participants felt better and thus were able to participate more. Other challenges of using systematic reviews to evaluate a dose–response relationship include underreporting of dose parameters; underreporting of

both intended and actual dose; heterogeneity in control groups, interventions, and dosing schedules; and low power [1, 10]. If these issues are so prevalent that they preclude retrospective analysis of dose parameters from completed RCTs, then investigators may wish to conduct prospective studies.

Assessment of Perceived Optimal Intervention Dose via Prospective Survey or Interview of Key Stakeholders—Researchers can gain information regarding optimal intervention dose via a second descriptive approach: surveys or interviews of key stakeholders. With the advent of the Patient-Centered Outcomes Research Institute, an emphasis has been placed on gaining perspectives and input from a variety of stakeholders [11]. For a behavioral intervention, key stakeholders may include members of the target patient population, caregivers and other informal support persons, clinic staff, clinic or hospital administrative personnel, operations partners, or insurers. Each type of stakeholder may be asked different questions to provide insights on intervention acceptability, efficacy, effectiveness, adherence challenges, or feasibility of implementation. For example, members of the target patient population may provide information about acceptability of a proposed intervention dose, such as whether a proposed intervention schedule would be too frequent or infrequent and too long or short in duration. This information could inform effectiveness but is less likely to inform efficacy. Clinicians may provide insights based on clinical experience as to the frequency or duration of contact with patients that they believe would be needed to effect change and that is feasible for their clinic configuration, staffing, and throughput to inform efficacy and effectiveness. Administrators may provide information on the dose that would be feasible and affordable if the intervention were implemented on a broad scale. As the ultimate goal is to incorporate behavioral interventions into clinical practice, it is important to design interventions that can be implemented with available resources.

Advantages of surveys and interviews include the involvement and perspectives of various stakeholders and flexibility to assess a broad range of issues in an efficient manner using open- or close-ended questions. Open-ended questions can be useful when there is little known or poor consensus about initial values of dose parameters, whereas close-ended questions can be useful to obtain feedback about possible initial values. One disadvantage of surveys and interviews is that results may also be difficult to interpret if stakeholder viewpoints contrast.

Additionally, interpretation of results may not be straightforward. For example, in a previous weight loss study [12], participants met every 2 weeks for months 1–6 and then monthly for months 7–12. Attendance was lower during months 7–12, which might lead one to conclude that attendance in months 7–12 would have been even worse if meetings had continued every 2 weeks, particularly because time conflict was the most common reason given for nonattendance. Yet, at study conclusion, retained participants indicated that reducing frequency in months 7–12 made dietary adherence difficult due to less social support and accountability (data available from WSY). Therefore, maintaining meetings every 2 weeks may have stimulated better attendance.

Other considerations are that clinician perceptions about optimal dose may not accurately reflect the dose that is necessary to effect change. Moreover, the dose for which an insurer or healthcare system would be willing to pay may be too modest to yield clinically meaningful changes in patient outcomes. Additionally, as surveys and interviews typically involve a single assessment, reports may be inaccurate due to recall bias, the availability heuristic, or social desirability [13]. Finally, perceived optimal dose may not match the timescale of the target behavior, a possibility that can be investigated via prospective, longitudinal, observational studies.

Assessment of Target Patient Behavior via Prospective, Longitudinal, Observational Studies—A third descriptive approach to narrowing the universe of dose values is via prospective, longitudinal, observational studies. Long-term longitudinal studies involve repeated measurements of individuals over intervals, such as years or decades (i.e., intraindividual change), whereas short-term longitudinal studies involve repeated assessments over many more closely spaced occasions, such as minutes, hours, days, and weeks. Short-term longitudinal studies include daily diary studies and experience sampling methods/ecological momentary assessment [14, 15]. Measurement burst designs combine these approaches, using repeated sequences of these intensive measurements separated by longer time intervals [16]. Information on how often, and for how long, unwanted behaviors or experiences occur (particularly when unknown or assumed from prior cross-sectional studies) may be informed by any of these observational study designs.

In such observational studies, patients may complete self-report questionnaires (e.g., self-reported medication adherence) or use devices that directly measure behavior, such as engaging in physical activity (e.g., pedometer) or opening pill bottles (e.g., Medication Event Monitoring System). Direct measurement via devices is advantageous because behaviors are observed rather unobtrusively following a period of acclimation [17–19], and patient contact/response is not required. These studies may also rely upon retrospective information extracted from patient electronic medical records (e.g., screening visits, refill adherence).

Decisions regarding optimal dose depend on the behavior or variable of interest and its identified time course (i.e., how often it occurs and/or changes if known) as well as the research question(s). Some behaviors occur less frequently, such as screening behaviors (e.g., pap smear, mammogram), whereas other behaviors and experiences generally occur more frequently, such as taking medications and experiencing physical health symptoms. To inform intervention dose from repeated measures, one could calculate the mean time span between occurrences of unwanted behaviors (e.g., medication nonadherence) within persons and then assess the mean and standard deviation for the full sample. This information could be used to determine how frequently to intervene over a specified duration and whether intervention frequency should be tailored (in the case of great variation around the sample average lag) or untailored (in the case of less variation around the sample average lag).

Prospective, longitudinal, observational studies offer the ability to examine intraindividual change and variability and the ability to determine how frequently unwanted thoughts, feelings, and behaviors occur. Disadvantages of longitudinal studies include selection bias,

attrition, and time burden. Additionally, causation cannot be determined from observational studies due to lack of randomization and potential confounding.

Inferential Approaches to Validate Dose Parameters

Once investigators have identified and narrowed the universe of dose parameters to a subset of plausible values, perhaps via one of the aforementioned approaches, then they can conduct studies to validate the most effective dose parameters. Nonrandomized and randomized inferential approaches can be applied.

Nonrandomized, Early-Phase Trials—Behavioral intervention dose parameters could be validated using nonrandomized, early-phase, dose-finding designs such as the continual reassessment method (CRM) [13] that was initially developed to test the maximally tolerated doses of novel medications. The CRM is especially useful in cases in which randomization is not feasible, sensible, or ethical. The rationale for conducting these early-phase trials is to concentrate as many patients as possible at doses at, or close to, an optimal dose. Furthermore, these trials can be used to simultaneously determine that a “minimally effective dose” is less than a “maximally tolerated dose.” Proponents of the CRM and other early-phase designs believe them to be superior to traditional dose-finding designs because they “learn” from information obtained at earlier time points in the study, are more likely to treat patients at efficacious doses, and are less likely to treat patients at harmful doses. These designs assume that the probabilities of both efficacy and toxicity increase as dose increases [20]. As such, the goal of these designs is to identify the highest (i.e., most efficacious) dose whose risk of toxicity or harm is tolerable. In behavioral interventions, toxicity is a lesser concern or no concern at all, so these methods can be applied to identify the maximally effective dose.

These early-phase trials require specification of a dose–response function (i.e., an a priori estimate of the probability of response at each dose level) [21]. To begin, one to three participants are run through the intervention at a dose that is deemed by the clinical investigator(s) to be optimal, based upon previous literature regarding use of the intervention in other patient populations and clinical intuition. The occurrence of a response (or lack of response) in these participants provides information for the statistical model and is used to compute an iterative adjustment of the initial dose–response model for the next group of one to three participants. Dose can be escalated or de-escalated at each adjustment. Importantly, these CRM models use data from all participants to estimate the dose–response curve, and they yield a confidence interval for the recommended dose for subsequent trials. Although nonrandomized, early-phase trials were initially designed for use with binary outcomes, they have since been expanded to incorporate time-to-event data [22], nonbinary outcomes [23, 24], and multiple outcomes [25].

As an example of how to apply the CRM to a behavioral intervention, consider a study that tests the efficacy of problem-solving treatment (PST) for improving quality of life in heart failure patients. To begin, the clinical investigator may select a range of number of contacts (e.g., 2, 4, 6, 8, or 12 sessions) of PST to be offered over a fixed duration of 8 weeks based on previous studies of PST in other populations [26], recommendations in the PST manual,

and clinical intuition. The investigators might further assume a dose–response model characterized by “diminishing returns” in number of contacts (i.e., an expectation that quality of life will increase monotonically across doses of PST but that quality of life improvements will be more modest with larger doses). All participants could receive assessments of quality of life at baseline and 8 weeks, and “response” could be defined as a 10-point improvement on a 100-point quality of life measure given that this amount of improvement has been previously identified as clinically significant.[27] The investigators might select a target probability of response of 66 %—that is, the investigators would be seeking to identify the dose of PST that will yield a 10-point improvement in quality of life for 66 % of participants after all participants have completed the trial. Previous studies of psychotherapy dose–response have included a 50 % response rate [28], but this convention harkens back to early dose-finding studies in pharmacology and may not be applicable to behavioral intervention dose-finding studies.

In this example, the first group of three participants would receive six sessions of PST, which is an a priori assumption of the optimal dose based on the response to PST seen in care-givers of breast cancer patients [26]. After the first group completes the intervention, the estimate of the optimal dose would be updated based on the number of participants who achieve a 10-point improvement in quality of life. The next group of three participants would then be assigned to this new estimate of the optimal dose, and the procedure would be repeated until the final group completes the intervention. Although dose escalation and de-escalation decisions are allowed in this design, not all doses within the prespecified range of doses may be tested in this trial. If, for example, only one out of three participants in the first group experiences a 10-point increase in quality of life, then the investigator would likely escalate the dose to eight sessions without testing the lower doses. If, on the other hand, all three participants in the first group that receives six sessions experience a 10-point increase in quality of life, then the investigator might de-escalate to four sessions for the next group.

Nonrandomized, early-phase designs such as the CRM that were developed for evaluating toxicity-efficacy tradeoffs for medication dosing have not yet been applied to behavioral interventions, possibly because of their statistical complexity and because toxicity may be less relevant in behavioral interventions [3]. Additionally, these designs may not identify an optimal dose given that they do not test the full range of possible dose values. Nonetheless, these designs offer promise for identifying the optimal dose of behavioral interventions because they may be more cost-effective than traditional randomized designs and other dose-finding approaches. More empirical and conceptual work is needed to understand the comparative costs of these designs versus traditional approaches.

Randomized Designs—A second inferential approach to validate the most effective dose parameters is via randomized designs, of which we discuss full and reduced factorial designs. The effect of different intervention doses on outcome(s) of interest may be evaluated in experiments in which participants are randomized to one level (value) of at least one dose parameter (e.g., 4-vs. 8-week duration). When two or more dose parameters are manipulated (e.g., frequency and duration), or when one dose parameter is combined with at least one other intervention component (e.g., receipt vs. no receipt of meal replacement), a factorial research design may be used [29]. Fully factorial designs are those

in which every level of an independent variable is crossed systematically with every level of the other independent variable(s). The number of conditions, or combinations, is denoted by 2^k where *k* refers to the number of factors (e.g., for a 2×2 design, there would be 2^2, or four, treatment combinations). Applied to the evaluation of dose parameters, one could construct a 2×2 factorial experiment to evaluate whether the effect of frequency (e.g., once monthly vs. twice monthly) differs by study duration (e.g., 6 vs. 9 months). In cases in which the number of combinations in a factorial design is so large that specific combinations are illogical or harmful, or evaluating every possible combination is unfeasible (e.g., logistically or due to cost), a reduced design may be used [30]. Reduced designs are those in which all *k* factors are manipulated, but not all combinations are evaluated. Three types of reduced designs for evaluating multiple independent variables include the individual experiment, single-factor design, or fractional factorial design. See Collins et al. [30] for a more complete discussion of factorial designs.

In the individual experiment approach, a two-condition (i.e., one experimental and one control) experiment is conducted for each independent variable *k* on samples that are typically much larger than CRM samples. In the single-factor approach, one experiment is conducted with several levels of a factor evaluated against a single control group. In the fractional factorial approach, an investigator tests only a carefully chosen subset of treatment combinations [30]. The choice of whether to conduct a fully factorial or a reduced factorial design depends on multiple considerations, including sequencing (i.e., whether the result of one study must be known to inform a subsequent study), resources (i.e., total budget available and cost of each combination and each participant), interest in interactions, number of conditions and participants required, and confounding of main effects and interactions [4, 30]. Another consideration is whether dose parameters are being examined in isolation or in combination with other dose parameters or intervention components. If one wished to examine the effect of a single dose parameter, then individual experiments with two conditions or single-factor experiments with many levels can be conducted.

To illustrate, suppose a researcher intends to examine the effect of financial incentives on physical activity levels. One of the first considerations in designing such a trial is the amount of financial incentive to deliver. In this particular example, incentive amount is manipulated while frequency (one payment per week) and duration (26 weeks) of incentive delivery are held constant. Published literature might suggest that amounts between $5 and $10 would be sufficient for inducing clinically significant increases in physical activity. With the individual experiment approach, one might compare an amount (e.g., $8) to $0. If the result were significant, then one might conduct a follow-up experiment to determine whether a smaller amount (e.g., $5) would be equally effective, whereas if the result were nonsignificant, then one might conduct a follow-up experiment to determine whether a larger amount (e.g., $9, $10) would be effective. Thus, a series of individual experiments could be conducted to identify the optimal financial incentive amount.

In contrast to conducting a series of individual experiments, one could take the single-factor approach and conduct an experiment in which several financial incentive amounts are compared to $0 (control), such as values ranging from $5 to $50 in $5 increments. Assuming $50 were the true upper limit, one could identify the lowest amount that differed from $0.

One risk of the single-factor approach is using valuable resources (e.g., personnel, money, participant time) on conditions in which the amount is too small to be effective or too large to be feasible for a full-scale RCT or dissemination. Yet, the amount of resources used on a carefully designed experiment prior to conducting an RCT may prevent the waste of hundreds of thousands or millions of dollars on a full-scale RCT in which an ineffective financial incentive amount is used.

Less common variations on randomized designs can be used to examine dosing parameters as well. In N-of-1 designs, for instance, an individual may receive individualized dose escalation of treatment followed by usual care to determine whether an improvement (e.g., depressive symptoms, pain levels, blood pressure) occurs during dose escalation [31–33]. Sequential, multiple assignment, randomized trial (SMART) designs may be used to evaluate treatment sequences that involve changes in dose parameters [5]. SMART designs have four components: (1) a sequence of decisions regarding what to do for responsive and nonresponsive participants; (2) a set of treatment options at each decision point; (3) tailoring variables that dictate responsive versus nonresponsive to treatment; and (4) a sequence of decision rules that determine which treatment to administer at the time of the decision [5, 34]. As an example, participants are randomized initially to receive one of two effective diets (e.g., low carbohydrate and low fat/low calorie) delivered via groups that meet every 2 weeks. After 16 weeks, weight is assessed; participants losing at least 4 kg could be randomized to receive the same dietary approach either at the same frequency or once per month (less frequently), whereas participants who do not lose at least 4 kg could be randomized to receive either the same diet approach weekly (more frequently) or the other diet delivered every 2 weeks (same frequency).

Randomized designs have a number of advantages. The effect of extraneous variables is assumed to be equally distributed between groups, so internal validity is maximized. Additionally, such designs allow the evaluation of many values for one or more dose parameters and intervention components in isolation, in conjunction with one another, or in combination with other intervention components. Despite these strengths, randomized experiments can be resource intensive and thus require separate funding sources. They may also be time intensive, particularly if several iterations are needed until an optimal dose is identified. Finally, they may not be adaptive.

Discussion

Intervention dose is the sine qua non of drug development, yet relatively little attention has been paid to explicit formulation of behavioral intervention dose, possibly because a framework of dosing parameters has been lacking until recently [1, 2]. We have outlined several descriptive and inferential approaches that can be used to identify and narrow the universe of possible dosing values to those that are plausible for testing as well as inferential approaches to validate which values are optimal. These approaches can be used to evaluate duration, frequency, and amount in isolation or in combination. Which descriptive or inferential approach to use depends on several factors, including availability of resources, how much is known about the area of interest, availability of primary data and literature on the topic, and research goals. For example, retrospective analysis of completed RCTs

provides information on what has or not worked, whereas surveys or interviews provide information on perceptions of what might work.

Assumptions

Application of the aforementioned approaches for establishing optimal efficacious and effective doses assumes two conditions, the first of which is optimal implementation fidelity. The best designed intervention, even if administered at the optimal dose, will not be efficacious or effective if not delivered with a high degree of fidelity [35]. Fidelity may vary by interventionist or practitioner background, experience, or perceived treatment acceptability and efficacy. Fidelity should be assessed during interventionist training and throughout the study duration to detect and avoid drift. When possible, standardized methods should be implemented to assess content fidelity (e.g., Motivational Interviewing Treatment Integrity code to assess therapist adherence to motivational interviewing techniques) [36]. Because many interventions comprise multiple components and have different theoretical foundations, however, intervention fidelity assessment tools may need to be tailored for each study [37]. When assessed, degree of fidelity can be evaluated across intervention dose parameters. For example, in a recent study of smoking cessation support services, transcripts of counseling sessions were coded by two independent raters using a taxonomy of 43 behavior change techniques. There was no correlation between proportion of behavior change techniques used and duration of counseling sessions (i.e., amount) [38].

Application of the aforementioned descriptive and inferential approaches also assumes that dosing parameters are not confounded with intervention content. This assumption may be more reasonable in some intervention contexts than others. A counseling session that takes 40 min instead of 20 min may incorporate additional behavior change techniques, such as motivational interviewing or relapse prevention [39]. In contrast, in a study of financial incentives for health behavior change, isolating financial incentive amount is straightforward. Another consideration is that, for some intervention techniques, amount may vary significantly from patient to patient and contact to contact.

Dosing Schedules

As mentioned previously, an intervention dose can be delivered at a fixed or variable interval and tailored to an individual’s needs or uniformly. A fixed, nontailored intervention dose might be a 20-min telephone call that occurs monthly for 6 months. A fixed, tailored intervention dose might be a 20-min telephone call that occurs monthly for 6 months only if the average weekly home blood pressure monitoring value exceeds 140/90 mm/Hg for any week in the preceding month. A variable, nontailored intervention dose might be 20-min telephone calls that occur weekly for the first 2 months and then twice-monthly for the next 4 months. Finally, a variable, tailored intervention dose might be 20-min telephone calls that occur weekly for the first 2 months and then twice monthly for the next 4 months only if the average weekly home blood pressure monitoring value exceeds 140/90 mm/Hg for any week in the preceding month.

Whether the dose should be delivered at a fixed or variable interval and in a tailored or nontailored fashion may be informed by clinical and theoretical considerations. When

monitoring a health status parameter such as blood pressure or blood glucose, safety or medication management considerations may determine the schedule. When an intervention comprises reinforcement or punishment for desired or undesired behavior, respectively, then the schedule may be dictated by learning theory. For example, fixed or variable-interval schedules may be selected, depending on whether the goal is to initiate or maintain behavior. Learning theory also suggests that withdrawing incentives is likely to result in extinction of the new behavior; indeed, this has been observed in studies of financial incentives for health behaviors [40, 41]. Thus, in some circumstances, theory may inform intervention dosing.

These various dosing schedules can be evaluated using the aforementioned descriptive or inferential approaches. For example, stakeholders can be asked about perceptions of feasibility, acceptability, efficacy, or effectiveness of interventions in which the dose is tailored to an individual's needs or uniformly. Likewise, a systematic review or experiment can involve the comparison of fixed versus variable-interval schedules.

Timeline and Budgetary Considerations

In drug development, dose-finding occurs prior to evaluation of safety and effectiveness. Behavioral intervention studies should also identify the efficacious dose of a behavioral intervention prior to a full-scale RCT to conserve resources [4]. Budgetary issues also practically inform behavioral intervention dose at the validation and translation stages. In the validation stage, budgets for intervention resources are often fixed, and it may not be clear whether a single dosing parameter (e.g., duration) or multiple parameters (e.g., duration and frequency) should be changed from the pilot intervention or whether increases or decreases should be made to each domain. Given a fixed budget, budget neutrality may require an intervention with frequent contacts of brief duration to be changed to less frequent contacts of longer duration.

Challenges in maintaining an efficacious dose also likely arise during translation and implementation when budget constraints are even greater than during the validation stage and return on investment concerns become increasingly important. Translation of established interventions likely requires modification of dosing parameters to retain the effect size and cost-effectiveness while making the intervention less intensive. Such tradeoffs may require scaling back some dosing parameters from their validation-stage levels to be sustainable, such as less frequent contact or less amount for each contact. Ideally, a modestly less effective and modestly less costly intervention can remain as cost-effective as its more effective and more costly predecessor.

Decisions that occur during translation could be informed by formative work conducted in the initial stages of intervention development. For example, if an investigator determined that once monthly contact was more effective than once quarterly contact, then the investigator would be in a strong position to advise against once quarterly contact during implementation. More work is needed to understand how effectiveness, costs, and cost-effectiveness change with changes to dosing parameters of interventions found to be effective at the validation stage.

Generalizability

Application of the three dosing parameters may not be straightforward for every type of intervention. For interventions that involve interaction between at least one participant and interventionist, the three parameters are definable and measurable. In contrast, for interventions that are self-administered (e.g., web-based) or relayed via media campaigns, frequency and amount could be static and difficult to assess unless measures are put into place to control access and exposure, respectively. More work is needed to operationalize dosing parameters for these types of interventions.

Another consideration is that what works for one population at one period of time may not necessarily work for other populations or at different points in time. A behavioral intervention that is effective for one population may need to be delivered at a different dose to yield the same improvement in another population. Intervention effectiveness may vary across a range of physical, psychological, social, and demographic characteristics. Dose may also need to be adjusted to reflect secular trends, such as increased knowledge about a disease, effectiveness of public health campaigns, changes to guidelines, or availability of alternative interventions. Thus, as with other intervention components, dosing should undergo the optimization process, being updated and adjusted as information is learned [5].

Conclusion

When designing behavioral interventions, researchers make many decisions, including those regarding intervention content and dose. Determining content may be more straightforward due to the availability of theoretical models specifying constructs to target and associated processes of change [39]. Determining dose is less straightforward because there has been no framework for considering dose and little guidance on how to select dose values [2]. In theory, the potential values for each dose domain are infinite. We have outlined descriptive approaches to identify and narrow the universe of possible dose values and inferential approaches to validate specific values. By implementing these approaches in the early stages of intervention development, researchers will be better able to understand the unique contribution of intervention dose and be better equipped to inform translation efforts.

Acknowledgments

This article was supported by a post-doctoral fellowship to Dr. King from the Department of Veterans Affairs, Office of Academic Affiliations, Health Services Research and Development (HSR and D) Service (TPP 21–020), a Research Career Scientist award from HSR and D to Dr. Maciejewski (RCS 10–391), and with resources and facilities of the Veterans Affairs Medical Center in Durham, NC. The views expressed in this article are those of the authors and do not necessarily reflect the position or policy of the Department of Veterans Affairs or the United States government.

References

1. Voils CI, Chang YK, Crandell JL, et al. Informing the dosing of interventions in randomized trials. Contemp Clin Trials. 2012
2. Manojlovich M, Sidani S. Nurse dose: What's in a concept? Res Nurs Health. 2008; 31:310–319. [PubMed: 18231974]
3. Lilienfeld SO. Psychological treatments that cause harm. Perspect Psychol Sci. 2007; 2:53–70.

4. Collins LM, Baker TB, Mermelstein RJ, et al. The multiphase optimization strategy for engineering effective tobacco use interventions. Ann Behav Med. 2011; 41:208–226. [PubMed: 21132416]
5. Collins LM, Murphy SA, Strecher V. The multiphase optimization strategy (MOST) and the sequential multiple assignment randomized trial (SMART): New methods for more potent eHealth interventions. Am J Prev Med. 2007; 32:S112–S118. [PubMed: 17466815]
6. Waters LA, Winkler EA, Reeves MM, Fjeldsoe BS, Eakin EG. The impact of behavioural screening on intervention outcomes in a randomised, controlled multiple behaviour intervention trial. International Journal of Behavioral Nutrition & Physical Activity. 2011; 8:24. [PubMed: 21450069]
7. Dickens C, Cherrington A, Adeyemi I, et al. Characteristics of psychological interventions that improve depression in people with coronary heart disease: A systematic review and meta-regression. Psychomsomatic Medicine. 2013 epub ahead of print.
8. Moyer VA. Screening for and management of obesity in adults: U.S. preventive services task force recommendation statement. Ann Intern Med. 2012; 157:373–378. [PubMed: 22733087]
9. Services CfMaM. 2012 Obesity Counseling Reimbursement Fact Sheet. from http://www.ethicon.com/sites/default/files/Specialty/Bariatric%20Surgery/rhcp/2012-obesity-counseling-reimbursement-fact-sheet.pdf
10. Schünemann, HJ.; Oxman, AD.; Vist, GE., et al. Chapter 12: Interpreting results and drawing conclusions. In: Higgins, JPT.; Green, S., editors. Cochrane Handbook for Systematic Reviews of Interventions. Vol. 5.1.0. 2011.
11. Krumholz HM, Selby JV. Seeing through the eyes of patients: The patient-centered outcomes research institute funding announcements. Ann Intern Med. 2012; 157:446–447. [PubMed: 22847044]
12. Yancy W Jr, Westman E, McDuffie J, et al. A randomized trial of a low-carbohydrate, ketogenic diet versus orlistat plus a low-fat diet for weight loss. Arch Intern Med. 2010; 170:136–145. [PubMed: 20101008]
13. Willis, G. Cognitive interviewing: A tool for improving questionnaire design. Thousand Oaks: Sage; 2005.
14. Ram N, Gerstorf D. Time-structured and net intraindividual variability: Tools for examining the development of dynamic characteristics and processes. Psychol Aging. 2009; 24:778–791. [PubMed: 20025395]
15. Nesselroade, JR. The warp and woof of the developmental fabric. In: Downs, R.; Liben, L.; Palermo, D., editors. Visions of development, the environment, and aesthetics: The legacy of Joachim F. Wohlwil. Hillsdale, NJ: Erlbaum; 1991. p. 213-240.
16. Sliwinski M. Measurement burse designs for social health research. Soc Personal Psychol Compass. 2008; 2:245–261.
17. Wagner GJ, Ghosh-Dastidar B. Electronic monitoring: Adherence assessment or intervention? HIV Clin Trials. 2002; 3:45–51. [PubMed: 11819185]
18. van Onzenoort HA, Menger FE, Neef C, et al. Participation in a clinical trial enhances adherence and persistence to treatment: A retrospective cohort study. Hypertension. 2011; 58:573–578. [PubMed: 21825228]
19. Berg KM, Arnsten JH. Practical and conceptual challenges in measuring antiretroviral adherence. Journal of Acquired Immune Deficiency Syndrome. 2006; 43(Suppl 1):S79–S87.
20. Garrett-Mayer E. The continual reassessment method for dose-finding studies: A tutorial. Clinical Trials. 2006; 3:57–71. [PubMed: 16539090]
21. Le Tourneau C, Lee JJ, Siu LL. Dose escalation methods in phase I cancer clinical trials. J Natl Cancer Inst. 2009; 101:708–720. [PubMed: 19436029]
22. Cheung Y, Chappell R. Sequential designs for phase I clinical trials with late-onset toxicities. Biometrics. 2000; 56:1177–1182. [PubMed: 11129476]
23. Hu C, Cheung Y. On the efficiency of nonparametric variance estimation in sequential dose-finding. Journal of Statistical Planning and Inference. 2012; 143:593–602. [PubMed: 23329867]
24. Chen Z, Tighiouart M, Kowalski J. Dose escalation with overdose control using a quasi-continuous toxicity score in cancer phase I clinical trials. Contemporary Clinical Trials. 2012; 33:949–959. [PubMed: 22561391]

25. Thall P, Cook J. Dose-finding based on efficacy-toxicity trade-offs. Biometrics. 2004; 60:684–693. [PubMed: 15339291]
26. Toseland R, Blachard C, McCallion P. A problem solving intervetnion for caregivers of cancer patients. Soc Sci Med. 1995; 40:517–528. [PubMed: 7725125]
27. Spertus J, Peterson E, Conard M, et al. Monitoring clinical changes in patients with heart failure: A comparison of methods. Am Heart J. 2005; 150:707–715. [PubMed: 16209970]
28. Howard KI, Kopta SM, Krause MS, Orlinsky DE. The dose-effect relationship in psychotherapy. Am Psychol. 1986; 41:159–164. [PubMed: 3516036]
29. Fisher R. The arrangement fo field experiments. Journal of the Ministry of Agriculture of Great Britain. 1926; 33:503–513.
30. Collins LM, Dziak JJ, Li R. Design of experiments with multiple independent variables: A resource managemetn perspective on complete and reduced factorial designs. Psychol Methods. 2009; 14:202–224. [PubMed: 19719358]
31. Gabler N, Duan N, Vohra S, Kravitz R. N-of-1 trials in the medical literature: A systematic review. Med Care. 2011; 49:761–768. [PubMed: 21478771]
32. Larson E. N-of-1 trials: A new future? J Gen Intern Med. 2010; 25:891–892. [PubMed: 20632123]
33. Guyatt, GH.; Jaeschke, R.; McGinn, T. N-of-1 randomized controlled trials: Study design. In: Guyatt, GH.; Rennie, D.; Meade, M.; Cook, D., editors. Users' guides to the medical literature. New York: McGraw Hill Medical; 2008. p. 179-192.
34. Lei H, Nahum-Shani I, Lynch K, Oslin D, Murphy SA. A "SMART" design for building individualized treatment sequences. Annu Rev Clin Psychol. 2012; 8:21–48. [PubMed: 22224838]
35. Carroll C, Patterson M, Wood S, et al. A conceptual framework for implementation fidelity. Implement Sci. 2007; 2:40–49. [PubMed: 18053122]
36. Moyers T, Martin T, Manuel J, Hendrickson S, Miller W. Assessing competence in the use of motivational interviewing. J Subst Abus Treat. 2005; 28:19–26.
37. Song M, Happy M, Sandelowski M. Development of a tool to assess fidelity to a psychoeducational intervention. J Adv Nurs. 2010; 66:673–682. [PubMed: 20423402]
38. Lorencatto F, West R, Christopherson C, Michie S. Assessing fidelity of delivery of smoking cessation behavioural support in practice. Implement Sci. 2013; 8:40. [PubMed: 23557119]
39. Abraham C, Michie S. A taxonomy of behavior change techniques used in interventions. Health Psychol. 2008; 27:379–387. [PubMed: 18624603]
40. Volpp KG, Gurmankin Levy A, Asch DA, et al. A randomized controlled trial of financial incentives for smoking cessation. Cancer Epidemiol Biomark Prev. 2006; 15:12–18.
41. Burns RJ, Donovan AS, Ackermann RT, et al. A theoretically grounded systematic review of material incentives for weight loss: Implications for interventions. Ann Behav Med. 2012; 44:375–388. [PubMed: 22907712]

Table 1

Advantages and disadvantages of empirical approaches for evaluating behavioral intervention dose

Purpose	Approach	Advantages	Disadvantages
Identify and narrow the universe of dose values	Retrospective analysis of data from completed RCT(s)	• Provides quantitative evaluation of dose–response relationship • Permits identification of cases for post hoc analyses (e.g., qualitative interviews of participants who were not adherent to the protocol) to inform future studies • Comparison of intended and actual dose can be informative • Permits evaluation of moderators of dose–response relationship via meta-regression	• No randomization • Reverse causation: positive dose–response relationship may indicate that participants improved because they received more intervention OR that they participated more because they improved • Heterogeneity in control groups and interventions can obscure inferences
	Assessment of perceived optimal intervention dose via prospective survey or interview of key stakeholders	• Involves multiple stakeholders, including patients, providers, operations partners, and administrators • Evaluates perceived feasibility, acceptability, efficacy, or effectiveness of proposed doses • Includes open- or close-ended questions • Assesses broad range of issues efficiently	• No randomization • Stakeholder ideas may have little to do with efficacy or effectiveness • Feedback from various stakeholders may be inconsistent
	Assessment of target patient behavior via prospective, longitudinal, observational studies	• Determine how frequently unwanted thoughts, feelings, and behaviors occur (e.g., missed medication doses) • Examine long-term change or short-term variability in behaviors	• No randomization • Selection bias • Attrition • Time burden
Validate expectation of optimal dose	Early-phase nonrandomized methods	• Small sample size • Strong alternative when randomization is not feasible • Adaptive • Precise and provides confidence intervals around optimal dose • Considers both minimally effective dose and maximally tolerated dose	• No randomization • Statistical complexity • Scant evidence supporting its use for behavioral interventions
	Randomized designs	• Maximizes internal validity • Examines interactions between dose parameters or dose parameters and other intervention components • If more than one dose is efficacious, can distinguish	• May be resource intensive and difficult to obtain funding • May take several iterations until an optimal dose is identified

Purpose	Approach	Advantages	Disadvantages
		optimal dose based on resources required • Can evaluate sequences of dosing schedules	• May not be adaptive