#268

Adaptive and Other Innovative Designs for Effectiveness Studies of New Animal Drugs

Guidance for Industry

Submit comments on this guidance at any time. Submit electronic comments to https://www.regulations.gov. Submit written comments to the Dockets Management Staff (HFA-305), Food and Drug Administration, 5630 Fishers Lane, Rm. 1061, Rockville, MD 20852. All comments should be identified with the docket number FDA-2020-D-1401.

For further information regarding this document, contact AskCVM@fda.hhs.gov.

Additional copies of this guidance document may be requested from the Policy and Regulations Staff (HFV-6), Center for Veterinary Medicine, Food and Drug Administration, 7500 Standish Place, Rockville MD 20855, and may be viewed on the Internet at https://www.fda.gov/animal-veterinary, https://www.fda.gov/regulatory-information/search-fda-guidance-documents or https://www.regulations.gov.

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Veterinary Medicine (CVM)
October 2021

Table of Contents

Adaptive and Other Innovative Designs for Effectiveness Studies of New Animal Drugs

Guidance for Industry

This guidance represents the current thinking of the Food and Drug Administration (FDA or Agency) on this topic. It does not establish any rights for any person and is not binding on FDA or the public. You can use an alternative approach if it satisfies the requirements of the applicable statutes and regulations. To discuss an alternative approach, contact the FDA staff responsible for this guidance as listed on the title page.

I. Introduction

FDA is issuing this Guidance for Industry (GFI), as required under section 305 of the Animal Drug and Animal Generic Drug User Fee Amendments of 2018 (Pub. L. 115-234), to assist sponsors in incorporating complex adaptive and other novel investigation designs into proposed clinical investigation protocols and applications for new animal drugs under the Federal Food, Drug, and Cosmetic Act (FD&C Act). Section 305 of Pub. L. 115-234, among other things, directed FDA to hold a public meeting for interested parties to discuss innovative animal drug investigation designs and to issue guidance addressing the incorporation of the use of such elements of investigations as complex adaptive and other novel investigation designs, data from foreign countries, real-world evidence (including ongoing surveillance activities, observational studies, and registry data), biomarkers, and surrogate endpoints into clinical investigation protocols and applications to support the effectiveness of new animal drugs.

The contents of this document do not have the force and effect of law and are not meant to bind the public in any way, unless specifically incorporated into a contract. This document is intended only to provide clarity to the public regarding existing requirements under the law. FDA guidance documents, including this guidance, should be viewed only as recommendations, unless specific regulatory or statutory requirements are cited. The use of the word *should* in Agency guidances means that something is suggested or recommended, but not required.

II. Background

In the *Federal Register* of July 9, 2019 (84 FR 32749), FDA's Center for Veterinary Medicine (CVM) published a notice of a public meeting entitled "Incorporating Alternative Approaches in Clinical Investigations for New Animal Drugs" giving interested persons until August 17, 2019, to comment on the topics discussed at the public meeting and the questions published in the meeting notice (84 FR at 32750-32751).[1] On August 13, 2019, we published a notice announcing the extension of the comment period to September 16, 2019 (84 FR 40071). CVM

[1] https://www.fda.gov/animal-veterinary/workshops-conferences-meetings/public-meeting-incorporating-alternative-approaches-clinical-investigations-new-animal-drugs

received numerous comments on the topics discussed at the public meeting and the questions published in the meeting notice. Those comments were considered as draft guidance was developed.

This document describes recommendations for designing, conducting, and reporting the results for investigations or studies including adaptive design features to demonstrate substantial evidence of effectiveness or a reasonable expectation of effectiveness of drugs intended for use in animals and to support the approval of a new animal drug application (NADA) or an application for conditional approval of a new animal drug (CNADA).[2] This guidance also provides information about obtaining feedback from CVM with respect to incorporating adaptive design features in investigations and study protocols for new animal drugs. Other centers within FDA, including the Center for Drug Evaluation and Research (CDER), the Center for Biologics Evaluation and Research (CBER), and the Center for Devices and Radiological Health (CDRH), have released guidance documents on the topics of adaptive design and other innovative designs, such as master protocols, enrichment strategies, and the use of Bayesian statistics.

CVM will consider all established and accepted adaptive design methodologies in submissions to investigational new animal drug (INAD) files, new animal drug applications (NADA), and applications for conditional approval of a new animal drug (CNADA) to demonstrate substantial evidence of effectiveness or a reasonable expectation of effectiveness. This guidance document provides CVM's recommendations specific to investigations for animal drugs.

Some concepts and language in the recommendations for animal drugs are intended to be similar or the same as those in other guidance documents issued by FDA on the same or similar topics. Because these recommendations are specific to investigations for animal drugs, they have been tailored to the unique aspects of and considerations for animal drug development.

III. Scope

The purpose of this guidance is to provide recommendations to animal drug sponsors for whom it may be beneficial to use adaptive and other innovative designs to demonstrate substantial evidence of effectiveness or a reasonable expectation of effectiveness of drugs intended for use in animals and to support the approval of an NADA or a CNADA. The general recommendations may also be applied to non-pivotal studies (e.g., pilot studies and exploratory studies). This guidance does not address the use of adaptive and other innovative designs to support technical sections other than Effectiveness or Reasonable Expectation of Effectiveness. In addition, this guidance describes how sponsors may obtain feedback from CVM on technical issues related to the use of adaptive and innovative designs before the submission of an application.

The decision to use adaptive or innovative elements in a clinical effectiveness study design will depend on a number of factors, including the potential advantages and limitations described later in this guidance. CVM encourages sponsors to explore a variety of design options in the planning of clinical effectiveness study designs, and to proactively discuss their considerations

[2] 21 CFR 514.4; 21 U.S.C. 360ccc(a)(2)B)

with CVM. Some recommendations in this guidance may be technically involved and we recommend you consult with the appropriate CVM experts to help facilitate use of these recommendations.

The following sections outline considerations specific to adaptive and other innovative designs for effectiveness studies for new animal drugs.

IV. Adaptive Designs and Complex Adaptive Designs

For the purposes of this guidance, an adaptive design is defined as a clinical effectiveness study design that allows for prospectively planned modifications to one or more aspects of the design based on accumulating data from subjects in the study. These modifications may affect sample size, intended target population, treatment arm selection, allocation to treatments, endpoint selection, and other design features. A complex adaptive design may include more than one of these modifications.

A. General Recommendations

An adaptive clinical study should be consistent with the following general recommendations when it is intended to provide substantial evidence of effectiveness or a reasonable expectation of effectiveness. There should be procedures to adequately control the chance of erroneous conclusions and minimize the risk of statistical and operational biases. Additionally, the details of the study design should be pre-specified.

Study adaptations from one or more interim analyses can introduce multiplicity concerns, which refers to the potential inflation of the Type I error rate as a result from multiple testing (Type I error is the probability of erroneously concluding an effect when the truth is that there is no effect). Multiplicity concerns can arise from repeated testing of multiple endpoints, subgroups, or interventions (e.g., dosages). All adaptive study proposals should address the possibility of inflating the Type I error probability and control for this error. In order to ensure that the study has adequate power, the proposal should also consider Type II error (Type II error is the probability that a true effect will not be detected by the test). The likelihood of these types of erroneous conclusions should be investigated as part of the examination of the operating characteristics of the design. Analytical statistical methods and simulations can be used to evaluate the operating characteristics of an adaptive study design.

Some adaptive design elements, such as response adaptive randomization, can lead to statistical bias in the estimation of treatment effects and related quantities such as p-values, confidential intervals, etc. Also, operational bias[3] may be introduced when the changes resulting from the interim analysis results are known (e.g., increased sample size) to study participants, resulting in different behaviors before and after the interim

[3] For the purposes of this guidance, operational bias is the bias that arises if some or all participants (e.g., investigators, owners, and caretakers) in the study have access to study results and this information influences the ongoing operation of the study.

analysis. It is important to reduce the potential for statistical and operational bias because these biases can distort the findings of a study and undermine its scientific validity. For some designs there are known methods for adjusting estimates to reduce or remove bias associated with adaptations (Jennison and Turnbull, 1999; Wassmer and Brannath, 2016). However, operational bias may not be quantifiable and cannot be overcome by statistical adjustments to account for its presence. Measures to reduce the potential for bias (e.g., masking procedures and controlling access to interim results) should be developed during study design.

To help avoid such biases, the details of an adaptive design should be completely pre-specified in the protocol prior to initiation of the study. The protocol should describe the number and timing of the interim analyses, the analysis methods to be used, and the type(s) of adaptations, the specific algorithm governing adaptation decision(s), and how information from interim analyses will be controlled. Complete pre-specification is critical to ensure that appropriate statistical methods can be applied to control the chance of erroneous results and provide reliable estimates for treatment effect. Unplanned adaptations based on accumulated data may compromise the validity of the statistical inference and interpretability of the results.

1. Advantages

There are advantages that an adaptive design can provide over a non-adaptive design. Some of these advantages of a successful study adaptation are due to promotion of some of the principles of the three Rs in animal research: Replacement, Reduction, and Refinement (Russell and Burch, 1959).[4]

- Design adaptations can improve efficiency of studies by saving time, money, and resources, as well as supporting the reduction principle for animal research. A study with interim analyses could stop early for effectiveness or lack of effectiveness (futility). In some cases, an adaptive design can provide a greater chance to detect a true drug effect (i.e., greater statistical power) than a comparable non-adaptive design. For example, when assumptions regarding the parameters used for designing the study are somewhat uncertain, the study may benefit from an adaptive approach to re-estimate sample size based on interim data. Sample size re-estimation can correct an under-powered study and improve the chance of detecting a clinically relevant effect, thus potentially avoiding the cost, and further risk to animals, with a new clinical effectiveness study.

- An adaptive design can provide ethical advantages over a non-adaptive design, consistent with the reduction principle. For example, the ability to stop a study early, if it becomes clear that the study is unlikely to demonstrate effectiveness, can reduce the number of animals exposed to the unnecessary risk of an ineffective investigational treatment.

[4] We support the principles of the 3Rs. The concept of adaptation is to potentially reduce the number of animals and to make prospective refinements to the study design to increase inferential value.

- An adaptive enrichment design may make it possible to demonstrate effectiveness in either a given population of animals or a targeted subgroup of that population, where a non-adaptive alternative might require infeasibly large sample sizes, and potentially expose certain subgroups of animals to treatments that are unlikely to be beneficial to those subgroups. (See section V.A. *Enrichment Strategies*.)

2. Limitations

The following are some of the possible limitations associated with an adaptive design study:

- Adaptive study designs that are overly complicated can be difficult to plan. Pre-planning adaptive design modifications can require more effort at the design stage. Additionally, analytical methods to investigate the operating characteristics of the design and establish adequate control of erroneous conclusions may not be readily available or implementable.

- The use of an adaptive design adds logistical challenges in ensuring appropriate study conduct and study integrity. Approaches to appropriately limit access to interim results may be complex and add to study costs.

- If not done correctly, adaptive designs can introduce operational or statistical bias, making it difficult to characterize the true effect of the investigational new animal drug.

- An adaptive change to a study design may lead to results before the adaptation that substantively differ from the results after the adaptation that may lead to challenges in interpretability of the overall results. These differences in the results should be addressed in the final study report. See section VI.B. *Documentation*.

- Adaptations predominantly benefit studies where outcomes are assessed shortly after enrollment so that study changes have time to take effect on future enrollments.

- The maximum sample size and/or study duration with an adaptive design may be greater than a non-adaptive design.

B. Group Sequential Design

Group sequential designs allow the total sample size of the study to be flexible, as data is sequentially evaluated over time. These designs allow for one or more prospective interim analyses of the outcomes that use treatment group information, with pre-specified criteria for stopping the study for success or futility. This design can provide ethical and efficiency advantages by reducing the expected sample size and calendar time of studies.

If there is a real possibility that the treatment may perform better than expected, using a group sequential design may be considered to allow for the possibility of stopping the study early due to successful demonstration of effectiveness. The preferred approach would be to design an adaptive study to allow for an interim look with a possibility of stopping early if there is a sufficient number and composition of subjects to support a scientifically valid inference to the target population.

In a group sequential study, performing each of the multiple statistical hypothesis tests for effectiveness would inflate the Type I error probability and increase the chance of erroneous conclusions. Therefore, the design should include a pre-specified statistical plan that accounts for the interim analyses and makes appropriate adjustments to distribute the alpha so that the overall Type I error is controlled. For example, the O'Brien-Fleming approach tends to require very persuasive early results to stop the study for effectiveness (O'Brien and Fleming, 1979). Approaches such as those proposed by Pocock require less persuasive early results and have higher probabilities of early stopping (Pocock, 1977). The Lan-DeMets alpha-spending approach allows for specifying a function for how the Type I error probability is spent throughout the study, while also allowing for flexibility in determining the number and timing of interim analyses (Lan and DeMets, 1983).

It is important to adhere to the prospective analysis plan and terminate the study only if the stopping criteria are met. In some cases, there should be a limit on how early group sequential interim analyses can occur or whether they should occur at all because a minimum sample size is needed for generalizability of effectiveness results to the target population, for inferential value and independent substantiation of evidence, as well as a reliable evaluation of safety.

C. Sample Size Re-estimation (SSR)

One common study adaptation is to prospectively modify sample size based on interim analysis results. SSR may be conducted with or without using treatment group information (also called unblinded versus blinded SSR).

Adaptive design using SSR can help avoid under-powering studies, particularly in situations where substantial uncertainty exists concerning the variance or effect size. For example, the SSR could indicate that a larger sample size is needed because the effect size is more modest, although still clinically relevant, than initially anticipated. It is crucial that the discussion concerning the clinically important effect size occur during the study planning stage and not after outcome data are available. As a result, an adaptive SSR study design is not intended to fix or salvage a completed study that has failed to provide conclusive results, but instead can help prevent a study from failing to provide conclusive results in the first place.

SSR that does not use treatment group information is generally believed to have limited or no effect on the Type I error probability.[5] However, SSR that uses treatment group information, like the group sequential design, may inflate the Type I error probability. Specifically, if the SSR results in a sample size increase, the Type I error rate may be inflated, depending on the conditional power of the study at the time of SSR and the percent of the sample size increment (Chen et al., 2004; Mehta and Pocock, 2011; Broberg, 2013). There are a variety of existing methods that can be used to appropriately control Type I error probability. For example, hypothesis testing approaches have been developed based on combining test statistics or p-values from the different stages of the study in a preplanned manner or through preservation of the conditional Type I error probability (Bauer and Kohne, 1994; Fisher, 1998; Cui et al., 1999; Denne, 2001; Müller and Schäfer, 2001; Chow and Chang, 2011).

Under certain conditions (Chen et al., 2004; Mehta and Pocock, 2011), SSR without the intention to decrease the sample size may be acceptable without statistical adjustment for Type I error. In this case, the SSR should be performed by an independent party with the only allowable outcomes being stopping for futility, continuing as is, or increasing the sample size.

In addition to Type I error rate control, there are also challenges in maintaining study integrity in the presence of sample size adaptation. Usually, knowledge of the adaptation rule and the adaptively chosen sample size allows at least a qualitative deduction of the interim estimate of treatment effect. Operational bias might be introduced into the study if such knowledge were somehow revealed to the investigators and lead to deviations from the intended study protocol. Therefore, additional steps should be taken to limit personnel with this detailed knowledge so that study integrity can be maintained. For example, one way to reduce the risk of this type of operational bias is by using an independent statistician for SSR in accordance with the prespecified interim analysis procedures.

The additional considerations that were discussed in the section on group sequential designs regarding adherence to the adaptation plan, the evaluation of safety, and generalizability of conclusions also apply to designs with sample size adaptations that use treatment group information. The design should include pre-specification of the statistical hypothesis testing method that will be used and the specific rules governing the sample size modification.

D. Other Study Design Adaptations

Other study design adaptations may be considered as appropriate and are briefly discussed here.

1. Adaptive treatment arm selection

[5] Some exceptions may apply to noninferiority and equivalence tests where SSR, even if it does not use treatment group information, may lead to some inflation of the Type I error rate (Friede and Kieser, 2003).

In cases where more than one dose (or dose regimen) may be effective, an adaptive design with sequential analyses allowing early termination of one or more treatment arms may be more efficient than a non-adaptive design (similar to expansion cohorts described below). Such an adaptive design could, in principle, allow interim modifications to additional aspects of the design, such as the number of additional animals that will be enrolled and the randomization ratio for the treatment arms carried forward.

Statistical hypothesis testing methods (e.g., multiplicity adjustments) should account for the adaptive selection of a best dose or doses from among the multiple doses evaluated in the study, as well as any additional adaptive modifications, such as the potential to stop the study early or to modify sample size.

2. Adaptations to endpoint selection

This is a design that allows adaptive modification to the choice of the primary endpoint based on interim results that use treatment group information. Such a design might be motivated by uncertainty about the treatment effect sizes on multiple treatment outcomes that would be considered acceptable primary endpoints by CVM. As with other adaptive designs, the adaptation rule should be pre-specified, and statistical hypothesis testing methods should account for the adaptive endpoint selection, including appropriate alpha adjustments to control for Type 1 error. Because endpoint selection involves important clinical considerations, early discussions with CVM is recommended when such designs are being considered. In particular, the criteria for selecting the primary endpoint should also consider the clinical relevance of the magnitude of the treatment effects observed at the interim analysis.

3. Adaptations to subject allocation

There are two types of adaptations to subject allocation: adaptations based on comparative baseline characteristics data and adaptations based on outcome data that use treatment group information (e.g., response-adaptive randomization).

A covariate-adaptive treatment assignment is an adaptation intended to promote balance between treatment groups on baseline covariates. In this adaptation, a subject's treatment assignment may depend on its baseline characteristics, (e.g., breed, sex, disease type or stage) and the baseline characteristics and treatment assignments of previously enrolled subjects. Studies that employ covariate adaptive treatment assignments should utilize techniques to reduce predictability of treatment assignments. Additionally, results should be analyzed using randomization or permutation tests or other appropriate statistical methods that control for Type I error probability.

Subjects can also be assigned to treatment using response-adaptive randomization, in which the chance of a newly-enrolled subject being assigned to a treatment arm varies over the course of the study based on accumulating outcome data for subjects

previously enrolled. There are potential statistical, ethical, and pragmatic advantages in a design where a newly-enrolled subject is more likely to be assigned to a treatment arm with favorable outcomes (Rosenberger and Lachin, 1993). However, concerns have been raised about using inconclusive interim results to alter randomization in an ongoing study (Hey and Kimmelman, 2015). Sponsors should discuss with CVM the potential advantages and risks associated with a specific proposed design at the planning stage.

Response-adaptive randomization alone does not generally increase the Type I error probability of a study when used with appropriate statistical analysis techniques. It is important to ensure that the analysis methods appropriately take the design of the study into account.

4. Adaptations in time-to-event studies

There are additional considerations specific to adaptive studies in which the primary endpoint is the time to occurrence of a certain event, such as time to death or time to progression. In these studies, power is dependent on the number of events rather than the number of subjects. It is, therefore, common to target a fixed number of events rather than a fixed number of subjects. Sample size adjustment in these studies has the purpose of modifying the number of events and, therefore, may take the form of modifying the number of subjects, the length of the follow-up period for each subject, or both.

E. Combining Several Adaptive Features

It is possible to employ more than one adaptation in a study. For example, a group sequential design for a study with more than one dose that includes interim looks for potential stopping of treatment arm(s) can also include a sample size reassessment. Typical group sequential testing methods can be used, along with a multiple testing approach to control the Type I error probability across the multiple doses evaluated.

Studies with multiple adaptations should be carefully considered, because the operating characteristics of these complex studies may be difficult to ascertain. Additionally, study logistics and the statistical analysis will also increase in complexity, along with higher risks of operational bias or unforeseen complications in study implementation.

V. Other Innovative Designs

A. Enrichment Strategies

The purpose of enrichment strategies is to select the study population in which the potential effect of a drug can more readily be demonstrated. Enrichment is the prospective use of any characteristic to select a study population in which detection of a drug effect (if one is in fact present) is more likely than it would be in an unselected population. These characteristics are over and above the typical characteristics used as inclusion criteria. Enrichment is directed at improving the ability of a study to detect a drug's effectiveness, resulting in increased study efficiency or feasibility in drug

development. Depending on the enrichment strategy selected, specific label language may be needed regarding the enrichment factor, including appropriate limitations, if the enrichment factor is important for informing the end user of the appropriate use of the drug in the intended target population.

The enrichment strategies are discussed primarily in the context of randomized controlled studies. Enrichment strategies are prospectively planned and fixed prior to study initiation. These strategies generally do not compromise the statistical validity of the studies or the meaningfulness of the conclusions reached for the population studied. A study intended to provide substantial evidence of effectiveness to support approval could include a broad range of subjects but be prospectively designed to evaluate in its primary analysis the effect in the enriched population subset.

The principal concerns with the use of enrichment strategies are the generalizability and applicability of the study results to the population expected to receive the drug once approved (i.e., inferential value to the intended target population). Given the potentially complex interpretation of studies using enrichment designs, plans to use them should be discussed with CVM early in product development.

Three broad goals of enrichment strategies are described listed below.

1. **Decrease variability**

   Approaches to increasing study power (the ability of a clinical study to demonstrate a true treatment effect) by decreasing heterogeneity (nondrug-related variability) include the following:

   - Selecting subjects with consistent baseline measurements used to diagnosis the disease of interest or using baseline measurements within a narrow range. This type of selection is intended to decrease variability within or among study animals;

   - Defining entry criteria carefully to ensure that enrolled subjects actually have the disease that is being studied;

   - Using a challenge infection or infestation in a model study to produce a study population with a more consistent level of disease or infestation;

   - Using placebo lead-in periods before randomization to eliminate subjects that improve spontaneously, have large placebo responses, or improve for reasons other than response to treatment;

   - Identifying and selecting study animals/owners likely to adhere to treatment administration to decrease variability in drug exposure;

   - Excluding study animals unlikely to tolerate the drug;

- Excluding study animals/owners likely to drop out for non-medical reasons (e.g., pet owners who have difficulty bringing the pet to the study site for follow-up evaluations); and

- Excluding study animals with a comorbid illness that would make completing the treatment period unlikely.

Some strategies to decrease variability can result in studies that provide too little inferential value to the intended target population that will receive a drug in clinical practice. Therefore, prior to study conduct, the limitations of using these strategies should be carefully considered and balanced against the need for information in the broader population.

2. Prognostic enrichment strategies

Prognostic enrichment strategies are designed to increase the proportion of study animals likely to have a particular disease-related endpoint event or a worsening condition, in other words identifying study animals at high risk. These strategies allow a treatment effect to be more readily determined. Drugs more frequently evaluated using this design are those intended to slow disease development or progression. Prognostic enrichment is intended to increase the number of events occurring within a shorter time period, generally allowing for a smaller sample size. For any given desired power in an event-based study, the appropriate sample size will depend on effect size and the event rate in the placebo group.

Prognostic indicators include clinical and laboratory measures, medical history, genomic, and/or proteomic measures, among others. For example:

- Animals with cardiovascular disease can have clinical findings more often associated with rapid worsening of their disease;

- Animals with certain stages of cancer that is more likely to progress; or

- Farm history may be used to select a herd at high risk to evaluate drugs intended to control an infectious disease such as bovine respiratory disease.

3. Predictive enrichment strategies

Predictive enrichment strategies are designed to include subjects that are more likely to respond to the treatment than other subjects with the same condition being treated. Identification of a population with a high rate of response or a larger response increases the chance that a study evaluating an effective investigational new animal drug will be able to detect a treatment effect, if one exists. Such selection can lead to a larger effect size and can permit use of a smaller study population than would be needed for a study in an unselected population.

Identifying a more responsive population does not necessarily indicate that a benefit does not exist for the remaining population. It may be desirable to include and collect

data on the non-selected (non-enriched) population to determine the response rate in that population. These data also can provide an assessment of safety in the non-selected population if a broader population of animals is expected to receive the drug after approval.

There are many possible ways to identify study animals more likely to respond to a particular treatment, such as selection of study animals based on a specific aspect of pathophysiology, past history of response to a similar class of drugs, a biomarker, or a disease characteristic that is related in some way to the study drug's mechanism of action. Identification of a responder population may avoid exposure and potential toxicity in study animals that would not benefit from the drug. The strategy may be useful for pilot effectiveness studies or for studies that may demonstrate a reasonable expectation of effectiveness for a conditional approval.

For example:

- Antibacterial drug effects are evaluated only in study animals from which a target organism has been detected. Study animals may have been randomized before the specific infecting organism is known. The detection of the target organism from the study animal is used as a baseline measure or characteristic for inclusion of the case in the final analysis, even though the assessment occurs post-randomization.

- Enroll study animals appearing to have previously responded to a drug in the same class (e.g., dogs that previously responded to an NSAID are enrolled when evaluating an investigational new NSAID in dogs).

- Protein or genetic markers related to a drug's mechanism of action can be used to identify potential responders. For example, a drug targeting the cKit mutation in mast cell tumors may be more effective in the study population with cKit mutations.

- A protein or genetic marker shown to predict response, even without a documented mechanism of action.

Enrichment Study Design Considerations

The study protocol should explicitly describe an enrichment design. For both prognostic and predictive enrichment factors (markers), understanding the accuracy and performance characteristics of the method used to identify subjects or marker-defined subgroups for enrichment is important. In addition to assay validity, for any biomarker used to select subjects, even a familiar one, the biomarker's clinical

sensitivity, specificity, and positive and negative predictive values should be well characterized.[6]

The principal concerns with the use of enrichment strategies are the generalizability and applicability of the study results to the population expected to receive the drug once approved (the intended target population). Enrichment strategies will typically result in a study population that is a well-defined subset of the broader population for which a new animal drug may be intended or in which an approved drug may be used. The expected response rate in the broader population may be different than that in the enriched study population, particularly when the enriched population is expected to be more responsive. The inference of the study results and the best use of the drug in the broader population should be carefully considered.

A critical question when enrichment is used is whether to include the marker-negative population in the study, particularly with predictive enrichment designs. Studies can be designed to include either: (1) only marker-positive subjects; or (2) both marker-positive and marker-negative subjects. If the marker-negative subjects are not expected to respond, their inclusion may dilute the effect response. Also, the presence of significant toxicity associated with the drug should be considered when using it in a marker-negative study population that may not respond to the drug.

B. Designs to Mitigate the Placebo Effect

One significant problem in many placebo-controlled studies for the clinical assessment of investigational new animal drugs intended for symptomatic improvement, such as pain, is a high rate of placebo response, also called placebo effect. High placebo responses may induce false expectations regarding the drug's effectiveness and result in an erroneous conclusion that a drug is ineffective when the drug actually is effective. This section discusses some proposed study designs to mitigate the placebo effect, including placebo lead-in studies and randomized withdrawal studies. In addition, Ivanova and Tamura (2015) proposed a two-way enriched design, which combine the placebo lead-in design and the randomized withdrawal design.

1. Placebo lead-in studies

One possible approach to mitigate the placebo effect is a placebo lead-in design (also referred to as placebo run-in). In this design, all subjects receive placebo first, which is referred to as the lead-in period. After this period, only subjects with signs or symptoms that remain above some threshold value (i.e., placebo non-responders) remain in the study and are randomized to receive drug or placebo. Subjects who show a meaningful symptomatic reduction are excluded from study. The intention is that by eliminating subjects that would respond to placebo before randomization, there should be a reduction in the percentage of subjects responding among those

[6] See CVM GFI #267, "Biomarkers and Surrogate Endpoints in Clinical Studies for New Animal Drugs," (October 2021)

randomized to the post-randomization placebo arm (Faries et al., 2001). Also, through the lead-in period, because many signs and symptoms vary spontaneously, subjects with initial screening values representing random high intensity of the disease condition then showing regression to the mean may be identified and excluded.

2. Randomized withdrawal studies

Another potential design to mitigate the placebo effect is a randomized withdrawal design. In a randomized withdrawal study, subjects that have an apparent response to treatment in an open-label period or in the treatment arm of a randomized study are randomized to continued drug treatment or to placebo treatment. The logic behind the design is that a subject that has shown symptomatic improvement to an active drug is more at risk to lose that benefit when switched to placebo as opposed to remaining on drug (Ivanova and Tamura, 2015). Because such studies generally involve only subjects that appear to have responded, this is a study enriched with apparent responders. A randomized withdrawal design in which the study population is on treatment for an extended duration followed by masked, randomized withdrawal of treatment for a short duration can provide evidence of prolonged effectiveness with only brief exposure to the placebo. Also, this design can allow a subject to be removed from the study (for having reached an endpoint) when the condition returns at a specified severity, avoiding long-term exposure to an ineffective treatment.

C. Expansion Cohorts

Expansion cohort study designs are studies that use multiple, concurrently accruing study animal cohorts, where individual cohorts assess different aspects such as the safety, pharmacokinetics, or anti-tumor activity of the drug. These studies can utilize a single protocol with different cohorts that have cohort-specific objectives. For example, objectives can be to assess anti-tumor activity for different tumor types or stages of disease, assessment of different doses for safety or effectiveness, evaluation of different dosing schedules, or evaluation of a predictive value for a potential biomarker. In general, comparison of activity between cohorts is not planned except where a prespecified randomization and analysis plan are part of the protocol design.

Challenges associated with use of expansion cohorts include timely dissemination of safety information to the clinical investigators including updating owner informed consent forms, exposure of enrolled subjects to potentially toxic or ineffective therapies based on the dose assessed, and potential inappropriate interpretation of study cohort results. Typically, these study designs should not be used for investigational new animal drugs with a narrow margin of safety.

The protocol should incorporate a priori, stopping rules based on lack of effectiveness or toxicity. The protocol should also address planned sample size for each cohort and adaptions to the sample size in each cohort.

D. Master Protocols

A master protocol is a single protocol with multiple sub-studies designed to evaluate multiple hypotheses regarding one or more investigational new animal drugs in one or more (sub)population(s) within the overall study structure. Subpopulations may be defined by disease subtypes, tumor types, histologic types, biomarkers, etc. The study(ies) may be exploratory (dose finding) or support substantial evidence of effectiveness, may have one common control arm or multiple control groups, and may be fixed or adaptive. Examples of master protocols are basket designs and umbrella designs which are described below.

These protocols provide flexibility during the study to allow adaptation such as discontinuation of ineffective therapies, adjustments in randomization, or changes in standard of care. These studies can share a control group which may reduce the number of animals enrolled in the overall study.

Challenges associated with master protocols include potential over-interpretation of the findings due to the multiple study groups; attribution of adverse events to one or more investigational new animal drugs when multiple drugs are administered within arms and the study does not employ a single control arm; or limited assessment of the safety profile of one of the investigational new animal drugs.

1. Basket designs

A master protocol for a basket design evaluates a single investigational new animal drug in different populations defined by, for example, tumor type, disease stage, histology, number of prior therapies, or genetic or other biomarkers.

2. Umbrella designs

A master protocol for an umbrella design evaluates multiple investigational new animal drugs in a single disease population. These studies can use randomized controlled study arms to compare the activity of the investigational new animal drugs with a common control group.

E. Bayesian Adaptive Designs

The term Bayesian adaptive design refers to clinical study designs that use Bayesian statistical reasoning to facilitate the studies in various ways (Berry, et al., 2010). Some examples are:

- Use of Bayesian predictive models to plan the timing and decision criteria for interim analyses;

- Use of external information (e.g., previous studies, observational studies, case studies, or expert opinion) via informative prior distributions to improve the efficiency of a study; and

- Use of posterior probability distributions to form study success criteria.

Different from studies based entirely on Bayesian inference, a Bayesian adaptive design is inherently a frequentist study governed by Type I error control and power consideration. Therefore, the same statistical principles apply to Bayesian adaptive designs as to adaptive designs without Bayesian features. As with any complex adaptive design proposal, early discussion with CVM is recommended for Bayesian adaptive designs that formally borrow information from external sources.

One common usage of most Bayesian adaptive designs is to establish the operating characteristics of the study through numerical simulations. Many simulations in Bayesian methods rely on computationally demanding algorithms such as Markov chain Monte Carlo. It is sometimes advisable to use less resource-intensive techniques such as conjugate priors to overcome the limitation.

VI. Other Considerations

Pre-planning and meetings with CVM – As stated above, adaptive designs should be planned a priori and incorporated into the study protocol. CVM recommends that sponsors discuss these potential designs with CVM early in the development phase of the investigational new animal drug. See section VII. *Obtaining CVM Feedback on Use of Adaptive and Other Innovative Designs*.

A. Protocol

The protocol should clearly specify the type of adaptation or enrichment strategy proposed during the study conduct, provide appropriate methods to control for bias, and control Type I error. If the study will utilize an interim analysis, the protocol should specify if a third party will be involved with assessing the adaptation and whether the assessment will be conducted with prior knowledge of the treatment assignments. If treatment assignments are known by the third party, the protocol should specify how the results will be communicated to the sponsor and clinical investigators to minimize bias and protect masking.

B. Documentation

During the planning and design stage, study simulations are likely needed to choose the number and timing of interim analyses and adaptations, and to evaluate study operating characteristics. These simulations should be carefully constructed over a reasonable range of study scenarios and statistical assumptions to demonstrate that the operating characteristics, for example Type I error and power, meet the desired levels.

It is important that sponsors document the simulations that support the selected design. CVM will review the documentation to confirm that the estimates were taken over a robust set of conditions and may also perform its own simulations.

The final study report should fully detail the adaptive design, enrichment strategies, or other features used in the study and their effect on the interpretation of results. The final study report should address potential reasons and explanations for differences in pre- and

post-adaptation results, such as changes in the study populations or introduction of operational biases. See section IV.A.2. *Limitations*.

An important consideration in the conduct of these simulations is the choice of software. FDA does not endorse any particular software package. However, there are advantages to using commonly available (commercial and non-commercial) software packages that can be freely exchanged or inspected both so that simulation programs can be developed and shared by sponsors and CVM.

VII. Obtaining CVM Feedback on Use of Adaptive and Other Innovative Designs

There are various approaches that sponsors may take to open a discussion with CVM on the use of adaptive and other innovative designs as part of their development program to demonstrate effectiveness or a reasonable expectation of effectiveness. The sponsor's decision regarding which approach to select may be affected by where the project is in the development process. Communication about adaptive and other innovative designs may occur at any point in the development process.

The Office of New Animal Drug Evaluation (ONADE) project managers (PMs) serve as a central point of contact for drug sponsors and can provide information about the new animal drug review process and ONADE's regulatory procedures. If you have questions about the approval process and do not have an ONADE PM assigned to your company, you can contact the PMs through the CVM mailbox AskCVM@fda.hhs.gov.

A. When to submit information regarding the use of adaptive and other innovative designs

There are a variety of points in the development process and a variety of submission types that can be used to obtain feedback. CVM encourages sponsors interested in using adaptive and other innovative designs as part of their development program for a new animal drug to inform CVM as early in the product development process as possible.

Sponsors planning to incorporate adaptive and other innovative designs to demonstrate effectiveness or reasonable expectation of effectiveness are encouraged to inform CVM of their intent either as part of their initial request to open a General Correspondence (GC) file or an INAD file (A-0000), or as part of their initial presubmission conference with CVM to discuss the drug product development plan (Z-submission product development meeting). If one or more studies incorporating adaptive and other innovative designs are already complete, CVM recommends sponsors submit the information described in section VII.B. *How to submit information regarding the use of adaptive and other innovative designs* prior to the initial presubmission conference. While CVM cannot make a determination if existing data satisfies technical section requirements outside of a data submission, if the sponsor submits sufficient information about the existing data early, we can provide feedback to help inform the development

plan.[7] Sponsors are also encouraged to contact their assigned PM for assistance in determining the most appropriate method for obtaining feedback from CVM.

B. How to submit information regarding the use of adaptive and other innovative designs

There are several ways that sponsors may submit detailed information about plans for incorporating adaptive and other innovative designs into their development program to demonstrate substantial evidence of effectiveness or reasonable expectation of effectiveness. The regulatory pathway selected (CNADA versus NADA), the stage of development, the information available, and the feedback being sought from CVM, among other factors, may influence the submission type selected.

Sponsors may seek general guidance on the use of adaptive and other innovative designs in a GC file prior to opening an INAD file, Sponsors may submit information to support use of adaptive and other innovative designs as part of their initial request to open an INAD file; as part of a meeting request for a presubmission conference (Z-submission) to discuss the Effectiveness technical section requirements; or as part of an information submission (H-submission) or meeting request (Z-submission) to discuss study protocol design.

Sponsors considering incorporating adaptive and other innovative designs into future studies to demonstrate effectiveness or reasonable expectation of effectiveness should, prior to conducting a study, submit a study protocol for review (E-submission). Obtaining CVM input regarding study design will make reaching protocol concurrence more efficient.

Sponsors may also open a Veterinary Master File (VMF) to hold detailed information regarding a specific study design, including those regarding pre-investigational discussions about the use of adaptive and other innovative designs, or if the information will be used in the development of multiple applications.[8] The VMF is confidential and is typically used when a holder wishes the material in the VMF to remain proprietary, although the material may be referenced by multiple third-party products or files (INAD, NADA, or CNADA). Alternatively, if multiple sponsors are cooperating on product development, sponsors may establish a Public Master File (PMF) to allow all cooperators to reference the information. As suggested by the name, the information in a PMF is publicly available.

Regardless of how information is submitted to CVM, sponsors should submit an organized and focused information package. This will allow CVM the best opportunity to provide appropriate recommendations in response. Although full information may not

[7] See CVM Program Policy and Procedures (P&P) Manual 1243.2200 Submission and Review of Early Information (EI) Prior to Presubmission Conferences and Protocol Review (June 2020) and CVM P&P Manual 1243.3050 Determining Technical Section Requirements for New Animal Drug Product Approval (May 2019)

[8] See CVM P&P Manual 1243.2400 Veterinary Master Files with Manufacturing Information (August 2019)

be available in the early stages of the development process, the amount of information provided and the level of detail of the information provided should be commensurate with the submission type. The information should address some or all of the following elements, as appropriate for the submission type:

1. The proposed study design that will utilize an adaptive design and how the study fits within the development plan for the product, including a discussion of how an adaptive design would be beneficial. See section IV.A. *General Recommendations*.

2. The type of adaptation(s) proposed (e.g., sample size estimation, dropping treatment arms, determining futility, etc.), and the timing and number of the adaptation(s). The algorithm/criteria for determining the specific change.

3. The operating characteristics of the design and the analytical methods or simulations used to explore these characteristics, if appropriate. The description should provide for adequate control of Type I error and minimize the risk of statistical and operational biases. The programs used should be included.

VIII. Glossary

The following definitions are supplied to provide the reader with an understanding of the specific terms used in this guidance as applicable to new animal drugs. These definitions should not be construed to be new interpretations or clarification of the use of similar words or phrases in the FD&C Act, related code or regulation, other Federal, State, or local laws, or other guidance documents.

A general definition of **bias** can be found in Szklo and Nieto (2000). In this guidance, bias is specifically used in the context of a systematic tendency for the estimate of the treatment effect to deviate from its true value and bias arising from differences in study conduct (operational bias).

Conditional power: The conditional probability of a statistically significant treatment effect at the end of the study calculated based on the results observed at an interim analysis.

A fixed sample trial: A clinical study with a targeted total sample size, or a targeted total number of events, that is specified at the design stage and not subject to prospectively planned adaptation.

Frequentist inference: A type of statistical inference that draws conclusions for the population from sample data by emphasizing the frequency or proportion of the data, using methodologies of statistical hypothesis testing or confidence intervals. In comparison, **Bayesian** inference is a method of statistical inference in which Bayes' theorem is used to update the probability for a hypothesis as more evidence or information becomes available.

In Bayesian statistics, an **informative prior distribution** provides specific, definite information about a variable of interest before a study. The **posterior probability distribution** is the probability distribution of the variable estimated based on the prior distribution and the evidence from the study.

Generalizability: The degree to which inference, based on the study or studies, is applicable to actual conditions of use in veterinary practice or animal management.

An interim analysis: Any examination of data obtained from subjects in a study while that study is ongoing and is not restricted to cases in which there are formal between-group comparisons. The observed data used in the interim analysis can include one or more types, such as baseline data; safety outcome data; pharmacokinetic, pharmacodynamic, or biomarker data; or effectiveness outcome data.

Multiplicity: The potential inflation of the Type I error rate as a result from multiple testing. This arises when the study data are examined and analyzed multiple times during the study without appropriate statistical preplanning.

A non-adaptive trial: A clinical study without any prospectively planned opportunities for modifications to the design.

Operating characteristics: The properties for a given study that describe its performance. For example, properties of interest might include Type I error probability; power; expected, minimum, and maximum sample size; bias of treatment effect estimates; and coverage of confidence intervals (the probability the confidence interval would include the true treatment effect if the clinical effectiveness study were repeated many times). Typically, operating characteristics are examined by simulating studies that use the proposed design under a range of reasonable scenarios (e.g., effect size, sample size, and/or various subject characteristics).

Prospective: For the purposes of this guidance, the term "prospective" means that the adaptation is planned and details specified before any comparative analyses of accumulating study data are conducted. In nearly all situations, potential adaptive design modifications should be planned and described in the clinical effectiveness study protocol (and in a separate statistical analysis plan, if used) before initiation of the study.

Reliability: The extent to which statistical inference from the clinical effectiveness study accurately and precisely evaluates the treatment effect.

In a clinical effectiveness study, a statistical test is applied to assess the strength of evidence against a null hypothesis, where the null hypothesis is typically a statement of non-effect (e.g., no difference between two groups in the study). If the null hypothesis is rejected at a specified level of significance (typically a two-sided level equal to 0.05), with demonstration of a clinically meaningful effect of the drug, the evidence generally supports a conclusion of effectiveness. Sometimes, however, the null hypothesis is rejected even though the drug is ineffective. This is called a **Type I error**. We use the term **Type I error probability** to refer to the maximum probability of rejecting the null hypothesis when it is true (e.g., concluding that there is an effect when the truth is that there is no effect). **Type II** error is the error of not rejecting the null hypothesis when it is false.

IX. References

Bauer, P and K Kohne. 1994. Evaluation of Experiments with Adaptive Interim Analyses. Biometrics 50(4):1029–1041.

Berry, SM, BP Carlin, JJ Lee, and P Muller. 2010. Bayesian Adaptive Methods for Clinical Trials. CRC Press, ISBN 9781439825488.

Broberg, P. 2013. Sample size re-assessment leading to a raised sample size does not inflate type I error rate under mild conditions. BMC Med Res Methodol 13:94.

Chow, SC and M Chang. 2011. Adaptive Design Methods in Clinical Trials. CRC Press, ISBN 9781439839874.

Cui, L, HM Hung, and SJ Wang. 1999. Modification of Sample Size in Group Sequential Clinical Trials. Biometrics 55(3):853–857.

Chen, YH, DL DeMets, and KKG Lan. 2004. Increasing the Sample Size When the Unblinded Interim Result Is Promising. Statist Med 23(7):1023-1038.

Denne, JS. 2001. Sample Size Recalculation Using Conditional Power. Statist Med 20(17-18):2645–2660.

Faries, DE, JH Heiligenstein, GD Tollefson, and WZ Potter. 2001. The Double-Blind Variable Placebo Lead-in Period: Results from Two Antidepressant Clinical Trials. J Clin psychopharmacology 21(6):561-568.

Fisher, LD. 1998. Self-designing Clinical Trials. Statist Med 17(14):1551–1562.

Friede, T and M Kieser. 2003. Blinded Sample Size Reassessment in Non-inferiority and Equivalence Trials. Statistics in Medicine 22(6): 995-1007.

Hey, SP and J Kimmelman. 2015. Are Outcome-adaptive Allocation Trials Ethical? Clinical Trials 12(2):102-106.

Ivanova, A and RN Tamura. 2015. A two-way enriched clinical trial design: combining advantages of placebo lead-in and randomized withdrawal. Statistical Methods in Medical Research 24(6):871-890.

Jennison, C and BW Turnbull. 1999. Group Sequential Methods with Applications to Clinical Trials. New York: Chapman & Hall/CRC Press.

Lan, KG and DL DeMets. 1983. Discrete Sequential Boundaries for Clinical Trials. Biometrika 70(3):659–663.

Mehta, CR and SJ Pocock. 2011. Adaptive Increase in Sample Size When Interim Results Are Promising: A Practical Guide with Examples. Statist Med 30(28):3267-3284.

O'Brien, PC and TR Fleming. 1979. A Multiple Testing Procedure for Clinical Trials. Biometrics 35:549–556.

Pocock, SJ. 1977. Group Sequential Methods in the Design and Analysis of Clinical Trials. Biometrika 64(2):191–199.

Rosenberger, WF and JM Lachin. 1993. The Use of Response-Adaptive Designs in Clinical Trials. Controlled Clinical Trials 14:471-4.

Russell, WMS and RR Burch. 1959. The Principles of Humane Experimental Technique. London: Metheun and Co. [Reissued: 1992, Universities Federation for Animal Welfare, Herts, UK].

Szklo M, and F Nieto. 2000. Epidemiology: Beyond the basics. Gaithersburg, MD: Aspen Publisher, Inc.

Wassmer, G and W Brannath. 2016. Group Sequential and Confirmatory Adaptive Designs in Clinical Trials. Springer series in pharmaceutical statistics. New York: Springer.