STUDY PROTOCOLS

Overview

Protocols describe the design of proposed studies. This chapter focuses largely on the review of efficacy study protocols, but many of its key concepts are also applicable to other types of studies. Additional information regarding the design of other types of studies may be found in the manual chapters for specific topics such as assay validation, safety, or stability.

Firms are strongly encouraged to submit protocols for all licensing studies at least 60 days prior to initiating the study so that the CVB can provide comments on the proposed study design and analysis. This is the CVB's primary opportunity to guide the firms in designing appropriate studies that will generate scientifically meaningful data of acceptable quality. The CVB may comment on the protocol or, if necessary, ask for revisions. The CVB also may give some indication of what the applicant may expect from the CVB in response to a report of the completed study. To expedite review, firms should provide a list of differences from similar previously approved protocols, as indicated in CVB Notice 13-17.

Flow of Information

1. Protocols that involve statistical design, analysis, or inference should be routed to Statistics as soon as they are received. Protocol review is the Statistics section's #1 priority; turnaround time is generally less than a week (target= 2 days). See the Statistics chapter for details on routing submissions.

If a protocol follows a codified study design (e.g., 9CFR standard requirements for immunogenicity), then full statistical review may not be necessary. Reviewers should, however, contact the Section Leader for Statistics and explain the nature of the protocol. The Section Leader can then determine whether Statistics review is needed.

2. When the statistician has completed his/her review, the "route to Statistics" child loop is closed, and statistics comments are uploaded to the mail log submission.

Reviewing Study Protocols

The considerations for a particular protocol are highly dependent upon the type of study being performed. The following is a list of general considerations, not all of which apply to every study.

1. **Objective**

The protocol should clearly state the nature and purpose of the study. For confirmatory studies intended to support specific label claims, the proposed claim should be stated and an appropriate primary outcome should be specified. Determine whether the primary outcome could support the proposed claim. (Example: a protocol in which the primary outcome measures disease severity might support an "aid in the control/reduction" claim but would not be suitable if the firm wishes to obtain an "aid in the prevention of disease" claim.) Exploratory studies do not need to be as explicitly described or rigorously designed as confirmatory studies. Some protocols, however, propose study designs that blur the line between exploratory and confirmatory objectives. This may be done with the hope that if the results are "good enough" they could be used to support licensure. Since some prior knowledge or experience with the proposed clinical model is usually necessary to adequately design a confirmatory study, appropriate expectations should be communicated to the firm.

2. **Primary outcome**

Specifying the primary outcome is a critical early step in study design, and this is currently a prime area of focus in protocol review. (Also see VS Memorandum 800.202 for additional information.) Statistics will routinely comment on the proposed primary outcome.

Outcome. An outcome is an observation on an individual subject that includes the clinical event and the unit of measurement.

> Example. If the clinical event is diarrhea, the outcome may be specified as
> 1. Presence or absence of diarrhea at any time during the observation period,
> 2. Duration of diarrhea in days,
> 3. Severity of diarrhea by an explicit ordinal categorization (e.g. severe – explosive and bloody; moderate – watery; mild – unformed; normal.)

A study designed to support a claim related to disease prevention must have a case definition. The case definition determines whether or not an animal is affected by challenge, i.e. represents a case of disease. (E.g. #1 in the preceding example.)

When duration is the specified outcome, it is recommended that observations are made at equally spaced intervals. However, there are times when duration is calculated when observations are not equally spaced. Duration is defined as the time lapse between the first observation and the last observation. Each observation is made at the center of the interval. When the intervals span more than one day and are not evenly spaced, the initial observation is [the study day of the first positive – ½ the distance (in study days) to the previous observation], and the final observation is [the study day of the final positive + ½ the distance (in study days) to the next observation]. When observations are made daily duration reduces to first observation – last observation + 1. To

determine duration, the observation period must continue until clinical signs have ended. Consequently, if clinical signs are still present at the final observation time, duration cannot be determined.

Case definition. Disease case definitions should be explicit, clinically meaningful, and have a natural interpretation as representing a case of a particular disease. An arbitrary dichotomization of disease severity should not be presented to support a claim related to disease prevention. In other words, if what the proposed "case definition" actually does is differentiate between mild and severe disease, it is really a category definition rather than a case definition, and it does not support a claim related to disease prevention. On the other hand, in some circumstances it might be appropriate for animals displaying transient, clinically trivial signs to be considered unaffected, and in that case a prevention-type claim could be supported. Consequently, reviewers must use good judgment in deciding whether a proposed case definition could support a prevention-type claim.

a. It is sometimes necessary to define a compound case definition based on a combination of clinical signs.

> Example: animal is affected if it has a fever >2° above baseline AND concurrently develops mucopurulent nasal discharge

Any time that a compound case definition is created, there may be circumstances under which an animal does not technically meet the stated case definition for disease but realistically is affected. Give careful thought to scenarios under which this might occur.

> Example: animals spiked a fever one day before the nasal discharge (which persisted 5 days) became evident

Pay particularly attention to case definitions that call for combinations of arbitrary amounts such as the following.

> Example: An animal will be considered affected if it has at least two of the following: 3 daily observations of coughing, 2 daily observations of mucopurulent nasal discharge, 1 daily observation of fever.

A case definition is defective if there is a good possibility that an animal could escape the case definition yet still be clearly affected by disease. If so, comment on this in your response letter. Although the CVB is never bound to accept all possible eventualities in a study, it is best to provide feedback at the protocol stage if exceptions to the case definition are sufficiently likely to affect the CVB's conclusions regarding the study. Consider also its impact on the study design as a whole.

Outcome specification. When dealing with clinical observations, it is also important to consider how clearly the criteria defining the outcome are described. How likely is it that independent

scorers would consistently reach the same conclusion? Is the scale/categorization appropriate for the measurement, or does the scale/categorization suggest more precision than is practical for the observation?

Example:
Worst: Diarrhea will be ranked on a scale from 1 to 10. (Is it reasonable to expect an observer to differentiate among 10 degrees of severity for diarrhea, even if there was an attempt to provide objective criteria for each?)

Better (but still needs improvement): Diarrhea is scored as mild, moderate, or severe. (Three categories are more appropriate, but the protocol still lacks any objective means to determine which category is applicable. One person's "mild" may be the next person's "moderate.")

Best: Diarrhea is categorized in one of the three following categories: unformed semi-solid, totally liquid (no blood), or bloody liquid. (This protocol provides distinct criteria that are likely to be applied consistently by any reasonably trained observer.)

3. **Subject selection & randomized assignment to treatment groups**

The protocol should describe what type of animals will be used in the study and how they will be randomized to treatment groups.

- *Age*. Are the subjects all of similar age or a wide range of ages? If age may be a factor in response to vaccination and challenge, this should be considered when randomizing treatment groups. The age of the animal will also affect label claims—the age of the animals in an efficacy study determines the minimum age for which the product may be recommended. Field safety studies should include an appropriate number of minimum age animals (proportion varies depending on animal species). Historically, many labels did not specify a minimum age for administration, but all new products should include this information and it should be linked to the efficacy and safety studies.
- *Gender*. With the exception of vaccines directed towards gender-specific diseases (e.g., fetal abortion due to IBR), vaccine efficacy is not expected to differ between genders. Representation of multiple genders in the study is preferable, but not required. It is not uncommon for safety studies to be gender specific, particularly if they involve reproductive failure. If reproductive failure is a potential issue, animals representative of all three trimesters of gestation should be included. Also consider related issues like lactation status and parity where applicable.
- *Breed*. Safety studies ideally should utilize a wide variety of breeds.
- *Source*. Animals may be obtained from the sale barn for large animal studies. If several different sale lots are combined, it's important to consider source when randomizing animals to treatment groups.
- *Familial relationship*. The litter from which an animal is derived often has a material impact on the response to vaccination and/or susceptibility to challenge. In these situations, restricting the randomization of animals to treatment groups based on litter (often referred to as 'blocking on litter') results in a more precise estimate of vaccine efficacy, and thus is generally encouraged.

- *Experimental unit*. The study should clearly specify the experimental unit. In studies of passive maternal immunity, the experimental unit is the dam, and littermates are considered to be replicates of a single unit and not independent. In active immunity studies, each test animal is generally considered to be an experimental unit when animals are randomized appropriately to treatment group and housed in a manner that allows an independent assessment of efficacy for each animal. As a general rule, for an individual animal to be the experimental unit in challenge studies, treatment groups should be commingled, or animals should be housed individually prior to, and during, the challenge phase of the study. If the treatment groups are not commingled and animals are not housed individually but in shared housing units such as cages or pens, there must be replication of housing units within each treatment group. In aquaculture studies, many times the experimental unit is the tank, and not the individual fish.

Immunologically naïve animals should be used for efficacy and reversion to virulence studies. If the firm states that naïve animals cannot be found, ensure that they have checked beyond their usual source herds before approving protocols that utilize seropositive animals. Cross-check similar, recent studies from other firms to see if others have had similar problems. If seropositive animals must be used, ensure that the randomization and statistical analyses take this into account.

The protocol should specify how the exposure status of animals is determined prior to enrollment in the study. If the protocol states that the animals will be seronegative, the type of assay should be specified. (For certain diseases, animals may have a low (nonspecific) titer and still be considered "seronegative"; ensure that any arbitrary cut-off points are reasonable.) A study is often strengthened if additional measures, such as negative virus isolation results, are combined with the results of serology to demonstrate that the animals are naïve.

The method of randomizing animals to treatment groups should be explicitly described. Some firms may think that they are randomizing when actually they are not. Example: Gate cutting is not randomization. Randomization to housing unit, when applicable, should also be described.

Randomization that considers only the treatment assignment is termed simple randomization. Often additional blocking factors must be considered in the randomization scheme. In swine studies, for example, litter and pen are factors that often must be considered in the randomization scheme. By default, the CVB Statistics Section considers both litters and pens as possible stratification variables in the analysis. It is important for the study design to plan for conducting randomization with both factors kept in mind. To keep things simple, a good strategy is to house littermates together throughout the study. In studies where it is not possible to house all littermates together, one solution is to create blocks consisting of litter × pen and randomize to treatment within the litter × pen blocks. These blocks can also be randomly assigned to pens. Ideally the plan should include as much balance in the study as possible. For instance, every litter is represented in 2 of the 4 pens in the study, and each litter × pen combination contains about the same number of vaccinates and controls.

4. **Blinding (Masking)—Lack of bias**

Personnel making any kind of observations (including laboratory results) during the study should be suitably blinded. Blinding should occur on two levels—the observer should not know which treatment an individual animal received, and they also should not know which animals belong to the same treatment group. Knowledge of either has the potential to bias results.

The housing of animals can, at times, preclude blinding (e.g., housing animals vaccinated with a modified live vaccine separately from sham-vaccinated controls). In such cases, the individuals involved in the assessment of vaccine efficacy after challenge should not be involved in the vaccination phase. Other alternatives often exist as well.

When a study cannot be blinded (e.g., field safety study where the whole herd is vaccinated), those making critical observations should not have a vested interest in the outcome of the study. [redacted]

When the measurements are extremely subjective and based on a scale that gives the illusion of more precision than actually exists (e.g. lung lesion scores measured as a visual estimate of percent of lung involved), the study is greatly strengthened by using multiple blinded scorers working independently. [redacted]

Blinding of necropsies or laboratory assays often requires that specimens be handled in random order. Potential bias can also arise from the clustering effect of running many specimens on a single device (e.g., don't run all of the vaccinates on one plate or gel and all of the sham-vaccinated controls on another).

5. **Acceptable controls**

Studies should employ the proper controls. Vaccination-challenge studies should include a group of animals that are sham vaccinated before they are challenged. The use of a placebo vaccine is critical, unless the firm can justify why it is not necessary. Placebo vaccines may be immunologically inert (e.g. saline), antigen deficient (e.g. adjuvant), or product-matched. Product-matched placebos include all vaccine components except the antigen being assessed in the study. Product-matched placebos are particularly desirable for vaccines containing multiple antigens that protect against diseases of a similar nature (e.g., *Mannheimia haemolytica* and *Pasteurella multocida* pneumonia, both assessed by percent lung consolidation). The placebo should be physically indistinguishable from the experimental product.

Similarly, the laboratory assays used within a study also should have proper controls. This aspect is often overlooked in study protocols, as often only the type of assay (e.g., serum neutralization assay) is specified. If the lab result is critical to the outcome of the study (e.g., a label claim related to virus isolation after challenge), it is important to ensure that the assay has

been acceptably validated [redacted]

6. **Statistical power**

Since animals are costly and animal care concerns are increasing, firms often attempt to minimize the number of animals in a study. Statistics will inform the reviewer if the study is grossly underpowered to achieve the desired conclusion. Sometimes, however, it is *possible* to achieve the desired outcome with the proposed group sizes, but achieving the outcome would be more *likely* with larger groups. In general, the more variable the expected response, the larger the group sizes must be to provide the desired precision (as measured by confidence interval) for the estimate of vaccine efficacy. It may be prudent to recommend enrolling a few additional animals to reduce the risk of failing to achieve the necessary precision.

A well-designed and executed study need not be a total loss, however, simply because it was too small. The findings can be analyzed together (not pooled) with the findings of a subsequent study. Multiple smaller studies may in fact provide greater evidence than a single larger study, provided the studies were designed and conducted similarly and the results are reasonably concordant.

Be aware that the CVB often requires more animals than the bare minimum necessary to achieve statistical significance. If the number of animals proposed for a study seems excessively small, cross-check other similar studies to ensure that the firm is proposing a study consistent with what the CVB has expected of others. As an *extremely general* rule of thumb, prelicense efficacy studies in mammals generally require a minimum of approximately 20 independent vaccinates; many more are required for poultry and fish.

When animals are handled or housed in clusters (e.g. herd, pen, isolator), there *must* be replication of the clusters within treatment groups. Otherwise, it is impossible to differentiate potential cluster effects from treatment effects. It is more important to increase the number of clusters than the number of subjects. For example, one isolator of 36 birds per treatment is unacceptable. Three isolators of 12 birds are better, but six isolators of six birds would be even better.

If the firm wishes to demonstrate efficacy by two routes of administration, they are permitted to divide the vaccinates and vaccinate half by each route. If the study results demonstrate that there is no material difference in the response to each route of administration, then the power of the subgroups may be combined for statistical analysis, *provided that* significant efficacy is shown by each route separately. If, however, one route proves to be superior to the other, then the data from each route cannot be combined. Moreover, most "split-route" studies are not powerful enough to ensure that estimated efficacy can be distinguished between routes.

> Example: Number of sick: 8/10 nonvaccinates, 2/10 IM vaccinates, 5/10 SC vaccinates. Prevented fractions: IM = 0.75 (*95% CI:* 0.26, 0.93), SC = 0.38 (*95% CI:* –0.24%, 0.72), combined = 0.56 (*95% CI:* 0.12 , 0.79). There is not enough power to determine whether the

SC route is efficacious ***or*** if efficacy by the IM and SC routes is the same. Consequently, it would be reasonable to run another small study for SC and use the results of both studies.

It is often prudent to remind firms of these potential pitfalls when they elect, at their own risk, to use small subgroups.

7. **Confirming etiology of post-challenge disease**

The study protocol should include procedures to demonstrate that disease observed after experimental challenge is due to the challenge organism and not a coincidental co-infection. (This is especially important in large animal studies that utilize animals from sale barns or sources with unknown disease status.) If the disease under study causes pathognomonic signs or lesions (e.g., diamond skin lesions of erysipelas in pigs), the presence of the lesions themselves may be acceptable evidence of etiology. Far more commonly, however, the primary variable in a vaccination-challenge study is a composite of nonspecific clinical signs (e.g., cough, nasal discharge, fever). In such cases, other methods, such as virus or bacterial isolation, should be included unless otherwise justified.

8. **Measuring effect of challenge**

The protocol for an efficacy study should define the minimum disease that must be observed in sham-vaccinated controls for a valid study. The CVB requires results to be clinically relevant as well as statistically significant (sufficiently precise); thus, the challenge must have elicited a sufficient amount of disease so that a meaningful degree of protection afforded by vaccination can be demonstrated.

It may be difficult to confirm how much of the post-challenge pathology is due to the challenge if the possibility exists that the animals may have had pre-existing lesions. This is the case in many large animal respiratory disease studies where the primary variable is lung lesion scores. Animals enrolled in these studies may have pre-existing lung consolidation, and it is important to know the "baseline" level of lung pathology in the herd before making conclusions about the effect of the experimental challenge. In these cases, it is beneficial to include a nonvaccinated, nonchallenged control group to serve as a baseline for comparison. This control group is not only beneficial for confirming the validity of the challenge; it can be useful for the firm to explain an apparent vaccine "failure" if the observed lesions are not related to the study.

The post-challenge observation period should be sufficient to evaluate the entire course of disease. In general, the CVB does not grant label claims for products that merely delay the onset of disease or delay the time to peak disease severity. Post-challenge observations should continue until all relevant clinical signs have resolved in the surviving vaccinates. <u>This is especially critical for products being considered for claims of reduced disease severity or duration.</u> The proposed post-challenge observation period in the study protocol should be considered to be a *minimum* period, subject to extension as actual study conditions dictate.

9. **Location and critical dates of study**

The protocol should list specifically where the study is to be performed. Pivotal studies may be performed on licensed premises, or firms may utilize independent contractors, provided that the firm retains oversight of the study. Generally, pivotal studies should be conducted within the US, to ensure that the results are representative of disease conditions, genetics, and animal husbandry practices in this country. Requests to conduct studies elsewhere should be considered on a case-by-case basis. It is the responsibility of the firm to justify why a foreign locale is equivalent to the US. The justification should include evidence that:

- the challenge organism is representative of isolates found in the US
- the animal husbandry practices are equivalent to those found in the US
- the animal breed(s) are not materially different (for purposes of the study) than those in the US
- the study facilities are adequate
- personnel associated with the study are adequately qualified
- quality assurance procedures are adequate

In general, we must have no reason to suspect that the quality of the study (and its results) will be inferior to, or less meaningful than, those that might be generated domestically. Such determinations must be taken with due care and are left to the discretion of the reviewer. The reviewer should discuss proposed foreign locales at a reviewer staff meeting prior to giving approval to the firm.

The firm is expected to provide the CVB with critical study dates so that the CVB may, at our discretion, observe key parts of the study (CVB Notice 02-01). This information may be included in the protocol submission, but it also may be submitted separately after the protocol has been reviewed if the dates are not known well in advance.

10. **Statistical analysis methods**

This is covered more completely in the Statistics chapter. In addition to design issues, Statistics review of the protocol will often address discrepancies between the firm's proposed analyses and those that will likely be used by the CVB. These differences should be included in the response letter to the firm. In the past firms tended to rely excessively on p-values, and a p-value of 0.05 was the magical criterion against which most studies were judged. CVB currently emphasizes estimating the vaccine effect (e.g., prevented or mitigated fraction) and basing conclusions on its size and clinical relevance. Estimation also encourages better study design in general.

11. **Data capture forms/quality assurance**

Considerable insight into how the study will be conducted and recorded can be gained by reviewing data capture forms. This is especially true in the case of field studies, where the degree of quality assurance is evident in the forms that field cooperators must complete. Beware of forms with scant instruction and vague definitions—if they are not clear to you, they likely will not be clear to a cooperator who has far less understanding of quality assurance principles. General things to consider:

- Does the form capture all of the relevant observations?
- If the form is to be used by a contracting cooperator or an animal owner, does it contain sufficient instruction for the person to fill it out accurately?
- Is there a place to record negative, as well as positive, findings?
- Is documentation in place to account for every test subject?
- Does the form capture who made the observations and when?

CVB Response to Protocol Submissions

In responding to a protocol, do not use language that suggests blanket approval (such as "the protocol is acceptable"). Give relevant comments on various aspects of the protocol. Then wrap up with a conclusion such as one of the following.

Suggested Conclusions

1- For these reasons, it would be a good idea to submit a revised protocol before proceeding with the study.

2- The protocol is sketchy and there is insufficient detail to evaluate the likely suitability of the study findings. While you may elect to proceed with the study, you should have no particular expectations regarding CVB's response to the study report.

3- If the study is properly conducted with the recommended modifications, it is possible that the study findings will be suitable for evaluating [whatever]. That expectation does not necessarily imply CVB approval of all possible eventualities.

4- The protocol is fairly clear and complete. If the study is properly conducted, it seems likely that the findings will be suitable for evaluating [whatever]. This prospect does not necessarily imply CVB approval of all possible eventualities.