

B. COLLECTIONS OF INFORMATION EMPLOYING STATISTICAL METHODS

- 1. DESCRIBE (INCLUDING A NUMERICAL ESTIMATE) THE POTENTIAL RESPONDENT UNIVERSE AND ANY SAMPLING OR OTHER RESPONDENT SELECTION METHOD TO BE USED. DATA ON THE NUMBER OF ENTITIES (E.G., ESTABLISHMENTS, STATE AND LOCAL GOVERNMENT UNITS, HOUSEHOLDS, OR PERSONS) IN THE UNIVERSE COVERED BY THE COLLECTION AND IN THE CORRESPONDING SAMPLE ARE TO BE PROVIDED IN TABULAR FORM FOR THE UNIVERSE AS A WHOLE AND FOR EACH OF THE STRATA IN THE PROPOSED SAMPLE. INDICATE EXPECTED RESPONSE RATES FOR THE COLLECTION AS A WHOLE. IF THE COLLECTION HAD BEEN CONDUCTED PREVIOUSLY, INCLUDE THE ACTUAL RESPONSE RATE ACHIEVED DURING THE LAST COLLECTION.**

Large swine operations (1,000 head or more) component

The potential respondent universe of the large swine operations component of this study is all operations with 1,000 or more hogs in 13 States¹. The study population is limited to operations with at least 1,000 hogs in those States because the goal is to focus on operations with large commercial components. Operations with fewer than 1,000 hogs will be the focus of the small enterprise component of this study. The large enterprise swine operations component will focus on management and health issues more pertinent to operations with larger hog inventories. Operations and sites with 1,000 or more hogs account for approximately 15.6 percent of swine operations and sites in the U.S. but 97.2 percent of the swine inventory resides on these operations and sites (see Table A.1 in Appendix A for more information).

The 13 specific states were chosen because they account for 95.7 percent of operations and sites with 1,000 or more hogs and 93.0 percent of the hogs on operations and sites with 1,000 or more hogs in the U.S. Also, all 13 of the chosen States are included in the National Agricultural Statistics Service (NASS) Quarterly Hogs and Pigs Report and will therefore have the most complete population-level operation and inventory information available. States were chosen based on their percentage of hog inventory and hog operations, for their geographic representativeness, and expected response burden. See Table A.2 in Appendix A for operations and inventory estimates.

The swine industry tends to be vertically integrated within businesses. This means that there may be an operation that consists of a number of separate locations on which hogs are raised. Due to the nature of the NASS List Frame, operations are included on the frame, but sites are not. Thus, this component of the study will involve a two-stage sampling process. First, operations (business structures, contractors, independent operators) are chosen from the NASS list frame. Then, operations are visited and administered a short questionnaire in order to determine the number of sites that raise hogs owned by that operation in that state, and sites are then sub-sampled and administered the full study questionnaire, if willing.

¹ Illinois, Indiana, Iowa, Kansas, Michigan, Minnesota, Missouri, Nebraska, North Carolina, Ohio, Oklahoma, Pennsylvania, and South Dakota

The sampling design will be both multi-phase and multi-stage. There will be Phase I of the study, which consists of sampling Stages I and II, during which the primary data collection instrument will be the 2020 NAHMS Swine Large Enterprise Survey (VS Form 21-201). Phase I will be followed by Phase II of the study, during which the primary data collection instruments will be the 2020 NAHMS Swine Large Enterprise – VS Visit (VS Form 21-203) and the collection of biological samples (the 2020 NAHMS Swine Large Enterprise - Fecal Collection Form and the 2020 NAHMS Swine Large Enterprise - Saliva Collection Form, VS Form 21-205 and VS Form 21-206, respectively).

For Phase I, Stage I of the study, all of the operations on the NASS List Frame with 1,000 or more hogs in the 13 states will be selected and, thus, will be a census and not a sample of the large operations. The number of these operations totals approximately 2,700, but will have been updated by the time sampling occurs. Having an up-to-date list frame will help to ensure the most accurate number of operations with 1,000 or more hogs are included on the frame and also be the best chance to avoid out of scope, out of business, or otherwise operations with no animals present.

During the Phase I, Stage II of the study, operations selected during the first sampling phase from the NASS List Frame will be contacted by a NASS enumerator to set up a face-to-face interview using the short 2020 NAHMS Swine Large Enterprise Survey – Site Selection Form (VS Form 21-200). Using the information captured on the 2020 NAHMS Swine Large Enterprise Survey – Site Selection Form, the NASS enumerator will sub-select a number of sites that raise hogs for that operation within that state using simple random sampling. Simple random sampling of sites will be accomplished via random number generation from a website (www.random.org) and/or through selection of numbers from a random number table supplied to the enumerator (both approaches would include sufficient instructions and training to select random numbers using the chosen procedure). The number of sites sub-selected will be a function of the number of sites raising hogs under that operation (see Appendix B for more information).

Baseline response rates are expected to be approximately 54.0 percent for Phase I of the study and 51.0 percent for Phase II. The assumed response rates are taken from the NAHMS Swine 2012 study for operations with 2,000 or more hogs. The NAHMS Swine 2012 study was more similar to the NAHMS Swine 2020 study in design and material than was Antimicrobial Use (AMU) in Swine 2017. Using the response rates for operations with 2,000 or more hogs is more appropriate than the overall response rates because larger operations tend to respond at different rates than smaller operations (see Appendix D for more information on response rates). These response rates reflect the proportion of operations with complete survey data at each of the two study phases with respect to all operations sampled.

This estimate is used rather than a ratio of operations with complete survey data to in-scope operations because it is more useful in judging the actual study activities because it takes into account the extra effort exerted when out-of-scope operations are selected from the sampling frame. This estimate gives the most accurate representation of the study effort and costs, even though it is lower than the estimate after adjustment for out-of-scope or anticipated out-of-scope operations.

To increase response rates, a number of steps are being taken in the 2020 study. The questionnaires will be thoroughly reviewed and will be pretested to ensure a minimal amount of information is collected to meet study objectives using an effective instrument. Participants will be given pre-survey materials familiarizing them with the study, its potential benefits to them and their industry, and what would be expected of them. All enumerators will be trained on data and information security and applicable laws protecting respondent data and will be trained on data collection specific to the study. The sample will be taken from the NASS list frame after having been updated using information from the 2017 Census of Agriculture, which will likely result in a greater proportion of in-scope contacts being selected from the frame. Also, a new incentive (oral fluids biological sampling) is being offered. More efforts to maximize response are given under item 3 below.

Small swine operations (less than 1,000 head) component

The potential respondent universe of the small swine operations component of the study is all operations with fewer than 1,000 hogs in 38 states. The study population is limited to operations with fewer than 1,000 hogs because the goal is to focus on operations marketing swine products in niche markets. Operations with 1,000 or more hogs will be the focus of the large swine enterprise component of this study. The small enterprise swine operations component will focus on management and health issues more pertinent to operations with smaller hog inventories. Operations with fewer than 1,000 hogs account for approximately 2.8 percent of swine inventory in the U.S. but 84.4 percent of swine operations (see Table F.1 in Appendix F for more information).

The 38² specific states were chosen because they account for 95.2 percent of operations and sites with fewer than 1,000 hogs and 96.4 percent of the hogs on operations with fewer than 1,000 hogs in the U.S. The goal of representing at least 95 percent of operations and sites with fewer than 1,000 hogs and at least 95 percent of the swine inventory on operations with fewer than 1,000 hogs in the U.S. was set (rather than representing at least 70 percent of operations and inventory) because this segment of the swine population is studied less frequently and with less depth than operations with 1,000 or more hogs, which make up a larger portion of the commercial population of swine. States were chosen based on their percentage of hog inventory and hog operations, for their geographic representativeness, and expected response burden. See Table F.2 in Appendix F for operations and inventory estimates.

Based on unpublished NASS data, and unlike the large operations component, there was a small number of swine operations with fewer than 1,000 hogs owned for which hogs were raised on multiple sites. Thus, to minimize the effort spent to sub-select a number of sites from that operation, the number of sites sub-selected will be limited to one in all cases in which hogs are raised at more than one site. In the event that hogs are raised on more than one site, data will be collected on the closest site at which hogs and/or the site is owned by the operator.

² Alabama, Arizona, Arkansas, California, Colorado, Florida, Georgia, Idaho, Illinois, Indiana, Iowa, Kansas, Kentucky, Louisiana, Maine, Massachusetts, Michigan, Minnesota, Mississippi, Missouri, Montana, Nebraska, New York, North Carolina, Ohio, Oklahoma, Oregon, Pennsylvania, South Carolina, South Dakota, Tennessee, Texas, Utah, Vermont, Virginia, Washington, West Virginia, and Wisconsin

The sampling design will be two-stage, though there will only be one survey. Operations with fewer than 1,000 hogs, and especially operations with fewer than 100 hogs, tend to enter and exit the industry more frequently than the operations with larger inventories and will thus be more likely to have no hogs, be out of business, or otherwise inaccessible on a given day than larger operations. In this study, NASS will sample operations and will send the 2020 NAHMS Swine Small Enterprise Survey (VS Form 21-207) and will follow up with operations with unreturned questionnaires via a computer-assisted telephone interview (CATI) to determine whether the operation is still in business and has between 1 and 999 hogs on the operation and, if so, will administer the questionnaire over the phone.

Baseline percent of eligible operations is expected to be approximately 57.4 percent and 71.8 percent of the eligible operations are expected to complete the 2020 NAHMS Swine Small Enterprise Survey. The assumed most representative completion rates are averages of completion rates taken from the previous NAHMS Swine Studies (see Appendix H for more information on eligibility and completion rates). These rates reflect the proportion of operations with complete survey data with respect to all operations sampled.

This estimate is used rather than a ratio of operations with complete survey data to in-scope operations because it is more useful in judging the actual study activities because it takes into account the extra effort exerted when out-of-scope operations are selected from the sampling frame. This estimate gives the most accurate representation of the study effort and costs, even though it is lower than the estimate after adjustment for out-of-scope or anticipated out-of-scope operations.

To increase response rates, a number of steps are being taken in the 2020 study. The questionnaires will be thoroughly reviewed and will be pretested to ensure a minimal amount of information is collected to meet study objectives using an effective instrument. Participants will be given pre-survey materials familiarizing them with the study, its potential benefits to them and their industry, and what would be expected of them. All enumerators will be trained on data and information security and applicable laws protecting respondent data and will be trained on data collection specific to the study. The sample will be taken from the NASS list frame after having been updated using information from the 2017 Census of Agriculture, which will likely result in a greater proportion of in-scope contacts being selected from the frame. Skip patterns in the questionnaire will be used in order to identify which producers have the requisite number of hogs on hand at the time of the study and to reduce burden on those not considered in-scope. More efforts to maximize response are given under item 3 below.

The proposed paper and electronic (where applicable) questionnaires will be pretested by APHIS prior to field enumeration, involving 3 operations. It is planned that NASS will perform pretests of the Phase I questionnaires on 6 of operations and expert reviews by 8 reviewers.

Results of these pretests will be utilized to refine the questionnaires in order to reduce respondent burden and improve the usefulness of the information. Many revised questions from previous iterations of the NAHMS Swine study will be used. Final questionnaires will have been reviewed by a variety of experts, including academic researchers, individual representatives, extension agents, veterinarians, health specialists, epidemiologists, survey methodologists, and survey administrators.

2. DESCRIBE THE PROCEDURES FOR THE COLLECTION OF INFORMATION INCLUDING:

- STATISTICAL METHODOLOGY FOR STRATIFICATION AND SAMPLE SELECTION;

Large swine operations (1,000 head or more) component

Because all operations with 1,000 or more hogs will be selected from the NASS List Frame during Phase I, Stage I of the study, no sample selection will occur at this stage. During Phase I, Stage II of the study, sites within operations in the given State will be selected using simple random sampling and will thus not be stratified within operation.

Sites will be post-stratified in order to properly account for non-response among like operations and will take place along regional and size of operation (inventory of hogs) boundaries, of which there will be three levels of each. Regions will be defined as:

- Midwest (Iowa, Michigan, Minnesota, Nebraska, and South Dakota),
- East (Illinois, Indiana, Ohio, and Pennsylvania), and
- South (Kansas, Missouri, North Carolina, and Oklahoma).

Size categories will be defined as:

- Small (1,000-1,999 hogs),
- Medium (2,000-3,999 hogs), and
- Large (4,000 or more hogs).

If sampling at the operation level occurs, operations in strata with more hogs (as opposed to more operations) will be sampled at a higher rate to prioritize sampling larger operations to ensure they are included in the sample.

Small swine operations (less than 1,000 head) component

Operations will be selected from strata defined by State and size category. Size categories will be defined as:

- Very small (1-24 head),
- Small (25-49 head),
- Medium (50-99 head),
- Large (100-499 head), and
- Very large (500-999 head).

Operations in strata with more operations (as opposed to more inventory) will be sampled at a higher rate to prioritize sampling smaller operations.

We will properly account for non-response among like operations and will take place along regional and size of operation (inventory of hogs) boundaries, of which there will be five levels of each. Regions will be defined as:

- Central (Iowa, Kansas, Minnesota, Missouri, Nebraska, Oklahoma, and South Dakota),
- East (Illinois, Indiana, Kentucky, Michigan, Ohio, Pennsylvania, Virginia, West Virginia, and Wisconsin),
- Northeast (Maine, Massachusetts, New York, and Vermont),
- South (Alabama, Arkansas, Florida, Georgia, Louisiana, Mississippi, North Carolina, South Carolina, Tennessee, and Texas), and
- West (Arizona, California, Colorado, Idaho, Montana, Oregon, Utah, and Washington).

And size categories will be defined as above.

- **ESTIMATION PROCEDURE;**

Large swine operations (1,000 head or more) component

The sampling design is a multi-phase, multi-stage stratified sampling design with probabilities of selection at Phase I, Stage I equal to 1 and probabilities of selection of sites at Phase I, Stage II being proportional to the number of sites raising hogs in that state under the given operation. Statistical estimation will be undertaken using either SAS survey procedures and/or SUDAAN. Both software packages use a Taylor series expansion to estimate appropriate variances given the design, including the non-response-adjusted sampling weights.

Small swine operations (less than 1,000 head) component

The sampling design is a multi-stage stratified sampling survey design with unequal probabilities of selection across strata during Stage I and probabilities of selection of sites (where applicable) at Stage II being proportional to the number of sites raising hogs in that state under the given operation. Strata are defined by the State and size categories given above. Statistical estimation will be undertaken using either SAS survey procedures and/or SUDAAN. Both software packages use a Taylor series expansion to estimate appropriate variances given the design, including the non-response-adjusted sampling weights.

- **DEGREE OF ACCURACY NEEDED FOR THE PURPOSE DESCRIBED IN THE JUSTIFICATION;**

Large swine operations (1,000 head or more) component

The APHIS goal is to produce descriptive statistics (proportions or means) with a coefficient of variation (CV) of 20 percent or less. If possible given adequate response rates, estimates will be produced by region (Midwest, East, and West) and by size categories (Small, Medium, and Large).

In order to meet the precision criteria within each of the given reporting stratification cells, we require an overall sample size of approximately 616 sites assuming that a simple random sample with a perfect response rate is taken. However, due to practical considerations, we must account for the expected response of 54.0 percent at Phase I, a consent rate for Phase II of 53.6 percent, an expected response of 51.0 at Phase II, and an expected design effect of approximately 2.18 to

obtain estimates with appropriate standard errors. An overall sample size of approximately 4,875 sites is required after adjusting for these factors to produce estimates with CV not exceeding 20 percent within each of the reporting stratification cells. However, due to the fact that there are only approximately 2,700 operations in the population and because we do not want to overburden the population by collecting data that may be correlated, the overall sample size will be restricted to 4,085, of which 2,700 will be operations and the remaining 1,385 will be sites sub-selected from those chosen operations (see Appendix C for further explanation of the sample size estimation and allocation among phases).

Tables C.1 and C.2 in Appendix C show estimates of precision based on the total sample of 4,085 for both phases for breakouts by region and by size, respectively. Response rates from previous studies (see Table D.1 in Appendix D for response rate information from the NAHMS Swine 2012 and NAHMS Antimicrobial Use Swine 2017 Studies) and a design effect of 2.18 (calculated using information from the NAHMS Swine 2012 and 2006 studies) were assumed. All of the estimated CVs for Phase I and most of the CVs for Phase II are expected to be within the desired range with the given allocation.

Reporting strata for Phase II estimates may be adjusted depending on the number of respondents (e.g., similar regions may be merged). In general, if sample sizes are too small or CVs too large for any estimates, those estimates are not published or are reported at a more aggregate level.

Small swine operations (less than 1,000 head) component

NAHMS' goal is to produce descriptive statistics (proportions or means) with a coefficient of variation (CV) of no more than 20 percent. If possible given adequate response rates, estimates will be produced by region (Central, East, Northeast, South, and West) and by size categories (Very Small, Small, Medium, Large, and Very Large). Regions will be defined as given above.

In order to meet the precision criteria within each of the given reporting stratification cells, we require an overall sample size of approximately 1,114 sites assuming that a simple random sample with a perfect completion rate is taken. However, due to practical considerations, we must account for the expected percentage of operations that will be eligible (57.4 percent) and the expected percentage of operations that will complete the questionnaire (71.8 percent), and an expected design effect of approximately 2.18 to obtain estimates with appropriate standard errors. An overall sample size of approximately 5,880 operations is required after adjusting for these factors to produce estimates with CV not exceeding 20 percent within each of the reporting stratification cells.

Table G.1 in Appendix G show estimates of precision by region and nationally. Table G.2 shows the similar estimates by size category. Weighted average completion rates from previous studies (see Table H.1 in Appendix H for eligibility and completion rate information from the NAHMS Swine 2007 and 2012 Studies) and a design effect of 2.18 (calculated using information from the NAHMS Swine 2012 and 2006 studies) were assumed. All of the estimated CVs are expected to be within the desired range with the given allocation.

Reporting strata may be adjusted depending on the number of respondents (e.g., similar regions may be merged). In general, if sample sizes are too small or CVs too large for any estimates,

those estimates are not published or are reported at a more aggregate level.

- **UNUSUAL PROBLEMS REQUIRING SPECIALIZED SAMPLING PROCEDURES, AND**

There are no unusual problems requiring specialized sampling procedures and data collection cycles in either component of this study.

- **ANY USE OF PERIODIC (LESS FREQUENT THAN ANNUAL) DATA COLLECTION CYCLES TO REDUCE BURDEN.**

The data collection described is not planned to be executed on an annual or less than annual frequency basis for either component of this study.

3. **DESCRIBE METHODS TO MAXIMIZE RESPONSE RATES AND TO DEAL WITH ISSUES OF NON-RESPONSE. THE ACCURACY AND RELIABILITY OF INFORMATION COLLECTED MUST BE SHOWN TO BE ADEQUATE FOR INTENDED USES. FOR COLLECTIONS BASED ON SAMPLING, A SPECIAL JUSTIFICATION MUST BE PROVIDED FOR ANY COLLECTION THAT WILL NOT YIELD "RELIABLE" DATA THAT CAN BE GENERALIZED TO THE UNIVERSE STUDIED.**

Large swine operations (1,000 head or more) component

- Questionnaire design and training

The study minimizes collection of data to that which is absolutely necessary to meet the stated objectives. Questionnaires are extensively reviewed by APHIS staff, NASS staff, and experts both in industry and in academia.

For Phase I data collection, APHIS staff will develop training materials for NASS enumerators that explain the purpose of the study and address anticipated difficulties with questions. APHIS representatives will participate in the NASS enumerator trainings. Data collectors and data handlers will have been trained on data and information security guidelines.

For Phase II data collection, study State Coordinators (APHIS staff) will receive specialized training via APHIS study staff and in return train the other APHIS-designated data collectors in their State.

The Study Lead for this NAHMS study has made numerous contacts and has been involved in collaborative efforts to identify the information needs of the industry and the best way to ask for and incentivize that information collection via questionnaire.

APHIS will build on its efforts to collect data using an electronic data capture system, piloted in two recent NAHMS national studies. At least some of the data collection for this study will be performed using our electronic data capture system during in-person interviews during Phase II. APHIS-designated data collectors who collect data in this way will be trained by study staff on collection of data and data security using this system.

- Contacting respondents

NASS will send a pre-survey letter to selected operations to inform them of the study objectives and benefits as well as timelines and what they can expect if they decide to participate.

The study is supported by the National Pork Board, the National Pork Producer's Council, and the American Association of Swine Veterinarians. The support of these industry groups will help in marketing the study to the appropriate audience.

NASS enumerators will call producers up to seven times followed by an on farm visit (barring contact notes from the NASS list frame) before they are listed as a refused or inaccessible operation. NASS enumerators have gone through specific training to help them answer questions of reluctant producers so as to maximize response rates.

The APHIS data collectors will contact farms that have consented to continue in Phase II of the study and set up a convenient time for the producer to complete the questionnaire and conduct biological sampling. Training for the APHIS data collectors will include specific suggestions from the NASS trainers based upon their experience in working with specific producers to maximize response.

Phone scripts will be provided to NASS and APHIS data collectors in order to guide them through study specifics, benefits, and commitments by the producer in order to maximize participant conversion.

- Nonresponse adjustment

Baseline response rates are expected to be approximately 54.0 percent for Phase I of the study and 51.0 percent for Phase II. The assumed response rates are taken from the NAHMS Swine 2012 study for operations with 2,000 or more hogs. The NAHMS Swine 2012 study was more similar to the NAHMS Swine 2020 study in design and material than was Antimicrobial Use (AMU) in Swine 2017. Using the response rates for operations with 2,000 or more hogs is more appropriate than the overall response rates because larger operations tend to respond at different rates than smaller operations (see Appendix D for more information on response rates).

If the respondents differ substantially from the non-respondents, then there is potential for bias. NASS' List Frame data may be available for both respondents and non-respondents to allow for examination of potential differences in type of responding and non-responding operations. If needed, a nonresponse bias analysis will be performed to investigate unexpected response patterns to guide future sampling efforts.

Selection weights will be adjusted for nonresponse. Weights of eligible non-respondents will be transferred to responding operations that are most similar based on available data, including the region and size category post-stratification variables. The nonresponse adjustment will use the method of propensity scores, whereby a logistic regression model is constructed to predict the probability of responding. The inverse of this probability is the nonresponse adjustment.

- Sampling and design strategies

Sampling from the NASS List Frame after the NASS 2017 Census of Agriculture will help to maintain adequate response rates by avoiding operations that are out of scope, out of business, or otherwise do not have animals at the time of contact.

Because the 13 chosen States are included in NASS' Quarterly Hogs and Pigs Report, they will have the greatest likelihood of having been updated in the List Frame during the last quarter compared to other potential States.

APHIS plans to implement a face-to-face enumerated questionnaires. Face-to-face enumeration has historically been shown to produce response rates in NAHMS studies greater than those in studies implementing mail-out or phone-based surveys. In addition, this will allow for enumerators of the Phase II questionnaire to implement biological sampling at no cost to the producer.

The two-stage sampling design (sampling operations, then sampling sites within operation) was implemented, in part, to produce higher response rates among swine sites. Sites raising hogs belonging to a parent operation (contractor) that has already given consent to participating in the study are more likely to respond than sites for which no prior approval has been given by a parent operation (contractor).

Small swine operations (less than 1,000 head) component

- Questionnaire design and training

The study minimizes collection of data to that which is absolutely necessary to meet the stated objectives. Questionnaires are extensively reviewed by NAHMS staff, NASS staff, and experts in industry and in academia.

NAHMS staff will develop training materials for NASS enumerators that explain the purpose of the study and address anticipated difficulties with questions. NAHMS representatives will participate in the NASS enumerator trainings.

The swine specialist for NAHMS has made numerous contacts and has been involved in collaborative efforts to identify the information needs of the industry and the best way to ask for that information via questionnaire.

Data collectors and data handlers will have been trained on data and information security guidelines.

- Contacting respondents

A pre-survey letter will be sent to operations eligible as determined by the prescreening questionnaire to inform them of the study objectives and benefits as well as timelines and what they can expect if they decide to participate.

The study has is supported by the National Pork Board, the National Pork Producer's Council, and the American Association of Swine Veterinarians. The support of these industry groups will help in marketing the study to the appropriate audience

Producers will be called by the NASS enumerator 5-10 times if the 2020 NAHMS Swine Small Enterprise Survey (VS Form 21-207) is not returned to assist in interviewing the operator and ensure all of the operator's questions are answered. NASS enumerators have gone through specific training to help them answer questions of reluctant producers so as to maximize response rates.

- Nonresponse adjustment

Baseline percent of eligible operations is expected to be approximately 57.4 percent and 71.8 percent of the eligible operations are expected to complete the 2020 NAHMS Swine Small Enterprise Survey (VS Form 21-207). The assumed most representative completion rates are

weighted averages of completion rates taken from the previous NAHMS Swine Studies (see Appendix H for more information on response rates).

If the respondents differ substantially from the non-respondents, then there is potential for bias. NASS' List Frame data may be available for both respondents and non-respondents to allow for examination of potential differences in type of responding and non-responding operations. If needed, a nonresponse bias analysis will be performed to investigate unexpected response patterns to guide future sampling efforts.

Selection weights will be adjusted for nonresponse. Weights of eligible non-respondents will be transferred to responding operations that are most similar based on available data, including the region and size category stratification variables. The nonresponse adjustment will use the method of propensity scores, whereby a logistic regression model is constructed to predict the probability of responding. The inverse of this probability is the nonresponse adjustment.

- Sampling and design strategies

Sampling from the NASS List Frame after the NASS 2017 Census of Agriculture will help to maintain adequate response rates by avoiding operations that are out of scope, out of business, or otherwise do not have animals at the time of contact. In addition, a request has been submitted to include operations in this population on the NASS Criteria Survey to aid in keeping contact information up to date and maintain adequate list frame coverage.

Because 16 of the chosen States are included in NASS' Quarterly Hogs and Pigs Report, they will have the greatest likelihood of having been updated in the List Frame during the last quarter compared to other potential States.

The two-stage sampling design (sampling operations, then sampling sites within operation) was implemented, in part, to produce higher response rates among swine sites for the small portion of the target population raising hogs at more than one site. Sites raising hogs belonging to a parent operation (contractor) that has already given consent to participating in the study are more likely to respond than sites for which no prior approval has been given by a parent operation (contractor).

4. DESCRIBE ANY TESTS OF PROCEDURES OR METHODS TO BE UNDERTAKEN. TESTING IS ENCOURAGED AS AN EFFECTIVE MEANS OF REFINING COLLECTIONS OF INFORMATION TO MINIMIZE BURDEN AND IMPROVE UTILITY. TESTS MUST BE APPROVED IF THEY CALL FOR ANSWERS TO IDENTICAL QUESTIONS FROM 10 OR MORE RESPONDENTS. A PROPOSED TEST OR SET OF TESTS MAY BE SUBMITTED FOR APPROVAL SEPARATELY OR IN COMBINATION WITH THE MAIN COLLECTION OF INFORMATION.

The proposed paper and electronic (where applicable) questionnaires will be pretested by APHIS prior to field enumeration, involving 3 operations. It is planned that NASS will perform pretests of the Phase I questionnaires on 6 of operations and expert reviews by 8 reviewers.

Results of these pretests will be utilized to refine the questionnaires in order to reduce respondent burden and improve the usefulness of the information. Many revised questions from previous iterations of the NAHMS Swine study will be used. Final questionnaires will have been

reviewed by a variety of experts, including academic researchers, individual representatives, extension agents, veterinarians, health specialists, epidemiologists, survey methodologists, and survey administrators.

5. PROVIDE THE NAME AND TELEPHONE NUMBER OF INDIVIDUALS CONSULTED ON STATISTICAL ASPECTS OF THE DESIGN AND THE NAME OF THE AGENCY UNIT, CONTRACTOR(S), GRANTEE(S), OR OTHER PERSON(S) WHO WILL ACTUALLY COLLECT AND/OR ANALYZE THE INFORMATION FOR THE AGENCY.

The statistical aspects of the design were coordinated by:

- Mr. Matthew Branan, Mathematical Statistician, USDA, APHIS, VS, NAHMS, Fort Collins, CO (970-494-7349).

For questionnaire design and methodology, NAHMS will coordinate with:

- Kathy Ott, USDA, NASS, Methodology Division Standards and Survey Development Methodology Branch, Washington, DC (202-720-1114).

NASS review of the OMB package submission will be coordinated with:

- David Hancock, USDA, NASS, Methodology Division Standards and Survey Development Methodology Branch, Washington, DC (202-690-2388).

The actual data collection will be conducted by NASS enumerators (Phase I) and APHIS-designated data collectors (Phase II). Contact persons for data collection are:

- Mr. Gerald Tillman, Chief, Survey Administration Branch, USDA, NASS, Washington, D.C. 20250, (202-720-3895).

- Dr. Bruce Wagner, Director, Center for Epidemiology and Animal Health, USDA, APHIS, VS, Fort Collins, CO (970-494-7256).

- Dr. Jack Shere, Deputy Administrator, Veterinary Services, USDA, APHIS, Washington, D.C. (202-799-7146).

Analysis of the data will be accomplished by APHIS veterinarians, epidemiologists, and statisticians under the direction of:

- Dr. Amy Delgado, Director, Monitoring and Modeling, USDA APHIS, VS, CEAH, Fort Collins, CO (970-494-7302).

Appendix A: State selection

State selection was performed based on the percentage of operations and of inventory by State, for geographic representativeness, and expected response burden.

Table A.1. Number and percentage of operations and number and percentage of swine inventory by size class. Note, Percent Inventory doesn't sum to 100.0 percent due to value suppression at the State level.

Size class	Operations	Inventory	Percent Operations	Percent Inventory
1,000 or more hogs	10,340	70,336,346	15.6	97.2
Fewer than 1,000 hogs	56,099	2,044,661	84.4	2.8

* Population-level information is taken from the NASS 2017 Census of Agriculture.

Table A.2. Number and percentage of operations and of swine inventory on operations with 1,000 or more hogs, by State. Note, columns may not sum to totals due to rounding.

State	Operations and sites	Inventory	Percent Operations and sites	Percent Inventory
Iowa	3,648	22,266,119	35.3	31.7
North Carolina	1,062	8,853,993	10.3	12.6
Minnesota	1,548	8,256,924	15.0	11.7
Illinois	772	5,155,790	7.5	7.3
Indiana	655	3,852,482	6.3	5.5
Nebraska	405	3,477,015	3.9	4.9
Missouri	328	3,040,454	3.2	4.3
Ohio	616	2,457,846	6.0	3.5
Oklahoma	61	2,138,667	0.6	3.0
Kansas	107	2,063,303	1.0	2.9
South Dakota	220	1,528,790	2.1	2.2
Michigan	175	1,188,415	1.7	1.7
Pennsylvania	298	1,135,305	2.9	1.6
Total for 13 chosen States	9,895	65,415,103	95.7	93.0
Total for remaining States	445	4,921,243	4.3	7.0
Total for U.S.	10,340	70,336,346	100.0	100.0

* Population-level information is taken from the NASS 2017 Census of Agriculture.

Appendix B: Site sub-selection

Because this is a two-stage sampling design with the first stage involving the selection of operations and the second stage involving the selection of sites within operation, we need to distribute the sample among the operation and site levels. In order to do this systematically, we analyzed previous NAHMS Swine study data to compute the intraclass correlation coefficient (ICC), which we use below to determine the proportion of the sample that should be selected at the operation level versus the site level due to the correlation in animal health and management information among sites within a particular operation.

Since there are only approximately 2,700 operations in the chosen states, we can determine the proportion of the sample that should consist of operations and then compute the total sample size that should be chosen without overburdening the population (since sampling more sites than needed will not yield independent observations). Below is the calculation for the proportion of the overall sample size that should be at the operation level and at the site level, using population counts and the estimated ICC. We have,

$$\hat{\lambda} = \sqrt{\frac{\hat{\omega} * \hat{\rho}}{(1 - \hat{\rho})}} = \sqrt{\frac{0.56 * 0.87}{(1 - 0.87)}} \approx 1.94$$

$$\hat{\psi} = 1 - \frac{1}{1 + \hat{\lambda}} = 1 - \frac{1}{1 + 1.94} \approx 0.66$$

$$\hat{k} \equiv 2,700$$

$$n = \frac{\hat{k}}{\hat{\psi}} = \frac{2,696}{0.66} \approx 4,085$$

$$\hat{m} = n - \hat{k} \approx 4,085 - 2,700 = 1,385$$

where parameters are defined in Table B.1.

Table B.1. Parameter descriptions, symbols, and values in computation of the number of operations and number of sites to sample given population count and variability information.

Parameter	Symbol	Value
Desired sample size	n	4,085
Estimated ratio of operations to sites in the population	$\hat{\omega}$	0.56
Estimated number of operations in the population	K	2,700
Estimated intraclass correlation coefficient ¹	$\hat{\rho}$	0.87
Estimated proportion of overall sample size that should be operations (rather than sites)	$\hat{\lambda}$	1.94
Estimated ratio of operations to sites to select	$\hat{\psi}$	0.66
Estimated number of operations to sample	\hat{k}	2,700
Estimated number of sites to sample	\hat{m}	1,385

¹ Estimated from the NAHMS Swine 2012 and 2006 studies using variables such as inventory, carcass disposal method, management, housing type, waste management, veterinarian use, and disease presence indices.

Therefore, when selecting operations from the NASS List Frame, approximately 2,700 will be chosen, which will leave approximately 1,385 to be sub-selected within chosen operations. This procedure produces a sample that balances undue burden on the public while still attaining precision goals (see Appendix C). When actually selecting from the NASS List Frame, if there are fewer than 2,700 within-scope operations on the frame, then a census will be taken at the operation level, with the remaining sample size being sub-selected within operation.

Using the NAHMS Swine 2012 results, we can compute the expected distribution of operations belonging to site size categories (1 sites, 2 sites, 3-10 sites, etc.). Then, given the number of sites to sample per operation, the total number of sites expected to be sampled in NAHMS Swine 2020 can be computed. The total amount matches the target sample size of 4,085.

Table B.2. Number of sites to select based on the percent of operations in given site category in NAHMS Swine 2012 and the resulting number of sites that will be sampled in NAHMS Swine 2020 under the given sampling scheme.

Site category (number of sites)	Percent of operations from Swine 2012	Number of sites to sample per operation	Total number of sites sampled in category in 2020
1	58.6	1	1,579
2	13.3	2	716
3-10	22.9	2	1,236
11-50	4.3	3	347
50 or greater	1.0	8	208
Total	100.0	-	4,085

Appendix C: Expected estimates of precision

Estimates of percent of operations and percent of animals will be reported at the national level and for region and size categories. Estimates of precision are shown for proportions of 0.5, 0.25, and 0.1. As an example, for the Midwest category and an expected proportion of 0.5, the coefficient of variation (CV) is 5 percent in Phase I. Only three of the estimated CV values for the Region breakouts and two for the Size breakouts exceed 20 percent, all for estimates of 0.1 in Phase II.

Table C.1. Precision of estimates by region and by expected proportion, at 95 percent confidence for Phase I.

Region (States)	Estimated overall sample size	Proportion estimate	Phase I CV estimate (%)	Phase II CV estimate (%)
Midwest (IA, MI, MN, NE, SD)	1,429	0.50	5.1	7.3
		0.25	8.8	12.6
		0.10	15.2	21.8
East (IL, IN, OH, PA)	1,349	0.50	5.0	7.3
		0.25	8.6	12.7
		0.10	14.9	21.9
South (KS, MO, NC, OK)	1,309	0.50	4.9	7.3
		0.25	8.5	12.7
		0.10	14.7	22.0
Total	4,085	0.50	2.9	4.2
		0.25	5.1	7.4
		0.10	8.8	12.7

Table C.2. Precision of estimates by region and by expected proportion, at 95 percent confidence for Phase II.

Size of operation (number of hogs)	Estimated overall sample size	Proportion estimate	Phase I CV estimate (%)	Phase II CV estimate (%)
Small (1,000-1,999)	702	0.50	6.6	9.9
		0.25	11.4	17.2
		0.10	19.7	29.8
Medium (2,000-3,999)	774	0.50	6.3	9.5
		0.25	10.9	16.4
		0.10	18.8	28.4
Large (4,000 or more)	2,609	0.50	3.7	5.4
		0.25	6.4	9.3
		0.10	11.2	16.1
Total	4,085	0.50	2.9	4.2
		0.25	5.1	7.4
		0.10	8.8	12.7

Appendix D: Response rates

Table D.1. Response rates from relevant, previous NAHMS studies.

Survey	Phase I			Phase II		
	# Sample	# Complete	% Complete	# Sample	# Complete	% Complete
Antimicrobial Use (AMU) in Swine 2017 ¹	1,725	388	22.5	388	199	51.3
Swine 2012 (all operations)	5,237	2,119	40.5	944	474	50.2
Swine 2012 (operations with 2,000 or more hogs)	2,810	1,516	54.0	812	414	51.0

¹In-scope operations for this study included operations with 1,000 or more hogs. Phase I of this study is not reflective of the sampling that will take place for NAHMS Swine Larger Producer 2020 because the NASS List Frame had not been updated with NASS 2017 Census of Agriculture data, sampling was performed during a poor time of year, and because of the sensitive nature of the questionnaire and thus there were more out of scope, out of business, inaccessible, or otherwise operations with no animals than we expect for the upcoming study.

Appendix E: Burden estimates

Table E.1. Response burden estimates from the NAHMS Swine 2012 Studies (in minutes).

Survey	Median burden for respondents (min.)	Median burden for nonrespondents (min.)	Median travel time for all travelers combined (min.)
Swine 2012 General Swine Farm Report	66.0	10.0	.
Swine 2012 Initial Visit Questionnaire	60.0	15.0	90.0

Appendix F: State selection

State selection was performed based on the percentage of operations and of inventory by State, for geographic representativeness, and expected response burden.

Table F.1. Number and percentage of operations and number and percentage of swine inventory by size class.

Size class	Operations	Inventory	Percent Operations	Percent Inventory
1,000 or more hogs	10,340	70,336,346	15.6	97.2
Fewer than 1,000 hogs	56,099	2,044,661	84.4	2.8

* Population-level information is taken from the NASS 2017 Census of Agriculture.

Table F.2. Number and percentage of operations and of swine inventory on operations with fewer than 1,000 hogs, by State. Note, columns may not sum to totals due to rounding.

State	Operations and sites	Inventory	Percent Operations and sites	Percent Inventory
IA	2,012	464,421	3.6	22.7
MN	1,677	210,437	3.0	10.3
IN	1,915	151,906	3.4	7.4
MO	2,359	109,679	4.2	5.4
NE	825	107,741	1.5	5.3
PA	2,479	103,996	4.4	5.1
OH	2,868	103,406	5.1	5.1
IL	1,381	102,329	2.5	5.0
TX	5,879	83,068	10.5	4.1
WI	2,127	65,229	3.8	3.2
MI	1,842	48,152	3.3	2.4
NC	1,364	45,466	2.4	2.2
KS	836	37,461	1.5	1.8
SD	351	31,732	0.6	1.6
TN	1,860	29,754	3.3	1.5
OK	2,203	26,885	3.9	1.3
KY	1,766	25,274	3.1	1.2
CA	1,380	23,512	2.5	1.1
AR	981	22,279	1.7	1.1
NY	1,730	21,564	3.1	1.1
VA	1,431	20,205	2.6	1.0
CO	1,206	16,445	2.1	0.8
SC	970	16,413	1.7	0.8
GA	1,075	15,063	1.9	0.7
AL	1,064	13,805	1.9	0.7
FL	1,809	13,167	3.2	0.6
OR	1,177	11,218	2.1	0.5
MS	757	7,706	1.3	0.4
MA	336	5,848	0.6	0.3
MT	416	5,305	0.7	0.3
LA	874	4,888	1.6	0.2
UT	553	4,753	1.0	0.2
VT	353	4,736	0.6	0.2
ID	602	4,659	1.1	0.2
WV	892	4,601	1.6	0.2
WA	1,124	4,369	2.0	0.2
ME	429	3,493	0.8	0.2
AZ	519	1,047	0.9	0.1
Total for 38 chosen States	53,422	1,972,012	95.2	96.4
Total for remaining States	2,677	72,649	4.8	3.6
Total for U.S.	56,099	2,044,661	100.0	100.0

* Population-level information is taken from the NASS 2017 Census of Agriculture.

Appendix G: Expected estimates of precision

Estimates of percent of operations and percent of animals will be reported at the national level and for region and size categories. Estimates of precision are shown for proportions of 0.5, 0.25, and 0.1. As an example, for the Midwest category and an expected proportion of 0.5, the coefficient of variation (CV) is 5 percent in Phase I. None of the estimated CV values exceed 20 percent.

Table G.1. Precision of estimates by region and by expected proportion, at 95 percent confidence.

Region ¹	Sampled ²	Eligible ³	Complete ⁴	Effective sample size ⁵	Proportion	Half-width	C V
Central	1,182	678	487	224	0.5	0.07	7
					0.25	0.06	11
					0.1	0.04	20
East	1,191	684	491	226	0.5	0.07	7
					0.25	0.06	11
					0.1	0.04	20
Northeast	1,143	656	471	216	0.5	0.07	7
					0.25	0.06	11
					0.1	0.04	20
South	1,189	683	490	225	0.5	0.07	7
					0.25	0.06	11
					0.1	0.04	20
West	1,176	675	485	223	0.5	0.07	7
					0.25	0.06	11
					0.1	0.04	20
Total	5,880	3,374	2,423	1,114	0.5	0.03	3
					0.25	0.03	5
					0.1	0.02	9

¹ Regions:

- Central IA, KS, MN, MO, NE, OK, SD
- East IL, IN, KY, MI, OH, PA, VA, WV, WI
- Northeast ME, MA, NY, VT
- South AL, AR, FL, GA, LA, MS, NC, SC, TN, TX
- West AZ, CA, CO, ID, MT, OR, UT, WA

² Approximate sample size estimated by computing the sample size necessary to achieve the precision criteria in each region.

³ Eligible rate assumed to be 0.574 (see Appendix H).

⁴ Complete rate assumed to be 0.718 (see Appendix H).

⁵ Effective sample size taken into account using a design effect of 2.18.

Table G.2. Precision of estimates by size category and by expected proportion, at 95 percent confidence.

Size ¹	Sampled ²	Eligible ³	Complete ⁴	Effective sample size ⁵	Proportion	Half-width	C V
Very small	1,199	688	494	227	0.5	0.07	7
					0.25	0.06	11
					0.1	0.04	20
Small	1,149	660	474	218	0.5	0.07	7
					0.25	0.06	11
					0.1	0.04	20
Medium	1,115	640	459	211	0.5	0.07	7
					0.25	0.06	11
					0.1	0.04	20
Large	1,060	609	437	201	0.5	0.07	7
					0.25	0.06	11
					0.1	0.04	20
Very large	1,003	576	413	190	0.5	0.07	7
					0.25	0.06	11
					0.1	0.04	20
Total	5,523	3,170	2,276	1,047	0.5	0.03	3
					0.25	0.03	5
					0.1	0.02	9

¹ Size categories are:

Very small 1-24 head
 Small 25-49 head
 Medium 50-99 head
 Large 100-499 head
 Very large 500-999 head

² Approximate sample size estimated by computing the sample size necessary to achieve the precision criteria in each size category.

³ Eligible rate assumed to be 0.574 (see Appendix H).

⁴ Complete rate assumed to be 0.718 (see Appendix H).

⁵ Effective sample size taken into account using a design effect of 2.18.

Appendix H: Response rates

Table H.1. Completion rates from previous NAHMS Swine Studies.

Survey	Screening phase		Questionnaire phase	
	% Eligible and Responding	Adjusted % Eligible and Responding ¹	% Complete	Adjusted % Complete ¹
Swine 2012 ² (100-2,000 hogs only)			24.8	53.5
Swine 2012 ³ (<100 hogs only)			26.9	83.2
Swine 2007 ⁴ (<100 hogs only)	31.9	41.2	69.3	78.7

¹ “Adjusted percentage” denotes the percentage of sites that were eligible/complete out of all sites that were sampled, omitting the operations without hogs, out-of-scope operations, out of business operations, and the proportion of inaccessible operations that were expected to be out of scope.

² In-scope operations for this study included operations with 100 or more hogs. The numbers presented are for the operations in the 100-2,000 hog size category for Phase I of the study (since NAHMS Swine 2020 CATI will focus on the operations with 1,000 or fewer hogs).

Questionnaires were administered via CATI for operations with fewer than 1,000 hogs and via face-to-face interview for operations with more than 1,000 hogs.

³ In-scope operations for this study included operations with fewer than 100 hogs. Operators were contacted by phone and questionnaires administered via (CATI) follow-up for non-responding operations. There was only one phase of data collection for this portion of the 2012 study.

⁴ In-scope operations for this study included operations with fewer than 100 hogs. Operators were first mailed a prescreening questionnaire with a CATI follow-up for non-responding operations to determine if they had the requisite number of hogs on the operation within a time window. Questionnaires were administered via mail with a CATI follow-up for non-responding operations.

Because the sampling design assumes sampling from the NASS list frame that will have been updated using the NASS 2017 Census of Agriculture information, we anticipate having fewer inaccessible operations and operations with no hogs in the sample. Also, the adjusted completion rates for the Swine 2012 components can be interpreted as an approximate completion rate to a screened study because it ignores the out of scope/out of business/no hogs/inaccessible operations.

The expected response rate for eligible operations for the Swine 2020 CATI study is taken to be an average of the adjusted complete rates from Table A.1. Let

- R denote the event that an operation responds, and
- E denote the event that an operation is eligible for the study.

Then,

$$P(R|E) = \frac{0.535 + 0.832 + 0.787}{3} = 0.718$$

The expected proportion of eligible and responding operations out of those sampled operations is derived from the percent eligible and responding in the Swine 2007 study. Then, assuming that

$$P(R|E) = 0.718$$

$$P(R \cap E) = 0.412$$

We want to estimate the probability of eligibility, or $P(E)$,

$$P(R|E) = \frac{P(R \cap E)}{P(E)}$$

$$P(E) = \frac{P(R \cap E)}{P(R|E)} = \frac{0.412}{0.718} = 0.574$$

Thus, we assume that 57.4 percent of the sampled operations will be eligible for the Swine 2020 CATI study, and, of those, 71.8 percent are expected to complete the questionnaire.

Appendix I: Burden estimates

Table I.1. Response burden estimates from previous NAHMS Swine Studies (in minutes).

Survey	Estimated burden for respondents (min.)	Estimated burden for nonrespondents (min.)
Swine 2012 General Swine Farm Questionnaire (CATI, <100 hogs only)	30.0	5
Swine 2012 General Swine Farm Questionnaire (CATI, >=100 hogs only)	30.0	5
Swine 2000 General Swine Farm Report – Screening (CATI)	10.0	.