UNITED STATES FOOD AND DRUG ADMINISTRATION

Survey of Drug Product Manufacturing, Processing, and Packing Facilities

OMB Control No. 0910-NEW

Part B. Statistical Methods

1. Respondent Universe and Sampling Method

The universe for the survey is all finished dose drug product manufacturers, processors, and packers of drugs who are subject to 21 CFR Part 210 and 21 CFR Part 211, hereinafter referred to as in-scope facilities. This includes manufacturers, processors, and packers of animal and human over-the-counter and prescription drug products who are registered with FDA and conduct one or more of the following types of in-scope activities (as indicated in the FDA registration data):

- Analysis
- Analytical testing
- Labeling
- Manufacturing
- Packing
- Relabeling
- Repacking
- Sterilizing

For the purposes of the survey, we subdivide the in-scope respondent universe into four groups:

- Group 1: Facilities **in U.S. engaged** in drug product manufacturing (in addition to other possible activities)
- Group 2: Facilities **in U.S. *not* engaged** in drug product manufacturing but engaged in other forms of in-scope activity (e.g., labeling, repacking, etc.).
- Group 3: Facilities **outside U.S. engaged** in drug product manufacturing (in addition to other possible activities)
- Group 4: Facilities **outside U.S. *not* engaged** in drug product manufacturing but engaged in other forms of in-scope activity (e.g., labeling, repacking, etc.).

1.1 Sample Frame

Section 510 of the Federal Food, Drug, and Cosmetic Act (FD&C Act) requires firms that manufacture, prepare, propagate, compound, or process drugs in the U.S. or that are offered for import into the U.S. to register with the FDA. To select our survey sample, we will use registration data submitted to FDA that includes the Dun & Bradstreet DUNS number (a unique nine-digit identifier for businesses), name, phone number, address, and contact name and email

for each FDA-registered facility, as well as the type of activity conducted at the registered facility. FDA will also use DUNS employment data with each registered facility to stratify the sample by size.

The following types of facilities are out-of-scope for the purposes of this survey and will be excluded from our sampling frame:

- Human drug compounding facilities, also referred to as outsourcing facilities
- Medicated animal feed manufacturers
- Animal drug compounding facilities
- Particle size reduction
- Positron emission tomography drug production
- Recovery
- Salvage
- Medical gas manufacturing, including medical gas transfilling
- Third-party logistics provider
- Wholesale drug distributor

The survey will include screener questions related to the type of activity (e.g., manufacturing, sterilization, packing, etc.) conducted at the target respondent's facility to ensure that the facility is in-scope.

Table 1 shows the potential respondent universe by type of drug product (human, animal, or human and animal), size, location, and type of facility (based on activity).

Table 1: Target Population[a]

Facility Employment	Facilities in U.S.			Facilities Outside U.S.			Total Number of In-scope Facilities
	Group 1: Engaged in Drug Product Manu-facturing	Group 2: Not Engaged in Drug Product Manufacturing but Engaged in Other In-Scope Activity	Total	Group 3: Engaged in Drug Product Manu-facturing	Group 4: Not Engaged in Drug Product Manufacturing but Engaged in Other In-Scope Activity	Total	
Human Drugs							
1-19	109	63	172	63	34	97	269
20-99	94	64	158	132	67	199	357
100-499	115	43	158	254	61	315	473
500+	29	10	39	123	22	145	184
Unknown[b]	519	309	828	449	95	544	1372
Animal Drugs							
1-19	14	10	24	15	2	17	41
20-99	23	1	24	19	1	20	44
100-499	5	2	7	26	0	26	33

Facility Employment	Facilities in U.S.			Facilities Outside U.S.			Total Number of In-scope Facilities
	Group 1: Engaged in Drug Product Manu-facturing	Group 2: Not Engaged in Drug Product Manufacturing but Engaged in Other In-Scope Activity	Total	Group 3: Engaged in Drug Product Manu-facturing	Group 4: Not Engaged in Drug Product Manufacturing but Engaged in Other In-Scope Activity	Total	
500+	2	1	3	10	0	10	13
Unknown[b]	61	19	80	61	3	64	144
Both Human and Animal Drugs							
1-19	2	1	3	2	1	3	6
20-99	17	8	25	5	4	9	34
100-499	16	2	18	20	3	23	41
500+	5	1	6	26	0	26	32
Unknown[b]	40	14	54	30	4	34	88
Total Affected	**1,051**	**548**	**1,599**	**1,235**	**297**	**1,532**	**3,131**

Source: FDA, 2020.

[a] The respondent universe figures provided on January 24, 2020 by CDER's Office of Quality Surveillance. This data is updated quarterly. Prior to starting survey field work, FDA will use the most recent version of the data to sample from.

[b] The DUNS data used to provide a count of employment at each facility is incomplete and therefore we have also included a count of facilities for which employment is unknown in Table 1.

1.1.1 Sample Allocation

Table 2 shows the total number of facilities, target completions, expected response rate, and sample size for each survey estimation cell (i.e., Group 1 through Group 4). The derivation of these numbers is discussed in detail below.

Table 2: Survey Universe

Respondent Group	Total Number of In-scope Facilities[a]	Completed Surveys Needed[b]	Expected Response Rate	Sampling Frame Deficiency	Sample Size[c]
Group 1: Facilities **in U.S. engaged** in drug product manufacturing	1,051	213	60%	10%	394
Group 2: Facilities **in U.S. *not* engaged** in drug product manufacturing but engaged in other in-	548	180	60%	10%	333

Respondent Group	Total Number of In-scope Facilities[a]	Completed Surveys Needed[b]	Expected Response Rate	Sampling Frame Deficiency	Sample Size[c]
scope activity (e.g., labeling, repacking)					
Group 3: Facilities **outside U.S. engaged** in drug product manufacturing	1,235	220	60%	10%	407
Group 4: Facilities **outside U.S. *not* engaged** in drug product manufacturing but engaged in other in-scope activity (e.g., labeling, repacking)	297	141	60%	10%	261
Total	**3,131**	**754**	**60%**	**10%**	**1,396**

[a] Based on estimates provided in Table 1 (see last row).
[b] Assumes each estimate is an independent estimation cell with a precision target of 95% confidence and 6% margin of error. Because variance estimates for potential continuous variables are not available, the precision target reflects that for binary variables.
[c] Computed by dividing the number of completes needed by estimation cell by the expected response rate divided by 1 minus the sampling frame deficiency.
Totals may not add up due to rounding.

2. Procedures for the Collection of Information

The sample cells will be sufficiently large to yield statistically valid estimates of beneficiary experience with +/- 6 percent margin of error, *e*, at a 95 percent confidence level (i.e., $\alpha = 10$ percent). The desired sample size for each cell, $n_{Group\ i}$, where i the sample cell, is calculated as (Stat Trek, 2015):

$$n_{Group\ i} = \frac{\left(z^2_{\alpha/2}\right)\left(p_{Group i}\right)\left(1-p_{Group i}\right)+e^2}{e^2+\frac{z^2_{\alpha/2}\left(p_{Group i}\right)\left(1-p_{Group i}\right)}{N_{Group i}}}$$

where $n_{Group\ i}$ is the desired sample size for each group; *z* is the critical value (or z score) associated with the desired confidence level α; *e* is the margin of error; $p_{Group\ i}$ is the response distribution; and $N_{Group\ i}$ is the population size of each group.

Because variance estimates for potential continuous variables are not available, ERG assumed the precision target reflects that for binary variables (i.e., 50 percent or $p_{Group\ 2} = 0.5$),[1] the desired sample size based on the above equation and our statistical precision target is:

[1]The assumption yields the maximum sample size estimate.

$$n_{Group\,1} = \frac{(1.96)^2(0.50)(1-0.50)+(0.06)^2}{(0.06)^2+\frac{(1.96)^2(0.50)(1-0.50)}{1{,}051}} = \frac{0.964}{0.0045138} = 213$$

$$n_{Group\,2} = \frac{(1.96)^2(0.50)(1-0.50)+(0.06)^2}{(0.06)^2+\frac{(1.96)^2(0.50)(1-0.50)}{548}} = \frac{0.964}{0.00535255} = 180$$

$$n_{Group\,3} = \frac{(1.96)^2(0.50)(1-0.50)+(0.06)^2}{(0.06)^2+\frac{(1.96)^2(0.50)(1-0.50)}{1{,}235}} = \frac{0.964}{0.00437765} = 220$$

$$n_{Group\,4} = \frac{(1.96)^2(0.50)(1-0.50)+(0.06)^2}{(0.06)^2+\frac{(1.96)^2(0.50)(1-0.50)}{297}} = \frac{0.964}{0.00683367} = 141$$

2.1 Minimum Sample Size Needed

This section explores the *minimum* sample size necessary to achieve a desired power and effect size for hypothesis testing. For most surveys, 80 percent power and 20 percent effect size are typical assumptions used for these calculations. However, our planned sample sizes will exceed these standards. For example, our sample size of 213 completes for Group 1 (achieved by sampling 1,051 members of establishments) far exceeds the minimum sample size needed to achieve 80 percent power and 20 percent effect size, as discussed in detail below.

For establishments in Group 1, let p_i be the proportion of establishments that utilize a particular manufacturing practice. For the purpose of this discussion, we will refer to this as "manufacturing practice A" (e.g., testing for potential hazards in potable water used in the facility). Further assume that $p_i = p_{Group\,2} = p_{Group\,3} = p_{Group\,4}$ for simplicity. Then the sample size needed to compare a proportion of facilities engaged in drug product manufacturing in the U.S. ($n_{Group\,1}$) to p_i will be given by

$$n_{Group\,1} = \left(\frac{Z_{1-\alpha/2}+Z_{1-\beta}}{ES}\right)^2$$

where

$$ES = \frac{\left|p_{Group\,1}-p_i\right|}{\sqrt{p_i(1-p_i)}}$$

and $p_{Group\ 1.}$ is the proportion of establishments that are engaged in drug product manufacturing in the U.S. that are utilizing practice A. If we assume that the proportion of establishments in Groups 2, 3, and 4 that are utilizing practice A are:

$$p_{Group\,2} = p_{Group\,3} = p_{Group\,4} = 0.50$$

A 20 percent effect size, *ES* = 0.20 (i.e., ability to detect a 20 percent difference in the proportion of establishments in use of practice A between the Group 1 and Group 2 populations, the Group 1 and 3 populations, or between the Group 1 and Group 4 populations) implies:

$$0.20 = \frac{\left|p_{Group1} - p_i\right|}{\sqrt{p_i\left(1-p_i\right)}} = \frac{\left|p_{Group1} - 0.50\right|}{\sqrt{0.50\left(1-0.50\right)}} = \frac{\left|p_{Group\,1} - 0.50\right|}{\sqrt{0.25}}$$

$$p_{Group\,1} = 0.20 \times 0.50 + 0.50$$

$$p_{Group\,1} = 0.60$$

Given

$$\propto = 0.10$$

$$\beta = 0.80$$

The required minimum sample size for the drug product manufacturer group will be[2]

$$n_{Group\,1} = \left(\frac{Z_{1-\propto/2} + Z_{1-\beta}}{ES}\right)^2 = \left(\frac{1.96 + 0.84}{0.20}\right)^2 = 196$$

Note that different assumptions about the proportion of other drug processing facilities will lead to different minimum sample size estimates for the manufacturer group even if the power, significance, and effect size figures are unchanged. Below (Table 3) we present the power afforded by different sample sizes for the manufacturer group, at different effect size, *ES*, levels. Eighty percent power and 20 percent effect size are typical standards used in similar surveys, which imply a minimum sample size of 196 for this survey. However, our proposed sample size of 213 completes for the drug product manufacturers in the U.S. group affords an effect size of 20 percent at 83 percent power, exceeding these typical standards. Conducting a similar exercise for the other groups would also show that the planned number of completes for those groups exceed these standards.

Table 3: Power Associated with Different Size Samples for Drug Product Manufacturers in the U.S. (i.e., Group 1) at Varying Effect Sizes (*ES*)

Sample Size	Power					
[a]	***ES* = 10%**	***ES* = 15%**	***ES* = 20%**	***ES* = 25%**	***ES* = 30%**	***ES* = 35%**
50	11%	18%	28%	41%	55%	68%

[2] Note that the actual calculations are based on unrounded numbers.

Sample Size [a]	Power					
	***ES* = 10%**	***ES* = 15%**	***ES* = 20%**	***ES* = 25%**	***ES* = 30%**	***ES* = 35%**
75	14%	25%	40%	57%	73%	85%
100	17%	32%	51%	70%	84%	93%
125	20%	38%	60%	79%	91%	97%
150	23%	45%	68%	86%	95%	99%
175	26%	51%	75%	91%	98%	100%
200	29%	56%	80%	94%	99%	100%
213[b]	31%	59%	**83%**	95%	99%	100%
225	32%	61%	85%	96%	99%	100%
250	35%	66%	88%	98%	100%	100%
275	38%	70%	91%	99%	100%	100%
300	41%	74%	99%	99%	100%	100%
325	44%	77%	95%	99%	100%	100%
350	46%	80%	96%	100%	100%	100%
375	49%	83%	97%	100%	100%	100%
400	51%	85%	98%	100%	100%	100%
425	54%	87%	99%	100%	100%	100%
450	56%	89%	99%	100%	100%	100%
475	59%	91%	99%	100%	100%	100%
500	61%	92%	99%	100%	100%	100%
525	63%	93%	100%	100%	100%	100%
550	65%	94%	100%	100%	100%	100%
575	67%	95%	100%	100%	100%	100%
600	69%	96%	100%	100%	100%	100%
625	71%	96%	100%	100%	100%	100%
650	72%	97%	100%	100%	100%	100%
675	74%	97%	100%	100%	100%	100%
700	75%	98%	100%	100%	100%	100%

[a] This represents the sample size needed for the manufacturer group.
[b] This is the desired sample that would yield 83 percent power at a 20 percent effect size level, as noted in the earlier discussion.

2.2 Expected Response Rate

To estimate respondents' responsiveness to the survey, we examined the rates of response to other surveys similar in length, mode(s) of administration, and population sampled. For example, the Product Research Quality Institute (PRQI) conducted an online survey of FDA-registered domestic and foreign locations of firms that manufacture biological drug and device products. These firms routinely receive current good manufacturing practice (CGMP) inspections by FDA to obtain industry feedback on inspection and compliance aspects of program operations. The survey was sent to 163 registered manufacturing facilities and 26 percent of facilities responded (Buchholz et al, 2007). A smaller, thirty question web-based survey designed to assess aspects of

dossier development between pharmaceutical companies that was distributed to 26 pharmaceutical companies resulted in a 50 percent response rate (Estrada et al., 2008). A web survey conducted in Sweden of 47 member companies of the Swedish Association of the Pharmaceutical Industry regarding the cost of Good Clinical Practice related activities resulted in a response rate of 62 percent (Funning et al, 2009). A larger email survey, conducted to determine the advantages and disadvantages of outsourcing regulatory affairs tasks in the pharmaceutical industry in the EU, generated a response rate of 48 percent (Gummerus et al., 2016).

It is clear from these surveys that the response rate can vary widely, it is expected that the response rate of the current survey effort will likely be on the higher side, given the targeted contact data available in FDA's data and the methods that will be used to maximize response rates as described in Section 2.1. Therefore, FDA estimates the expected response rate to this survey at 60 percent, which is similar to the Swedish survey regarding Good Clinical Practice related activities.

2.3 Statistical Methodology for Stratification

A statistical method for stratification will be used for each of the four survey cells noted above, i.e., Groups 1 through 4. We will conduct proportional stratified random sampling based on the five employment class size groups within each survey cell. This results in the following strata within in each survey cell (i.e., Group 1 through Group 4):

- Facilities that manufacture, process, and(or) pack human drug products
- Facilities that manufacture, process, and(or) pack animal drug products
- Facilities that manufacture, process, and(or) pack human and animal drug products

- Facilities with 1-19 employees
- Facilities with 20-99 employees
- Facilities with 100-499 employees
- Facilities with more than 500 employees
- Facilities with unknown number employees

The design reflects simple proportionate sampling such that the sample size of each stratum within a survey cell is proportional to the size of the universe for that stratum in that survey cell. In other words, if a given stratum (e.g., U.S. human drug product manufacturing facilities with 20-99 employees) contains 20 percent of all establishments that manufacture drug products in the U.S. in the study universe, the sample size for that stratum will account for 20 percent of the sample size for that survey cell.

Assuming 213, 180, 220, and 141 targeted number completes in Group 1, 2, 3, and 4, respectively (see Table 2), a response rate of 60 percent and a statistical precision target of +/- 6 percent margin of error at 95 percent confidence level, we will use stratified random sampling by employment class size to select 394, 333, 407, and 261 (overall 1,396 facilities) to sample from Group 1, 2, 3, and 4, respectively. The sample allocation is shown in Table 4 below.

Table 4: Sampling allocation

Facility Employment	Facilities in the U.S.			Facilities Outside of the U.S.			Total Number of In-scope Facilities		
	Group 1: Engaged in Drug Manufacturing[a]	Group 2: Not Engaged in Drug Manufacturing but Engaged in Other In-Scope Activity[a]	Total	Group 3: Engaged in Drug Manufacturing[a]	Group 4: Not Engaged in Drug Manufacturing but Engaged in Other In-Scope Activity[a]	Total	Engaged in Drug Manufacturing[a]	Not Engaged in Drug Manufacturing but Engaged in Other In-Scope Activity[a]	Total
Human Drugs									
1-19	41	38	79	21	30	51	62	68	130
20-99	35	39	74	44	59	102	79	98	177
100-499	43	26	69	84	54	137	127	80	207
500+	11	6	17	41	19	60	51	25	77
Unknown	195	188	383	148	84	232	343	271	614
Animal Drugs									
1-19	5	6	11	5	2	7	10	8	18
20-99	9	1	9	6	1	7	15	1	16
100-499	2	1	3	9	0	9	10	1	12
500+	1	1	1	3	0	3	4	1	5
Unknown	23	12	34	20	3	23	43	14	57
Human and Animal Drugs									
1-19	1	1	1	1	1	2	1	1	3
20-99	6	5	11	2	4	5	8	8	16
100-499	6	1	7	7	3	9	13	4	16
500+	2	1	2	9	0	9	10	1	11
Unknown	15	9	24	10	4	13	25	12	37
Total	**394**	**333**	**728**	**407**	**261**	**669**	**802**	**594**	**1,396**

[a] Computed by dividing the total target number of completes needed by estimation cell in Table 2 by the expected response rate divided by 1 minus the sampling frame deficiency to calculate the target number of respondents and then distributing that estimate proportional to Table 1. Please note that totals may not add due to rounding.

2.3 Statistical Methodology for Sample Selection

The statistical method for selecting establishments to sample within each stratum will involve assigning each registered facility a random index number, using a random number generator. The registered facilities in each stratum will then be arranged in ascending order according to their random index number. If S_j is the size of the solicited sample in the j^{th} stratum, then those S_j registered facilities with the smallest index numbers will be selected and included in the sample.

2.4 Estimation Procedure

2.4.1. Analytic Methods

Survey data will be collected and maintained using an online survey system (Qualtrics). Final survey data will be downloaded in comma-delimited format for data cleaning and analysis. We will perform data cleaning and descriptive analysis in SAS v.9, and text analysis (for those questions that require verbatim responses) in MS Excel.[3]

Using the survey algorithms in SAS v.9 (e.g., PROC SURVEYFREQ, PROC SURVEYMEANS, etc.), the data analysis to be conducted will involve:

- A non-response bias analysis using variables such as establishment size, geographic location, type of manufacturing/processing/packing, and product manufactured/processed to assess any non-response bias (i.e., whether and how the non-respondents are different than the respondents).
- For each respondent, computation of:
  - Simple weights which are the inverse of the selection probability multiplied by the probability of response in the absence of non-response bias, or
  - Adjusted weights that account for non-response bias using the variables establishment size, geographic location, type of manufacturing/processing/packing, and product manufactured/processed/packed, if determined to influence response based on the findings of the non-response bias analysis.
- Tabulating weighted proportions and corresponding standard errors for each survey question in the manufacturing and non-manufacturing groups (e.g., weighted proportion of respondents who responded “Yes,” “No,” or “Don’t Know” for a given survey item).
- Testing to see if there are statistically significant differences in responses to each survey item among the manufacturing and non-manufacturing groups.

[3] Text analysis will involve a review and analysis of the verbatim responses to those questions that include an “Other – Please Specify” response category.

2.4.2 Simple Weights

2.4.2.1 Survey

Each respondent to the survey will be assigned a weight based on the inverse of the selection probability of the respondent's corresponding stratum multiplied by the probability of response. Below we discuss the method we will use in computing simple weights for respondents in each survey estimation cell (i.e., Group 1, Group 2, Group 3, and Group 4) that account for probability of selection and response, but do not incorporate the possibility of non-response bias. Thus, the derivation of these simple weights assumes that there are no significant differences with respect to such factors as establishment size, geographic location, and type of registered activity, between respondents and non-respondents to the survey in any of the survey estimation cells. Weights that deal with the possibility of non-response bias are discussed in Section Error: Reference source not found below.

For survey estimation cell, *i* (where *i* = Group 1, Group 2, Group 3, and Group 4), the probability of selection, $P^{i}_{S,j,k}$ for the j^{th} type of product manufactured/processed (human drug, animal drug, or both human and animal drugs) and k^{th} employment class size is given by:

$$P^{i}_{S,j,k} = \frac{S^{i}_{j,k}}{U^{i}_{j,k}}$$

where $U^{i}_{j,k}$ is the number of establishments in Group *i*, employment class size *j*, and establishment location *k*; $S^{i}_{j,k}$ is the size of the solicited sample in Group *i* employment class size *j*, and establishment location *k*.

Additionally, for survey estimation cell, *i*, the probability of response, $P^{i}_{R,j,k}$, for the j^{th} type of product manufactured/processed and k^{th} employment class size, can be calculated by dividing the solicited sample size in each stratum by the actual number of responses from the corresponding stratum, i.e.:

$$P^{i}_{R,j,k} = \frac{R^{i}_{j,k}}{S^{i}_{j,k}}$$

where $S^{i}_{j,k}$ is the size of the solicited sample in Group *i* and type of product manufactured/processed *j* and employment class size *k*; $R^{i}_{j,k}$ is the actual (responded) sample in Group *i* and type of product manufactured/processed *j*, and employment class size *k*. Then the simple sample weights, $W^{i}_{j,k}$ for Group *i*, type of product manufactured/processed *j* and employment class size *k* are computed as:

$$W^{i}_{j,k} = \frac{1}{P^{i}_{S,j,k} \times P^{i}_{R,j,k}}$$

where the terms are as defined above.

2.4.3 Degree of Accuracy Needed for the Purpose Described in the Justification

The accuracy required of the respondents poses no special demands on them. All data being requested can be readily supplied by respondents. The sample size was calculated to enable us to generate weighted sample estimates of proportions of interest in each group in the +/- 6 percent range of the true proportion with 95 percent confidence (i.e., α = 5 percent).

2.5 Unusual Problems Requiring Specialized Sampling Procedures

There are no unusual problems anticipated.

2.6 Use of Periodic (Less Frequent than Annual) Data Collection Cycles to Reduce Burden

This is a one-time data collection, which will minimize the burden on survey respondents.

3**.** Methods to Maximize Response Rates and Deal with Issues of Non-Response

The survey will be implemented both by mail and online. The process is summarized in Table 5. Survey respondents will receive an email containing the invitation describing the survey and providing each respondent with the URL to the survey, as well as their unique username and password. To ensure that we target the right respondents, the first few questions of the survey will ask about the activities conducted at the facilities. This will screen out any respondents that do not engage in drug product manufacturing or other forms of in-scope activities.

For the full survey, we will begin with a pilot of approximately 50 target respondents. Conducting a pilot is good practice as it helps identify and rectify unanticipated problems that might arise (e.g., inability to access the online survey using a particular browser). The survey pilot will take place over a two-week period (any lagging surveys will be handled in the same way as for the main group of respondents, discussed below).

Once most pilot surveys have been completed and all changes are made to the survey based on the pilot (assumed to take 1 week), we will send out the pre-notification by email (and mail if no email address has been identified) reminding respondents about the focus and extent of the survey, and providing each respondent with the URL of the online survey and their unique password. After two weeks, non-responders will receive a reminder email with their unique password and the survey URL or a reminder via USPS that includes a cover letter, a hard-copy of the survey, and a return envelope, if no email address is available. The second reminder will be sent two weeks after the first reminder and will include an email reminder and a postcard. The third reminder, sent one week after the second reminder, will be similar to the first reminder. The final reminder will be a telephone call, at which time the respondent will be offered the opportunity to complete the survey over the phone.

Table 5: Overview of Data Collection Steps to Maximize Response Rates

Data Collection Stage	Contact	Contact Type	Content
Full-scale Survey	Initial Contact	Email	Survey Pre-notification/Survey Link
	First Reminder	Email and mail	Survey Reminder with Survey Link/Survey Reminder Cover Letter with Hardcopy Survey (2 weeks after initial contact)
	Second Reminder	Email and mail	Survey Reminder with Survey Link and Reminder Postcard (2 weeks after first reminder)
	Third Reminder	Email and mail	Survey Reminder with Survey Link/Survey Reminder Cover Letter with Hardcopy Survey (1 week after second reminder)
	Fourth Reminder	Telephone	Survey Reminder/Caller Offers to Issue Survey by Phone (1 week after second reminder)

Multiple strategies will be employed to maximize response rates, including multiple contacts (i.e., an initial contact and several reminders), pre-notification, multiple modes of administration, and a survey help line. Text of the notifications and reminders are provided in the Appendix.

Multiple contacts. In this data collection, we plan to follow the Dillman Total Design survey method (Dillman, et al., 2014), which emphasizes multiple contacts with members of the sample as being one of the most successful techniques to increase response rates. This technique is now considered standard methodology for any survey. In this survey, we will use a survey invitation message with a link to the survey that includes questions in the beginning to eliminate out of scope respondents. This is followed by one or more contacts with non-respondents using a combination of email and mailed hardcopies of the survey (first and third reminder) or email and a reminder postcard (second reminder). Phone calls will only be made as a fourth reminder.

Pre-notification letters/emails that provide more information on the study increase respondent confidence in the validity and the importance of the study resulting in higher response rates. As such, we will send out pre-notification letters as part of this data collection effort.

Multiple mode administration (phone and mail, mail and Web, etc.) of a survey has been shown to increase response rates (Dillman, et al., 2014). Additionally, the use of multiple modes can also reduce non-response error and data collection costs. In this survey, respondents will be offered the option of completing the survey on-line and by mail. Respondents will also be offered the option of completing the survey by phone if a phone contact is made according to the reminder schedule.

Survey helpline. One tool we believe will be essential for a smooth survey administration is a survey helpline. Although a full vetting of the survey through expert review, QA/QC, pre-testing, and a pilot will be done, some questions will always arise. We will provide a contact

email/phone number for questions and assign a staff member to answer phones, respond to simple FAQ questions, and/or take messages for questions that require senior staff or FDA input. All calls and their content will be logged. A helpline email box will also be set up and staff will review the inbox daily.

Since widely accepted data collection techniques are being used and substantial resources are being devoted to minimizing non-response, we expect the response rate to this survey to be comparable or better than that achieved for surveys of similar size and scope.

Potential reasons for non-response include refusal, language barrier, and other circumstances, as well as the inability to contact the respondent. After the survey has been conducted, we will perform an analysis of non-response bias in the survey estimates. If non-response bias is detected, we will create adjusted weights based on establishment size, geographic location, and type of activity.

Using standard procedures, we will first construct a logistic model of the propensity for survey completion based on the following exogenous variables available for each target respondent (Lohr, 1999; Abraham, et al., 2006):

- Establishment size,
- Geographic location (e.g., EU, India, China, etc.),
- Type of activity conducted at the facility,

The general form of the logistic function (omitting the group superscripts for simplicity) is expressed as,

$$P_R = \frac{e^{x\beta}}{\left(1+e^{x\beta}\right)} = \frac{1}{\left(1+e^{-x\beta}\right)}$$

where P_R in this context is the probability of a respondent completing the survey and x and β are the vectors of explanatory variables (e.g., establishment size, geographic location, type of activity, etc.) and their respective coefficients. Given the above equation, the probability of survey nonresponse can be written as,

$$1-P_R = 1-\frac{1}{\left(1+e^{-x\beta}\right)} = \frac{1}{\left(1+e^{x\beta}\right)}$$

The odds of a positive survey response are, therefore,

$$\left(\frac{P_R}{1-P_R}\right) = e^{x\beta}$$

Taking the natural log of both sides, the above equation becomes,

$$\ln\left(\frac{P_R}{1-P_R}\right) = x\beta$$

For the purposes of nonresponse analysis for this survey, the logit model to be estimated can be specified as,

$$\ln\left(\frac{P_{R,l}}{1-P_{R,l}}\right)=a+b_1 x_{1l}+b_2 x_{2l}+\ldots+b_n x_{nl}+\epsilon_l$$

where $P_{R,l}$ is probability of response for target respondent *l*; *a* is the intercept term; b_i are the associated coefficient vectors for explanatory variables (e.g., establishment size, geographic location, type of activity, etc.); ϵ_k is the error term; and x_i are vectors of dichotomous dummy variables from the sampling frame corresponding to target respondent *l*.

Using maximum likelihood methods, we will estimate the above logistic relationship for target respondents and determine which, if any, estimated coefficients, are statistically significant. If none of the coefficient estimates are statistically significant, no adjustments to weights would be necessary as this would indicate lack of non-response bias. On the other hand, if some or all coefficient estimates are found to be significantly related to the probability of responding to the survey, then it will be necessary to adjust the weights for non-response.

Given the above regression model, the predicted probability of a positive survey response for a given potential respondent *l* will be calculated as:

$$\widehat{P_{R,l}}=\frac{1}{\left(1+e^{-\left(\hat{a}+\hat{b}_1 x_{1l}+\hat{b}_2 x_{2l}+\ldots+\hat{b}_n x_{nl}\right)}\right)}$$

We will then use these predicted probability estimates to recalculate the nonresponse bias adjusted weight to be applied to each respondent's responses. The nonresponse adjusted weights can be expressed as follows:

$$\widehat{\widehat{W}_{jkl}}=\alpha_{j,k}\left(\frac{1}{P_{S,j,k}\times\widehat{P_{R,j,k}}}\right)$$

where $\widehat{\widehat{W}_{jkl}}$ is the adjusted weight for respondent *l* that manufactures type of product *j* (human drug, animal drug, or both human and animal drugs), and employment class size *k*; $\widehat{P_{R,j,k}}$ is the estimated response probability for a respondent derived from the logistic regression, and $\alpha_{j,k}$ is the normalization factor. These factors are calculated to normalize the estimated response probabilities so that the set of nonresponse adjusted weights have the following property for each stratum within a given estimation group:

$$U_{j,k}=\sum_l \widehat{\widehat{W}_{jkl}} R_{jkl}=\alpha_{j,k}\sum_l\left(\frac{1}{P_{S,j,k}\times\widehat{P_{R,j,k}}}\right)R_{jkl}$$

where *l* is summed over all respondents in stratum j, k within a given survey estimation cell (i.e., Group 1, 2, 3 or 4).

Depending on the results of the above analysis, we will also consider using the multivariate regression-based imputation approach, to impute estimated values for non-respondents to address nonresponse bias.

3.1 Generalizing to the Universe Studied

Because we will obtain a stratified random sample of the population, we expect that the information collected will yield reliable data that can be generalized to the universe studied.

4**.** Test of Procedures or Methods to be Undertaken

As part of developing the mail and online survey instruments, the project team has conducted cognitive testing to get initial feedback on respondents' understanding of questions, consistency in interpreting questions and response options, ability to recall necessary information, how well the items reflect the measurement domains, and the flow of the survey tools and interviews.

We first beta-tested the survey instruments with an ERG employee. The ERG completed the survey as if they were a drug product manufacturer and it took slightly over one hour to complete the survey. For burden estimates, we assume that the survey will require one hour and six minutes to complete, whether in paper or online form. This is expected to be an overestimate, given that many respondents will skip over a significant number of questions that are not applicable to their operations.

We additionally conducted cognitive testing of the survey with eight members of the universe studied. In these interviews, respondents provided valuable feedback on how to improve question wording, simplify skip patterns, and otherwise make the elements of the survey package more interpretable. Based on respondent feedback during cognitive testing, we revised the survey to improve the questions – make them easier to comprehend and reduce the complexity of skip patterns.

5**.** Individuals Consulted on Statistical Aspects of Design and Individuals Collecting and/or Analyzing Data

Table 6 below provides the names, affiliation, and contact information for those consulted on the statistical aspects of the design and who will collect or analyze the information.

Table 6: Individuals Consulted on Statistical Aspects and Performing Data Collection & Analysis

Name	Affiliation	Contact Information
Aylin Sertkaya, Ph.D.	Eastern Research Group, Inc.	781-674-7227
Ayesha Berlind	Eastern Research Group, Inc.	781-674-7228
Andreas Lord	Eastern Research Group, Inc.	781-674-7381

Table 7 shows the name of FDA staff who advised on design.

Table 7: FDA Staff who advised on Design

Name	Affiliation	Contact Information
Andrew Estrin, Ph.D.	DHHS/FDA/OC/OPLIA/OEA/ECS	240-402-1829
Jonathan Bray	DHHS/FDA/CVM/OSC/DC	240-402-5623

References

Abraham, K. G., Maitland, A. & Bianchi, S. M., 2006. Nonresponse in the American Time Use Survey: Who Is Missing from the Data and How Much Does It Matter?. *Public Opinion Quarterly*, 70(5), pp. 676-703.

Buchholz, S., V. Gangi, A. Johnson, J. Little, S. Mendivil, C. Trott, K. Webber, and M. Weinstein. 2007. Results of a Survey of Biological Drug and Device Industries Inspected by FDA under the Team Biologics Program. PQRI Report. Volume 61. Number 3. Available at http://pqri.org/wp-content/uploads/2015/08/pdf/PDA_article_PQRI_Biologics_Survey.pdf

Dillman, D., Smith, J. & and Christian, L., 2014. *Internet, phone, mail, and mixed-mode surveys: The tailored design method.* New York: Wiley.

Estrada, P., J. Pocoski, J. Gricar, and M. Roychowdhury. 2008. Evaluation of Pharmaceutical Companies' Current Practices Surrounding Dossier Development. Presentation at the Drug Information Association 19th Annual Workshop on Medical Communications. Available at http://pharmafellows.rutgers.edu/wp-content/uploads/2018/08/2008-evaluation-of-pharmaceutical-companies-current-practices-surrounding-dossier-development.pdf

Funning, Sandra, A. Grahnen, K. Eriksson, A. Kettis-Linbad. 2009. Quality Assurance Within the Scope of Good Clinical Practice (GCP) - What is the Cost of CGP-related Activities? A Survey Within the Swedish Association of the Pharmaceutical Industry (LIF)'s Members. The Quality Assurance Journal. Volume 12. Pages 3-7. Available at https://onlinelibrary.wiley.com/doi/epdf/10.1002/qaj.433.

Gummerus, A., M. Airaksinen, M. Bengstrom, and A. Juppo. 2016. Values and Disadvantages of Outsourcing the Regulatory Affairs Tasks in the Pharmaceutical Industry in EU Countries. Pharmaceutical Regulatory Affairs. Volume 5. Issue 1. Pages 1-5. Available at https://www.omicsonline.org/open-access/values-and-disadvantages-of-outsourcing-the-regulatory-affairs-tasks-inthe-pharmaceutical-industry-in-eu-countries-2167-7689-1000161.pdf.

FDA. 2020. CGMP Survey Sites for ERG. Excel spreadsheet provided by Lakshmi Cherukuri, (CDER/OQS/DQDS). January 24.

Lohr, S., 1999. *Sampling design and analysis.* Pacific Grove: Duxburgy Press.

Stat Trek, 2015. *Sample Size: Simple Random Samples.* [Online]
Available at: http://stattrek.com/sample-size/simple-random-sample.aspx