Age-specific reporting bias

Because infection mortality is high among piglets only [1], we might expect that operations without piglets are less likely to perform diagnostic testing. We can gain some insight into such potential reporting biases from the data about the age classes of diagnostic samples. These age classes are suckling (less than 1 month old or still on sow), nursery (1–3 months of age), grower/finisher (3–8 months), and sow/boar (more than 8 months old). Although the report providing the data uses the term age class and gives those particular age ranges for each class, these terms are really names for production stages in the swine industry for which there may be some variation outside of those ranges, in particular for the time at which pigs are weaned and sent to a nursery.

The age-class distribution of cases varies among states (χ^2=210, simulation $P < 0.001$). But that variation is limited, as indicated by the small value of 0.05 for the uncertainty coefficient [2] of age class, given the state.

To determine whether the proportion of cases in the suckling age class may be explained by the distribution of farm types within a state, we compute expected proportions under a model of two-step random sampling as follows. In the first step, we sample a certain type of farming operation from a distribution of operation types. Table S2.1 gives the names of the available types. We obtain the distribution of these types for each state from census data [3]. In the second step we draw an age class from the age class distribution of the sampled farm type.

We derive an age-class distribution by first assuming that sows on average produced 2.31 litters of weaning size 10.3 every year and that pigs spent 21.5 days as suckling pigs, 46.0 days in the nursery stage, and 121.5 days in the grower/finisher stage. Those parameters are taken from 2012 averages from sow farms, nurseries, and conventional finishing farms participating in a U.S. benchmarking system [4, Tables 2, 4, and 5]. Larger farms tend to use artificial insemination [5, Table 3] and thus we assume that boars make up a negligible part of the total population on sow farms.

Given these parameters, we calculate an age-class distribution for the entire population by first calculating the rate of weaned pig production from the number of sows. The number of animals in all of the other age classes follow as the product of that rate and the average time spent in each class. Age classes on a farms with some subset of age classes follow as a subset of the age-class distribution for the entire population to those classes present on the farm. Table S2.1 shows which age classes are typically present on each type of farm. We normalize these age-class distributions to obtain sampling probabilities conditional on a farm type.

Table S2.1. Farm types and the age classes of swine typically present on them. Ones (zeros) indicate the presence (absence) of an age class on a particular farm type.

Farms type	Suckling	Nursery	Grower/Feeder	Sow/Boar
Farrow to wean	1	0	0	1
Farrow to finish	1	1	1	1
Finish only	0	0	1	0
Farrow to feeder	1	1	0	1
Nursery	0	1	0	0

Using logistic regression, we did not find that observed log odds of suckling cases

increased with the probabilities from our sampling model ($P = 0.64$).

To see if the expected and observed probabilities were different on average, we fitted an intercept-only logistic model with logits of expected probabilities as offset terms. The observed log odds of suckling cases was on average 3.54 natural logarithmic units above those predicted by random sampling (95% profile confidence interval = [3.15,3.97]). Removing highly influential observations (i.e., IA, NC, OK, KS, IL, and MN) resulted in somewhat lower interval estimate bounds of [1.86, 3.51]. These results indicate that farms with unweaned pigs are either more likely to choose unweaned pigs to be diagnostic samples than other pigs, more likely to seek laboratory confirmation of PEDV, or more likely to experience an outbreak than other farms.

References

1. USDA (2014). Summary of PEDV Actions. Available: `http://www.usda.gov/documents/pedv-summary-actions.pdf`. Accessed 10 May 2014.

2. Tang W, He H, Tu XM (2012) Applied Categorical and Count Data Analysis. CRC Press.

3. USDA NASS (2009). 2007 Desktop Data Query Tool 1.02. Available: `http://www.agcensus.usda.gov/Publications/2007/Online_Highlights/Desktop_Application/dataquery.zip`. Accessed 14 January 2013.

4. Kenneth J Stalder (2013). Pork industry productivity analysis. Available: `http://old.pork.org/filelibrary/research/ipafull.pdf`. Accessed 2 December 2014.

5. USDA ERS (2007). The changing economics of U.S. hog production. By N. Key and W. McBride. ERR-52. Available: `http://www.ers.usda.gov/media/244843/err52.pdf`. Accessed 2 December 2014.