Sponsors

University of Minnesota
College of Veterinary Medicine

College of Agricultural, Food and Environmental Sciences

Extension Service

Swine Center

Editors
W. Christopher Scruton
Stephen Claas

Layout
David Brown

Logo Design
Ruth Cronje, and Jan Swanson;
based on the original design by Dr. Robert Dunlop

Cover Design
Shawn Welch

Evaluating growing pig records in a system with high expectations

James F. Lowe, DVM

Carthage Veterinary Service, Ltd.; http://www.hogvet.com; lowe@hogvet.com

Introduction

Data evaluation in systems with high expectations can be challenging, rewarding, and fraught with unforeseen hurdles. As advisors in these systems, we are called upon to help the system succeed in an ever more competitive market by rendering opinions based on data analysis. If our analytical techniques and data are not robust, we risk making bad decisions. This paper will outline components of a good record system, philosophical approaches to data analysis, and a review of some basic data analysis techniques that have been employed in many large systems.

Record systems: The good, the bad, and the ugly

There are a variety of commercial recordkeeping systems for nursery/grow/finish records available in the marketplace. In addition, many production systems have their own internal databases that are more than adequate. Often, the challenge with records is not the computerized system that is used to capture the data; it is the raw data that can lead to significant challenges in the data analysis process. The old saying "The only thing worse than no data is bad data" could not be truer. The quality of records in many systems I have observed is excellent, while others are fraught with basic errors in data recording and systematic errors in data capture. The most common errors I see in records today are inaccurate inventories in both livestock and inputs. These inventory inaccuracies can lead to faulty interpretations because skewed data makes the divisor too large or small in many of the formulas critical to our interpretation of that data.

The best record systems are based on the financial accounting packages the system is utilizing. This eliminates double data entry and also, in general, increases the accuracy because of the tight inventory control required in most systems' financial departments. Deciding when and where animals will be counted in the system is extremely important as the record system is established. Wean-to-market systems have an inherent advantage in that only the placement count at the time of entry and the number of pigs at slaughter must be counted to gain an accurate total of what occurred during the grow-out period. Nursery-to-finish systems require an additional count of animals exiting the nursery and entering grow/finish, but, with sufficient dedication, will yield results as accurate as wean-to-market barns. The challenge in working with these systems is to understand and refine data capture so the livestock inventory process is easy and transparent. No system is fail-proof, but those systems that rely on "double-entry" (i.e., counts out of farrowing and into nursery being required to match) consistently produce more accurate results than those systems that rely on only one set of numbers. In addition, those systems that utilize the count at the slaughter plant as opposed to the count out of the barn should result in a higher degree of confidence in the data generated.

Feed inventory or delivery can also be a challenge. This is particularly true in large systems where multiple trucks are involved and inaccuracy in feed deliveries can occur. There is not an easy solution to this problem, but understanding the processes in the system will lead to a much better understanding of how much confidence that you can place in the data. Those systems that transfer feed manufacturing data directly to the financial record system tend to have fewer data entry errors. Hence, their growing pig systems tend to have fewer errors in their performance records because of the single entry at the time of feed purchase.

In summary, the most important thing that can be done in records analysis is to ensure the validity of the data prior to beginning any analytical process. This validity check often involves review of individual closeouts, histogram summaries for key production variables (average daily gain, feed conversion ratio, and average daily feed intake), a thorough understanding of the recordkeeping system in place and all the potential error points in the system, and discussions with data entry staff about the inaccuracies inherent in the system, as they are often your best source of information. As advisors, our primary—and most beneficial—role may be helping some systems increase the accuracy of their records before even attempting any in-depth analytical work for them.

Philosophical approaches to data analysis

I have been asked the question many times, "Could you please analyze these records for us?"—which leads to the follow-up question, "What is your goal?" The key thing in analyzing any set of data is understanding what the goal of the system is. I categorize systems with high expectations into two categories: low cost and optimum profit. These two categories may seem to be redundant, but there are distinct differences between them.

Low cost system

In low cost systems, upper management focuses strictly on the absolute cost of producing a pound of pork or the absolute cost of producing a pig in dollars per head. Common questions asked in these systems are:

- What was my total feed cost?
- What was my average cost per pound of feed?
- How many dollars per head did we spend on medication?

Many systems function in this mode. When the management team understands how cost impacts profit, it can be a highly successful approach to management.

Optimum profit system

The second group of high expectation farms fits in the optimum profit category. These farms are focused on generating value, not just lowering costs. Common questions in these systems are:

- What was my feed cost per pound of gain?
- What was my medication cost per pound sold?

These systems can also be highly successful, but require a different style of management to maintain a high level of profitability compared to those systems that are focused on absolute cost. The basic difference in these two systems is the cost system asks "What were my total expenditures per pig or per pound of feed?" and the optimum profit farms ask "What was my cost per pound of gain?" Understanding this difference when attempting to do records analysis is critical in providing the information with which the management team can make decisions in a style they find most effective.

Once you have determined what kind of system you are working in, it is important for you to understand what drives the system with respect to profit. Is it a southeastern system in which relatively low building costs but high feed costs mean feed efficiency and feed cost per pound of gain are the most important in determining profitability? Or is it a Midwestern system with relatively low grain costs, but relatively high facility costs where throughput through the buildings is the most important driver and feed efficiency becomes secondary? One of the most common mistakes I see in records analysis is the creation of many indices of production (e.g., average daily gain, average daily feed intake, feed conversion, etc.), but little focus on the key drivers of profitability; as a result, the system fails to make improvements in a rapid manner.

The second question I ask before I start an analysis projects is "Can we detect difference with the system we have in place?" Management teams often ask their record system to determine if the addition of pharmaceutical product or change in feed budget strategy resulted in lower cost or faster gains. It seems intuitive that you should be able to detect this with the record system. However, the analysis of a low number of closeouts may not provide enough statistical power to be a meaningful method of determining the value of changes in the system. As advisors, it is important that we understand the power any analysis has and be able to explain it clearly before we begin any analysis. Many successful management strategies have been aborted because a record system, even with good, accurate data, was not sensitive enough to detect the change induced as a result of those strategies. A simple way to work through this problem is to take the number of closeouts you believe management is going to look at prior to making a decision. In a weekly system, this may be as few as 10. In systems with multiple barns per week this could be 30–40, but they are willing to evaluate with the management change prior to either accepting or rejecting management interventions. Take the key drivers (average daily gain, feed efficiency, mortality) for that number of closeouts and calculate a mean and standard deviation for that group. The next step is to estimate the change of the mean you expect to see with the new intervention and determine the probability you will detect this level of change with the number of closeouts you are going to measure. Take this data and use a power analysis table found in any common statistics textbook to calculate the number of closeouts that need to be analyzed to determine whether a change has taken place.

This simple power analysis can give you a good starting point for discussion with management about how they are going to measure a known intervention. As advisors, it is often our job to understand when things need to be measured in a controlled research setting and when on-farm industrial research is appropriate to determine if we should or should not be intervening in a given matter.

Understanding the measures you use

All average daily gains are not created equal. Understanding how the outcome measures that are used in a given system are calculated can lead to dramatically different approaches to how the data is analyzed. Understanding if dead weight is included in weight gain is a simple example of differences that occur in systems and how they calculate gain. "Does average slaughter weight include

culls or not include culls?" is another key question that could be asked when starting any records analysis process. In addition, questions such as "Are cull pigs left in the barn or removed at placement?," "How many pigs are destroyed between nursery and finishing and have disappeared out of the system?," and "How is mortality calculated?" are important questions to ask as you begin doing data analysis. Unfortunately, there is not a correct answer for how to calculate any of these things in a given system. There are good arguments for and against almost any method of calculation used in systems today. We, as individuals, have our preferences and believe that the way things are done in our own system is correct. However, it is more important to understand what is being calculated than necessarily to change the method of calculation in the system. As you start your analysis, understanding the method of calculation will help you draw accurate conclusions as the analysis is completed and will enable you to make recommendations that have increased validity over time.

Charting the process

Process improvement and management philosophies have led to the development of a whole range of methods to chart and describe the production process. Although the purpose of this paper is not to describe those in detail, it is worth mentioning that being able to record what is normal in a system and then to document all the variations from normal over time is extremely valuable. This enables you to understand what is changing in the system and why variation has occurred. Some systems' recording methods can be quite elaborate, while others are rudimentary; all serve the same purpose of accurately documenting when changes occur. A system with which I am familiar uses a series of informal logbooks along with the pig flow recording system to chart when we make changes at given sites. These log books are kept by the feed mill and the production staff as a simple way to document what happened and what changes occurred. Although they are not easy to search, all of the data is there and, in my experience, the data capture is better than some systems with an elaborate electronic system that has excellent search capability embedded in it.

Another key charting process is to understand basic differences between sites. This may be a building design (total slat versus partial slat), intentional stocking differences (single- versus double-stocked wean-to-finish), or dedicated flow difference. Coding these classifications in your record system may be highly beneficial later if the goal is to make decisions about these types of categories and the value that they bring to the production process. The challenge with categorizing sites or batches is that many times systems get so carried away with categorization that they fail to discover basic differences because their categories are so fine that meaningful analysis cannot be completed. As with accounting, the world is divided into lumpers and splitters. I think in this case, lumping into larger groups offers greater value because the increased number in each group allows for more powerful comparisons to be completed at a later date.

Math: It's just the math

It is important to remember that, as we perform data analysis, we ought not to attempt to make the calculations too complicated or to attempt to over-analyze what is not there. While there are myriad fancy statistical techniques to tease-out differences in systems over time, these analyses are often not as beneficial as one would think to senior management teams' decision-making processes. The common questions I am asked are:

- Where are we at relative to the industry?
- Where are we at relative to where we could be?
- Are we getting better?

These three questions fit into two broad categories: internal and external benchmarking. Many systems tend to focus a great deal on external benchmarking. There is sometimes a lack of benchmarking against the biological potential, but that also can serve as a very useful form of data analysis. Internal benchmarking revolves around either SPC charting or means comparison over time to understand whether the system is making consistent progress.

External benchmarking

External benchmarking is a process that, while appearing to be simple, is fraught with many challenges. The primary challenges are inconsistency of measurements between systems, inaccurate reporting of data, and bias toward only reporting the best data available. External benchmarking is most commonly done by calculating a mean for each of the key production indices over a set period and comparing those means to other means from other systems. The Agrometrics(r) database is a prime example of external benchmarking and has proved very useful for many producers. I personally prefer to spend very little time on this method of benchmarking because I am never confident we are comparing things on an equal basis.

The form of external benchmarking I prefer is against biological potential. This can be extremely difficult because of the limited amount of biological potential data for a given genotype that has been reported in literature. However, many genetic suppliers have constructed quite accurate genetic models to predict the genetic potential of given offspring from a known cross at the commercial level. The question that then arises is, "What is the 'Ge-

netics by Environment' (G X E) interaction that occurs in a given production system?" This form of analysis leads to an understanding of what the environmental drag is on the known genetic cross and what the potential cost is of reducing that environmental drag, and a determination of whether the cost outweighs the benefit or the benefit outweighs the cost. This form of analysis is also done by comparing means. However, one must be very careful to inspect one's data to make sure the distributions are normal. Any skewing of the distribution may mean that comparison of the median to biological production may be more valid representation of where you are today. To make this analysis meaningful, it is best to compare the "good" to the "bad" groups for a given variable and to the biological potential. Assuming normal distribution, the good and bad sites are those that are 1.5 standard deviations above and below the mean or in non-normally distributed systems that are in the upper or lower 15%. Health, building design, feed delivery, and other factors can all be analyzed to determine what the differences are and, therefore, where changes might be made to move the bottom groups closer to the top. I have found this type of analysis to be much more useful in helping management make decisions than merely saying we are better, the same, or worse on a mean than is XYZ system. It also helps the management team understand which problems are easiest to solve with the highest potential returns, enabling them to make timely, high profit changes.

Internal benchmarking

Internal benchmarking is a very effective method of answering the question "Is the system making progress?" There are two basic methods of internal benchmarking: statistical process control (SPC) and mean comparison of period 1 to period 2 (which I prefer). It is outside the scope of this paper to fully discuss the methods behind SPC, but there are numerous published texts on the methods to calculate correct SPC charts. One thing I see in several systems is SPC charts being used that are not robust and, therefore, failing to illustrate true changes in the system. This is not helping the management team understand whether they are making progress or overemphasizing changes that would be expected in normal variation in the system. My recommendation is to spend a significant amount of time reading texts on correct implementation of SPC or continuous improvement programs before establishing a SPC program in a system with which you work. Next, consult one or more persons in the swine industry who have significant experience so you may come to understand the pitfalls and limitations prior to implementing a continuous improvement program.

SPC is a method to determine whether there are mathematically significant changes in the system over time for a given production index. It can be applied essentially to any index that is generated by the record system. One of the challenges in effectively implementing SPC in the nursery/grow/finish records is the time from placement to closeout, resulting in the long production cycle. Because most key production indicators that we use today are end-point measures, significant changes can occur in the system that will not be picked up for months (when those pigs close out) and, therefore, are not detected by SPC until it is too late. Being creative in selecting which points are plotted for SPC analysis might yield dramatically improved sensitivity in a system and could result in faster implementation of an intervention. An example of this would be charting total mortalities in a given nursery site on a weekly basis. One challenge with SPC analysis is that it does not tell you if the system is doing well or poorly, only whether it is improving. A bad system with high variation can look acceptable in SPC because it never varies out of its bad, high-variant zone.

Means testing between two periods is also a useful internal benchmark. It is obviously not as sensitive as SPC because it looks at large blocks of time instead of looking at every group and, again, is fraught with some of the same lack of sensitivity issues as SPC owing to the fact that calculations use end-point measures. Means testing is, however, an effective tool to compare this year to last year or the first quarter of one year to the first quarter of the next. You can use statistical tests to compare period-to-period. Owing to the nature of the data, it may be beneficial to look at raw means and variances to understand how close those data points are. One common error I see is the determination that two periods being compared are not statistically different, which may result from a lack of power in the statistical analysis or use of an incorrect statistical test, depending on the type of data being compared. Therefore, the art of being able to look at means and determining if a system is making headway may be as important as the hard science.

In all cases with internal benchmarking, being able to track the system and changes that occur in it over time is extremely important. Often, internal benchmarking becomes completely incomprehensible because it is impossible to remember when each of those things occurred. It is extremely important that internal benchmarking not be used alone but as part of a larger system improvement process so that the system itself is well understood prior to making these comparisons.

Summary

Data analysis is best when it is a simple process that defines the objectives in a clear, concise manner. The steps that I like to follow in performing analysis of any set of data are:

1. Understand the recordkeeping system and ensure the validity of the data.

2. Understand the philosophy of the system: Driven by cost or optimum profit?
3. Understand the questions that the system is asking: What are the key drivers, throughput, feed efficiency, etc.?
4. Complete both internal and external benchmarking if the appropriate comparisons are available.
5. Summarize the data in a clear concise manner for management to utilize.

Properly performed records analysis can be one of the most beneficial activities in advising a production system with high expectations. Unfortunately, improperly done, records analysis can also lead to incorrect decisions and management moving the system down a path that is costly and unprofitable.