ELSEVIER
SAUNDERS

VETERINARY
CLINICS
Food Animal Practice

Vet Clin Food Anim 22 (2006) 1–19

Basic Epidemiologic Concepts Related to Assessment of Animal Health and Performance

Pamela L. Ruegg, DVM, MPVM

Department of Dairy Science, University of Wisconsin, 1675 Observatory Drive, Madison, WI 53706, USA

Modern animal production systems are increasingly complex. Sustaining long-term performance on these farms depends on the ability of the farm manager to manage many assets, including people, animals, money, machinery, and land. Veterinarians have an important role as experts about animal health and performance. They depend on various types of data to monitor animal health, assure animal well-being, and assess farm profitability. Critical control points for farm profitability include the level of production and quality of the specific commodity (eg, meat or milk), management of animal health, animal comfort, reproductive management, and oversight of the replacement and nutritional programs. A variety of data sources is available to help assess these critical control points. Data sources range from notes written on calendars to sophisticated on-farm computer systems and external industry sources, such as Dairy Herd Improvement Association (DHIA) records and Cattle-Fax. Data sources used for farm analyses vary, depending on the size of the production unit and the technical ability of the farm owner. For example, a recent survey of 587 Wisconsin dairy producers asked managers about the use of computerized records to track antibiotic treatments. The survey found that $<4\%$ of those farms with ≤ 100 lactating cows used computerized records while 65% of Wisconsin dairy producers with >200 lactating cows used computers to track treatments. On small farms, handwritten records may be sufficient for assessing many key performance indicators. However, analysis requires a system designed to be easily accessible. On larger farms, analysis of data can be time-consuming and is unlikely to be performed on a regular basis unless the farm manager or consultant is highly motivated and the data are collected in a manner that allows for easy

E-mail address: plruegg@wisc.edu

0749-0720/06/$ - see front matter © 2006 Elsevier Inc. All rights reserved.
doi:10.1016/j.cvfa.2005.12.002

summarization. An unsuitable form of data collection can make herd performance problems hard to recognize. On some operations, regardless of herd size, data may be collected on paper but never entered into a computer or collected in a manner that accommodates easy summarization (Fig. 1). This type of data has limited value because it can't easily be analyzed to determine trends or summarized in a fashion that guides management decisions.

Veterinarians are important advisors. They help farmers make sound decisions based upon valid interpretation of data. The identification of key indicators of animal performance and the ability to differentiate between values that reflect normal biologic variation and those that require intervention can be challenging. As part of this advisory role, the veterinary consultant must understand the strengths and weaknesses of data, accurately assess production trends, and evaluate the results of management changes. This article describes some basic epidemiologic concepts about animal performance data. These concepts equip veterinary practitioners with the tools they need to give the best advice. More complex use of data is described in other articles in this volume.

General concepts of epidemiology relevant to animal agriculture

Epidemiology concerns the assessment and control of disease in populations of humans and animals. The related terms "epizootiology" and

Fig. 1. Calendar obtained from a commercial dairy farm. Data collected in this way is virtually unusable.

“epizootic” were once used to refer specifically to diseases occurring in animal populations. However, these terms are now considered outdated. The term “veterinary epidemiology” is generally preferred because many diseases can affect both animals and humans. Populations can consist of communities, distinct geographical areas, defined production units (eg, farms), or subunits (eg, pens). Many epidemiologic methods were initially developed to control highly infectious diseases that could rapidly spread through regional populations and that often resulted in devastating outcomes, such as decimation of entire animal populations. The characterization and role of infectious agents in transmission of disease is generally termed “qualitative epidemiology.” The term quantitative epidemiology commonly refers to the use of epidemiologic methods to measure the amount of disease and the use of comparative observational studies and statistical techniques to evaluate disease and the impact of risk factors on the occurrence of disease. Methods used in quantitative veterinary epidemiology have often originated from national policy programs used to control diseases of public health significance (eg, tuberculosis or brucellosis). Many of these methods are highly relevant for veterinarians involved in modern animal agriculture. Animal agriculture has evolved rapidly from small, isolated herds and flocks, where management was based on individual animals, to larger entities, where management is based on groups of animals. This trend has increased the necessity for veterinarians to develop quantitative skills. Likewise, many food animal veterinarians intuitively perform some type of qualitative epidemiology daily. The challenge for veterinarians is to identify inherent biases in data and to be able to accurately synthesize various data sources to arrive at sound recommendations for their clients.

Understanding performance and disease data

The practitioner should consider the following questions before beginning an assessment of animal health and production data:

- What type of data is available and what are the inherent characteristics of that data?
- What is the best way to summarize this data and arrive at solid conclusions about performance?
- What management decisions will I make based on this data?

Type of management decisions

In some herds, considerable collected information is rarely used for making decisions. Examples of such data include routine reports generated by herd management software or performance monitoring programs, such as DHIA. Often, this data have potential value for making daily decisions but are overlooked because of producer apathy, the perception that data are inaccurate, the complex analysis required to properly use the data, or insufficient time to assess herd performance. When data are routinely collected and not used, that data may become steadily less reliable because the people gathering the data perceive that their efforts are not valued. As data become unreliable a cycle is generated. The data are not used because the information is considered unreliable. Because the information is not used, data collectors become more careless, making data even less reliable. A typical example of unreliable data is the analysis of reasons for culling dairy cows. Many dairy herds record just one reason why each cow is sold. That reason could be, for example, low production, mastitis, or infertility. Analysis of this data often results in misleading conclusions because, while just one reason may be recorded, the decision to replace a dairy cow is generally based several reasons. A more complex assessment of the data could increase the value of the data and ultimately result in more accurate and complete recording of culling reasons. A small volume of accurate data is more useful than large volumes of unreliable data. Veterinary consultants should discourage the collection of data that are rarely used. Efforts at collecting rarely used data should be redirected to efforts at summarizing data that contribute to a decision-making process.

Thresholds versus averages

Before data are summarized, their underlying characteristics and expected use should be considered. Some management decisions are based on identification of animals that meet or exceed a specific threshold. Other decisions may require the determination of a population average. The analysis of nutritional programs by using body condition scores (BCSs) is an excellent example of when thresholds are more useful than population means. Body condition scoring is often performed for beef cattle using a scale of 1 to 9. A BCS of 5 to 6 is considered ideal for cows throughout the production cycle. However, BCSs are used to assess the adequacy of the nutritional program. Thus, the key management decision is based on identification of animals that are either too thin (ie, scores 1–3) or too fat (ie, scores 8–9). The use of an average score can lead to inaccurate decisions because important management decisions are based on an assessment of how many animals fall outside the desired BCS range rather than an assessment of a population average for BCS (Fig. 2). Unless the scores are all extreme, the average BCS of a group will rarely result in a defined management outcome.

Types of data

In general, data are considered either qualitative or quantitative. Qualitative data has to do with characteristics (eg, breed, gender), or yes-or-no data (eg, pregnancy, disease status). Sometimes, qualitative data are called categorical data. This type of data may be the result of a measurement but the

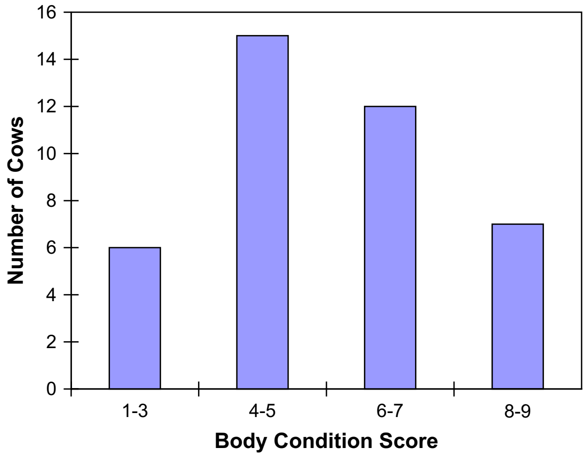

Fig. 2. BCS histogram obtained from beef cattle. This data has an average BCS of 5.4, but 33% of the scores are in unacceptable ranges. (The ideal BCS for all type of cattle is approximately 4–5 [1].)

measurement is used to place the animal within a category. For example, body temperature may be routinely measured on a group of cattle with the objective of categorizing the animals as healthy or febrile, depending on whether or not an animal's temperature exceeds a predetermined threshold, such as 102.5°F. Obviously, measuring the average body temperature of a pen of cattle would be meaningless. What is meaningful is determining how many animals exceed a certain standard for temperature. That determination can be made only by measuring each animal's temperature against that standard. Qualitative data can be referred to as nominal, which normally corresponds to a name or word (eg, breed, gender, pregnancy status). Quantitative data can also be referred to as ordinal, and is used if there is a comparative and progressive basis to the categories (ie, baby calf, heifer, adult). For example, fecal consistency of calves may be assessed using a score of 0 = normal; 1 = semiformed or pasty; 2 = loose but with enough consistency to stay on bedding; and 3 = watery. An increasing score on an ordinal scale will generally be associated with a trend such as deterioration or improvement. Analysis of all types of qualitative data is based on counting events. The data are best summarized by frequency distributions or rates. Generally, the categories have a clearly defined qualitative basis (Table 1). In production agriculture, qualitative data are often useful to set herd goals for animal performance. See related articles in this volume for commodity specific goals. The veterinarian comparing herd data against performance goals should thoroughly understand the context of industry goals. Many data sets used to create performance goals may have inherent biases. Also, the data may be obsolete because of changes in the industry.

Table 1
Frequency of udder hygiene scores for a dairy herd

Udder hygiene score	Frequency	Relative frequency
1: Very clean	40	9.7%
2: Clean	228	55.3%
3: Dirty	105	25.5%
4: Very dirty	39	9.5%
Total	412	100.0%

Quantitative data are based on actual measurements with numerical meaning. When measurements can take any value in a range, the data are termed "continuous." Examples include data on age at weaning, pounds of milk produced, number of pigs weaned, daily weight gain from birth to weaning, and scrotal circumference. When gaps occur between values, the data are referred to as "discrete." Examples include data on BCS, locomotion score, and parity. Some inherently discrete data are commonly summarized using averages. Examples of this kind of data include data on serum antibody titers and minimum inhibitory concentrations of antibiotics. For most discrete data, the most expeditious use is summarization using categories. When categories are used to group continuous data (eg, breeding intervals or percent linear score 0–4), the data are treated as though they were discrete. When clearly defined categories are meaningful, frequency distributions, rates, and ratios are often used to summarize discrete quantitative data (Table 2).

Quantitative data that are truly continuous are summarized by measures of central tendency and measures of dispersion. The objective of these summary statistics is to characterize group performance. Most continuous herd-performance data are summarized using averages. However, measures of variation are often omitted. This oversight can misrepresent group performance, especially the performance of small herds, and may result in incorrect management decisions. For example, assume hypothetical Group A consists of 100 cows. In this group, 50 of the cows produce 40 lbs of milk and 50 of the cows produce 60 lbs of milk. The population mean milk production for Group A would be 50 lbs per cow with a standard deviation of

Table 2
Categorization of continuous data for linear somatic cell score from a 172-cow Ohio dairy farm

	No. of cows			Percentage of cows		
Parity	LSCS 0–3	LSCS 4–6	LSCS $\geq$ 7	LSCS 0–3	LSCS 4–6	LSCS $\geq$ 7
1	41	15	7	68	25	7
2	26	8	3	70	22	8
3+	31	18	9	53	31	16
Total	98	41	16	63	26	10

Abbreviation: LSCS, linear somatic cell score.

10.1 lbs. The performance of Group A is generally consistent and the overall performance may be judged as poor or good, depending on group performance goals. In contrast, consider hypothetical Group B, also consisting of 100 cows. In Group B, 50 of the cows produce 10 lbs of milk and 50 produce 90 lbs of milk. The population mean milk production for Group B is identical to that of Group A (ie, 50 lbs per cow) but the standard deviation is 40.2 lbs/cow. The performance of Group B is highly variable and clearly some problem needs to be addressed. This hypothetical example is obviously extreme. However, it does illustrate the problem of relying solely on group averages. See the article by Reneau and Lukas elsewhere in this volume for more on the importance of understanding the role of variation in identifying changes in animal performance.

The difference between qualitative and quantitative data may be confusing because qualitative data are nonnumeric in nature but are summarized by counts. Another source of confusion is the tendency to categorize quantitative data. An example of assigning categories to quantitative data is when a dairy farm groups "low producers" as cows with daily milk yields of <50 lbs; "average producers" as cows with daily milk yields of 51 to 85 lbs; and "high producers" as cows with daily milk yields >85 lbs. The resulting categories are analyzed as qualitative data. The value of information is considered greatest for continuous data and progressively less valuable for discrete, ordinal, and nominal data, respectively.

Rates, ratios and proportions

Analysis of qualitative and discrete data is based on counting defined events and dividing by the appropriate population (eg, animals at risk for the event) within a specified time period. The most commonly used descriptive statistics for this type of data are proportions, rates, and ratios. Proportions relate the frequency of the animals with the condition of interest to the larger population to which they belong (eg, herd or flock). The numerator is always included in the denominator and the denominator is always the population at risk. For example, to characterize pinkeye treatments in a cow-calf herd at a point in time, the following proportion could be used:

Number of animals treated for pinkeye ÷ number of animals in the herd

In contrast to a proportion, the essential features of a rate are that the numerator is included in the denominator and a time period of observation is specified. Rates can be used to describe changes in the condition of interest over time. Animals included in the numerator are always limited to the same time period as the denominator. When the rate refers to the population in general, rather than a specific subpopulation, it is referred to as a crude rate. To calculate the crude rate of pinkeye treatments, the proportion described above would be changed as follows:

Number of animals treated for pinkeye in June ÷
average number of animals in the herd in June

The determination of which animals to include in the denominator depends on the rigor needed in the analysis. In most instances, when the population of the group is relatively stable, a simple average of the population is sufficient. The average in this case is determined by adding the number of animals present in the herd at the beginning of the time period to the number of animals present at the end the time period and dividing by 2. In other instances, more rigor may be required or the population may be very dynamic and the use of incidence-density (described below) may be more appropriate.

Ratios are used to compare the frequency of two different outcomes and are useful to describe conditions when determining an accurate denominator is difficult. The numerator is not necessarily part of the denominator. An example of a ratio would be one used to describe abortion patterns:

Number of cows that aborted ÷ number of cows with confirmed pregnancies

Not all pregnant animals would have been confirmed. Therefore, animals that aborted but had not yet been confirmed pregnant could be included in the numerator but would not be included in the denominator.

Defining rates

Defining the parameter of interest

Numerous rates are routinely calculated to monitor animal performance. These statistics may be calculated by herd management software, industry programs, or by hand. Often, the indices determined using different programs do not agree. Thus, the veterinary consultant must understand the underlying formulas used in the calculations.

The initial step is to define the population. Which animals should be included and, more importantly, which animals should be excluded? Definitions for both the numerator and the denominator need to be carefully considered. For the numerator, the key issue is defining the disease or production target of interest. For the denominator, the key issue is to define the population at risk. For many diseases, risk varies depending on stage of production cycle or age. The veterinarian may want to define the at-risk population narrowly to ensure that changes can be recognized. The incidence of postparturient paresis (ie, milk fever) in dairy cattle provides an excellent example of differential risk. Milk fever can occur in any age of animal or during any stage of lactation, but the risk of this disease is considerably higher in mature periparturient cows than it is for either first lactation animals or cows in the later stages of lactation. While all animals are technically at risk, the inclusion of the entire population of lactating cows in the denominator may hamper detection of important changes in disease incidence (Table 3).

The numerator should include animals that have experienced the event of interest. These events may be based on the occurrence of symptoms (eg, diarrhea, abnormal milk, retained placenta) or events (eg, twin births, failure to reach a performance target). Detection bias can be a significant problem on many farms. In many instances, treatments, rather than detection of disease or symptoms, are recorded in computerized software. This means that animals that exhibit mild signs of disease but are not treated may not be noted. The veterinarian should ensure that personnel responsible for recording events understand what to look for with each disease or event of interest. Practicing veterinarians should not feel constrained by technical ‘‘rules’’ related to the development of rates, but should feel free to evaluate herds based on useful definitions of populations at risk and diagnostic criteria for events of interest pertinent to making farm decisions.

Defining the time span

A critical step in the calculation of rates is to define the time span. A basic rule of rates is that each animal should only experience the event of interest once during a time period. Therefore, the time period should be defined so that it encompasses a unit that is meaningful to the production unit and the disease. Typically, rates are calculated using one of three methods. The first is the rolling time frame or rolling method. The second is the cohort approach. The third is the current-data method. When the rolling method is used, all of the herd’s animals that fit the defined population criteria are included in the statistic for every time period regardless of when the event

Table 3
Different populations at risk of disease in a hypothetical 500-cow dairy herd

Population at risk used in calculation	Number of animals calved in time period	Number of cases of milk fever	Estimated incidence
All lactating cows (n = 500)	NA	4	4/500 = 0.008 or < 1%
All cows that calved in the time period	80	4	4/80 = 0.05 or 5%
Only multiparous cows that calved in the time period[a]	40	4	4/40 = 0.10 or 10%

Abbreviation: NA, not applicable.
[a] These cows are at the highest risk for the disease.

occurred. With this method, computation is easy. However, the method is very slow to demonstrate changes in recent performance and may not be useful for evaluating recent management changes.

The cohort method is calculated by following a cohort of animals defined by some characteristic. For example, reproductive performance, production, or somatic cell count could be followed for cohorts of cows defined by month of calving. This method can be more abstract to interpret because it may require calculations involving the same statistics for several cohorts. The method is often useful for retrospective analyses but can be unwieldy and require complex interpretation.

The current-data method is calculated by including a cohort of only animals with the most current data in the index. Using this approach, peak milk production would be calculated only for the cohort of animals that were peaking. For example, peak milk could be defined as the test-day milk for animals that were 40 to 70 days in milk at the last test. Another use for the current method is in calculating days to first service for a cohort of cows that received a first service in the specified time period (eg, 1 month). This method has the advantage of being very current and reflecting the current herd situation. However, this method can be misleading if there are few animals that meet the criteria of the cohort. Table 4 illustrates differences for several management parameters calculated using each of three approaches in one herd. Each of the methods can be used successfully, provided the assessor is aware of the underlying definitions and populations.

Common morbidity rates

Morbidity rates describe the level of disease in a population (Table 5). Crude rates specify neither disease nor host characteristics and differ depending upon whether new cases or existing cases are of interest. Incidence rates describe the probability of a new case developing during a stated time interval. For most diseases, incidence rates are the most important measures for monitoring the success (or failure) of disease control strategies. Cases

Table 4
Management indices in one herd using three methods of determining the population

Parameter	Rolling[a] (n)	Cohort[b] (n)	Current[c] (n)
Peak milk (lb.)	110 (84)	102 (12)	113 (10)
Days to first service	79 (102)	54 (2)	80 (4)
Somatic cell count	NC	291 (20)	286 (148)

Abbreviations: N, number of cows; NC, not calculated.

[a] Data included for all lactating cows with non-zero values.

[b] Data included for cows that calved in January.

[c] Data for peak milk includes cows that were 40–70 days in milk at the last test date. Data for days to first service includes cows that received their first breeding in March; Somatic cell count data included for all cows that had somatic cell count values at the last test date.

Table 5
Rate of pneumonia treatments for calves (n = 10)

	Week 1								Week 2							
Calf	1	2	3	4	5	6	7	DAR	8	9	10	11	12	13	14	DAR
1	H	H	H	H	H	H	H	7	H	H	H	H	H	H	H	7
2	H	H	H	H	**RX**	RX	RX	4	RX	H	H	H	H	H	H	6
3	H	H	H	H	H	H	H	7	**RX**	RX	RX	RX	H	H	H	3
4	H	H	H	H	H	H	H	7	D[b]							0
5	H	H	H	H	H	H	**RX**	6	RX	RX	RX	H	H	H	H	4
6	H	H	H	H	H	H	H	7	H	**RX**	RX	RX	RX	H	H	3
7	H	D						1								0
8	H	H	H	H	H	H	H	7	**RX**	RX	RX	RX	RX	RX	H	1
9	H	H	H	H	H	H	H	7	**RX**	RX	RX	RX	H	H	H	3
10	H	H	H	**RX**	RX	RX	RX	3	H	H	H	H	H	H	H	7
								56								34

	Week 1	Week 2
Average population at Risk	(10 + 9)/2 = 9.5	(8 + 8)/2 = 8
Days at risk	56 calf-days	34 calf-days
New treatments	3	4
Total treatments	3	6
Incidence of treatment[a]	3/10 = 30%	4/6 = 67%
Prevalence of treatment	3/9.5 = 32%	6/8 = 75%
Incidence density	3/56 = .054 (5.4 treatments per 100 calf days)	4/34 = .118 (11.8 treatments per 100 calf days)
Relative Risk of Treatment in week 2		11.8/5.4 = 2.2

Boldface indicates initial treatment.

Abbreviations: D, died of other cause; DAR, days at risk in time period; H, healthy; RX, treatment.

[a] Denominator is animals at risk at the begining of the time period.

that exist at the beginning of the time period are not included. The denominator of an incidence rate is usually the population at risk at the beginning of the time period. The numerator is the number of animals that develop the attribute or disease during the time period. Animals that have the condition at the beginning of the time period are not included in either the numerator or the denominator. For example, the incidence of lameness could be calculated as:

Number of cows that became lame during June ÷
average number of cows in the herd in the beginning of June

It can be more difficult to define denominators when the risk period is not as clearly defined or when the disease exists in a subclinical state for long periods of time.

Incidence-density is a more precise measure of incidence and is calculated using the same numerator as above. The denominator is calculated by adding together all the days at risk for each individual in the population at risk. An animal ceases to be at risk once it develops the disease. The unit is therefore the number of cases, treatments, or events per animal-day at risk (Table 5). This figure is especially useful when used in large populations. However, collecting data for this method requires excellent record-keeping to monitor dynamic animal populations.

Prevalence rate is a static measure of disease frequency. It is the fraction of the population that is diseased or possesses the attribute at any one point in time. The point in time can be a single day or may be a defined time period, such as a month (period prevalence). Both new and existing cases are counted. For example, the monthly prevalence of lame cows would be calculated as:

Number of lame cows in June ÷ average number of cows in the herd in June

Obviously, for rates related to morbidity to be reliable, they must be based on precise definition and accurate detection of disease. For many diseases, the difference between a new case and an existing case may be subtle and depend on meticulous detection methods.

Chronic diseases can have higher prevalence than incidence, as the number of animals at risk will steadily decrease as more animals become diseased. The prevalence of disease within a herd is a function of incidence and duration. This relationship can be demonstrated by using hospital pen populations in animal production units to represent prevalence. Consider a hypothetical dairy herd containing 1000 lactating cows. This herd experiences one new case of clinical mastitis per day. To calculate incidence, the numerator would be 30 and the denominator 1000. Monthly incidence is thus 3%. The sick-pen population would represent prevalence. That is, the number of cows sick per 1000 cows per day. If the duration in the sick pen were 1 day, one cow would enter and leave the sick pen each day, so the prevalence of clinical mastitis would always be one case per 1000 cows per day. If the incidence doubled to two new cases per day but the duration in the sick pen remained 1 day, the sick-pen population (prevalence) would also double and would be stable at two cases per 1000 cows per day. Alternatively, if the incidence of mastitis remained stable at one case per day, but the duration of time in the sick pen doubled to 2 days, the prevalence (sick-pen population) would also increase to two cases per 1000 cows per day. Incidence is generally strongly influenced by herd control strategies. Duration can be influenced by such factors as death, culling, and treatment efficacy. These factors can have a strong influence on estimates of disease prevalence within a herd (Table 5).

Summarizing quantitative data

Computerized herd management programs and industry reports are filled with quantitative data. Interpretation of the data depends on an understanding of the underlying distributions of the population and the typical characteristics of the specific values.

Distributions

Data that are normally distributed form a bell shaped curve and have a mean, median, and mode that are approximately equal (Fig. 3). The mean, median, and mode for the mature equivalent milk yield data shown in Fig. 3 are 26,655, 26,690 and 26,333 lbs, respectively. Continuous data that have a wide range of values usually follow a normal distribution pattern. Most direct production and performance parameters are, to at least some degree, normally distributed.

Data that are not normally distributed are termed "skewed" (Fig. 4). Compared to the mean and the mode, the median is generally more representative of the central tendency of skewed data because it indicates the midpoint of the number of pieces of data and is unaffected by the underlying value of the parameter. The mean, median and mode for the somatic cell count data shown in Fig. 4 are 181,190, 65,000 and 11,000 cells per mL, respectively.

Measures of central tendency

Often, continuous quantitative data are used to characterize animal performance. Three basic measures of central tendency are used to analyze continuous quantitative data. Arithmetic means (ie, averages) are the most commonly used statistic. Means are simple to calculate and appear to be easily understandable. They have an important disadvantage in that they are easily influenced by extreme values, especially in small herds or flocks.

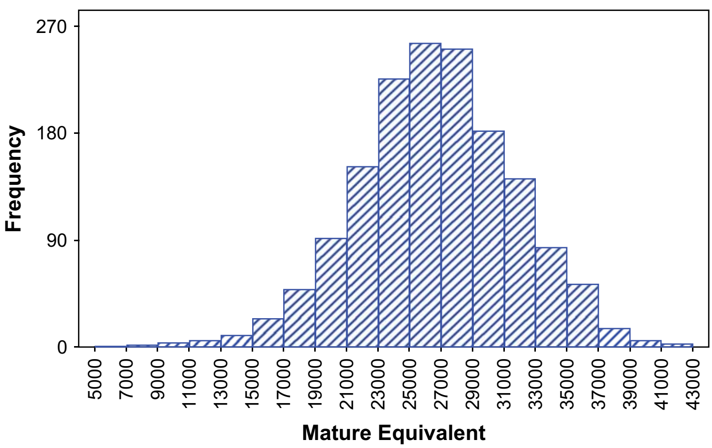

Fig. 3. Frequency histogram of mature equivalent milk yield that approximates a normal distribution on a dairy farm with 1600 cows.

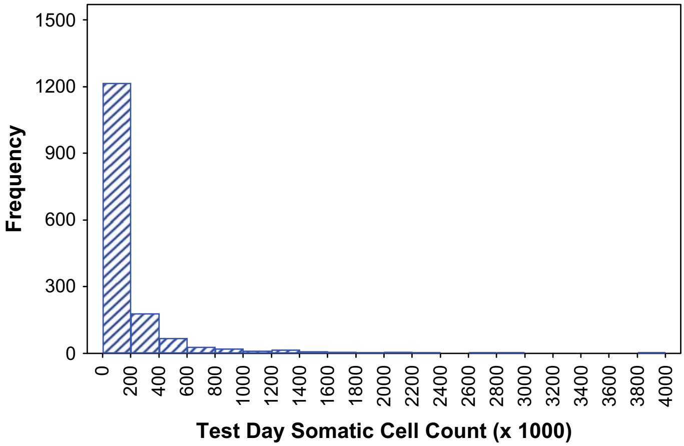

Fig. 4. Frequency histogram of somatic cell count data from a dairy herd with 1600 cows.

This disadvantage is especially important when data are not normally distributed or when few numbers contribute to the average.

When comparing averages over time, rolling averages, such as rolling herd averages, can help smooth out fluctuations stemming from small numbers or isolated extreme values. However, data used to compile rolling averages occurred retrospectively. Thus, rolling averages exhibit considerable lag for determination of herd performance problems. This means that rolling averages may not be as useful as other measures in recognizing and responding to immediate problems (Fig. 5).

Weighted averages are used for some statistics to ensure that each animal contributes equally to a statistic. Similarly, weighted averages may be useful for monitoring statistics on small subsets of animals where the number of animals contributing varies between measurement periods. For example,

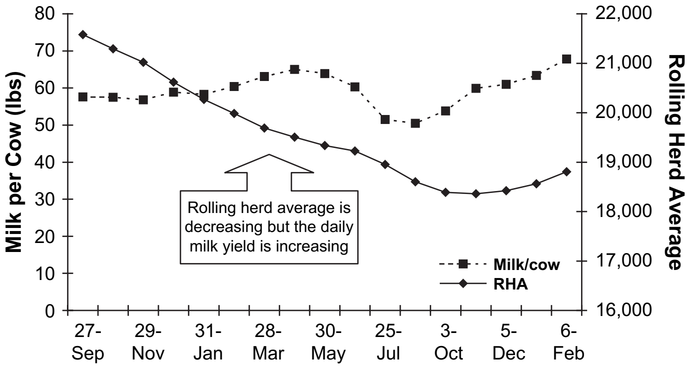

Fig. 5. Rolling herd average versus test-day milk from a 220-cow Indiana dairy.

a farm manager monitoring age at first calving for cows that calved during a specified time period in a small herd (< 100 cows) may find widely varying numbers of animals that contribute to the statistic each month (Table 6). By using a weighted rolling average, variation due to the individual unit (eg, animals, month) is minimized.

Medians (ie, 50th percentile) are an underused statistic. They are independent of the value of the underlying measurement and are not routinely found on most performance reports. A median is simply the midpoint of an ordered range of values. Any time that data is grouped by percentiles or ranked, a median can be estimated. In the field, medians can be easily calculated from data charted on histogram-type forms (Table 7). Medians are not affected by extreme values and are especially useful when data are not normally distributed. Many of the values routinely examined to assess animal performance are not normally distributed and are better represented by medians.

The mode is simply the most frequently occurring value. The mode is rarely used but may be useful if one value makes up a large portion of the data set. A disadvantage of using mode as a measure of central tendency is that there can be several different modes in any individual data set. Modes are not generally calculated in sources available to veterinarians.

Measures of dispersion

The standard deviation is one measure of variability. The standard deviation defines the width of a normal distribution. In a normal distribution, 65% of values are within one standard deviation of the mean, 95% of values are within two standard deviations of a mean, and 99% are within three

Table 6
Calculation of age at first calving using various methods

Month	N	Average days to first calving[a]	Four-month weighted rolling average[b]	Four-month rolling average[c]
July	4	754	NA	NA
August	1	735	NA	NA
September	3	763	NA	NA
October	1	784	758	759
November	5	769	765	763
December	8	746	758	766
January	8	756	767	764
February	9	755	755	757
March	2	725	750	746
Mean		754	754	

Abbreviations: N, number of cows calving for first time; NA, not applicable.

[a] Calculated as a simple average of the number of animals contributing each month.

[b] Calculated over 4 months by weighting each month by the number of animals contributing. For example for October: [(4 × 754) + (1 × 735) + (3 × 763) + (1 × 784)] ÷ 9.

[c] Calculated by adding 7 months of averages and dividing by 4.

Table 7
Body condition scoring form GROUP: CLOSE UP DRY DATE: Jan 2006 SCORER: PLR

	Number of cows																								
BCS	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25
>4.00	X	X	X	X	X	X	X	X																	
4.00	X	X	X	X	X	X	X	X	X																
3.75	X	X	X																						
3.50	X	X																							
3.25																									
3.00																									
2.75																									
2.50																									
2.25																									
2.00																									
<2.00																									

Median = 4.00.

standard deviations. The coefficient of variation is the standard deviation divided by the mean. One rule of thumb asserts that a coefficient of variation of more than 30 (ie, the standard deviation greater than 30% of the mean) indicates a great deal of variability in the data. The standard deviation is an important component of statistical control process analysis. See the related chapter in this issue by Reneau and Lukas about statistical control process analysis. Herd problems, as opposed to individual animal problems, can be identified by comparing the mean value to the standard deviation. In Table 8, the mean age at first calving for both farms is too old, but the standard deviation for Farm 1 indicates that the performance at that farm is consistently poor, as demonstrated by the relatively small standard deviation, in comparison with the highly variable performance of Farm 2. Probably Farm 2 has a small number of animals influencing the average while the performance of most animals may be acceptable. Many recommendations for herd management are directed at increasing uniformity of performance. Decreases in the standard deviation for a herd parameter may be a good indication that overall management of that area has improved.

Variability in data can make it difficult to interpret averages. The standard deviation of an average is called the standard error. Standard errors are calculated by dividing the standard deviation by the square root of the number of pieces of data in the average (N). A rough approximation of the standard error can be estimated for small data sets (ie, <15 data points) by dividing the range, which is the difference between the largest

Table 8
Age at first calving for two herds

Farm	No. of cows	Mean age at first calving (days)	Standard deviation (days)
1	39	1043	316
2	36	1192	710

and the smallest value, by N, the number of pieces of data. Obviously, the standard error decreases as N increases and this is why analysts can be more confident of the population means of larger data sets. Confidence intervals are calculated using standard errors. An average plus or minus approximately two standard errors will account for 95% of the possible values of that average. Confidence intervals can be used to differentiate a mean from a previous level. A value outside of the confidence interval is significantly different from the average. A value that is different than the average but within the confidence interval may be the result of chance. See the article by Slenning in this volume for more information about comparing averages. Understanding the influence of variation is important in interpreting responses to management change.

A large amount of variability in a performance parameter can make it difficult to determine if differences are due to management changes or other influences. For example, Table 9 demonstrates every-other-day milk shipments for one dairy with 100 cows for two separate months. Due to heat stress, milk shipments varied considerably more in August than in January. In this herd, a 500-lb increase above the average shipped in the bulk tank in August would still fall within the 95% confidence interval. By comparison, the same increase in January would exceed the upper bound of the 95% confidence interval. In other words, the response to a small management change will be easier to detect in January than in August because normal daily production variations are smaller in January.

As noted above, standard error is calculated by dividing the standard deviation by the square of the number of pieces of data. It is important to recognize that confidence intervals decrease as the sample size (ie, population or subpopulations included in the calculation) increases. This means that management indices are more useful for measuring change related to large herds than small herds because the effect of individual cow variations is minimized when many pieces of data are included in the calculation. An adequate sample size is especially important for parameters with more variations.

Confounders of production

When evaluating the relationship between two factors, such as between nutritional management and milk produced per cow per day, other

Table 9
Every-other-day milk shipments for one dairy in August and January

Parameter	August	January
Average milk shipped (lbs)	13,541	15,005
Standard deviation	1,701	503
Standard error	439	130
95% Confidence interval	12,680–14,402	14,751–15,260
Coefficient of variation	13%	3%

Table 10
Conception rate by artificial insemination technician

Technician	No. bred	No. conceived	Conception rate
A	360	198	0.55
B	200	88	0.44
Total	560	286	

extraneous factors may affect the relationship that is being studied. The epidemiologic term for factors that influence outcome variables, such as milk production, is "confounder". Confounders are risk factors associated with both the outcome variable being assessed and the risk factor being considered. Essentially, a confounder is a factor that when controlled may reduce or eliminate the effect of the variable being studied.

Many factors can confuse interpretation of production values associated with animal performance. In the dairy industry, confounders for milk production may include stage of lactation, number of times milked, percent of herd in first lactation, fat percentage, breed, culling patterns, and season. Other confounders exist for other animal production systems. When assessing management changes, the effects of these confounders must be considered. In the dairy industry, for example, if most of the herd is in early lactation when instituting a management change, production will naturally decrease as the herd average days-in-milk increases. Unless the confounder is recognized, the management change might be interpreted as detrimental rather than beneficial. Adjusted performance measures may be used to offset the bias stemming from confounders. A special type of confounding can occur when an unrecognized confounder can actually reverse the direction of an association. This type of confounding is referred to as Simpson's paradox. Simpson's paradox can be demonstrated by a hypothetical comparison of conception rates involving two artificial insemination technicians. A

Table 11
Conception rate by artificial insemination technician for repeat breeders

Technician	No. bred	No. conceived	Conception rate
A	120	18	0.15
B	120	24	0.20
Total	240	42	

Table 12
Conception rate by artificial insemination technician for first services

Technician	No. bred	No. conceived	Conception rate
A	240	180	0.75
B	80	64	0.80
Total	320	244	

practitioner reviewing such data in Table 10 may feel compelled to send Technician B back to breeding school. The practitioner would be wise, however, to explore the data in more depth before proceeding (Tables 11 and 12). The type of animal being bred is confounding the relationship between technician and conception rate. In this example, Technician B had a higher conception rate for both repeat breeders and first services. Technician A however, bred more first services. When the data was combined, the resulting conception rate was misleading.

Summary

Veterinarians have an important role in advising clients regarding the use of farm management data. An understanding of the basic concepts of veterinary epidemiology is important for accurate assessment of animal health and performance. Data collected should be used on a regular basis and evaluated in a simple fashion that properly characterizes the group, herd, or individual animal.

Further readings

Dohoo I, Martin W, Stryhm H. Veterinary epidemiologic research. Charlottetown (Canada): AVC Inc.; 2003.

Martin SW, Meek AW, Willeberg P. Veterinary epidemiology. Ames (IA): Iowa State University Press; 1987.

Noordhuizen JPTM, Frankena K, Thrusfield MV, et al. Application of quantitative methods in veterinary epidemiology. Wageningen (The Netherlands): Wageningen Press; 2001.

Schwabe CW, Riemann HP, Franti CE. Epidemiology in veterinary practice. Philadelphia: Lea and Febiger; 1977.

Smith RD. Veterinary clinical epidemiology. Boston: Butterworths; 1991.

Van Belle G. Statistical rules of thumb. New York: John Wiley and Sons; 2002.

Reference

[1] Browne MF, Hall JB, et al. Body condition scoring beef cows. Virginia Cooperative Extension. Available at: http://www.ext.vt.edu/pubs/beef/400-795/400-795.html. Accessed December 11, 2005.