Design and Performance of Statistical Process Control Charts Applied to Estrous Detection Efficiency¹

A. de Vries² and B. J. Conlin

Department of Animal Science,
University of Minnesota, Saint Paul 55108

ABSTRACT

Statistical process control (SPC) charts to monitor production processes have not been widely used in dairy management. Shewhart and cumulative sum (cusum) control charts were designed to determine true changes in estrous detection efficiency (EDE) amidst normal variation in dairy cattle. A stochastic simulation model was used to track performance over time of individual cows in herds of 100 and 1000 cows. Estrous detection ratios (EDR), calculated as observed estruses divided by estimated estrous days (in periods of 1 to 60 d), were used to monitor EDE. Control charts for EDR, using normal and binomial distributions, were designed at 0.65 EDE for both herd sizes; then EDE was set to 0.65 (no change), 0.55, 0.45, or 0.35 and average days to the first detection signal (ATS) in 400 runs was determined. Observed ATS at 0.65 EDE could differ from the target ATS, depending on the SPC chart design and estimated proportions of estrous days for inseminated cows. Observed ATS were shorter for larger changes in EDE and for the 1000-cow herd. Observed ATS for a change to 0.55 EDE were ~ 300 d (100 cows) or 60 d (1000 cows) with the cusum charts. For a change to 0.35 EDE, observed ATS were ~ 50 d (100 cows) and ~ 11 d (1000 cows). Shewhart charts performed similarly or took longer to signal changes depending on period length. Observed ATS on cusum charts were much longer than minimum when non-optimal reference values were used in the design. Observed ATS were also longer when SPC charts were designed with a longer target ATS and change in EDE was small. Control charts using normal and binomial distributions generally performed similarly. Statistical process control charts detected changes in estrous

detection efficiency soon enough to be potentially useful in dairy management.

(**Key words:** estrous detection, statistical process control, Shewhart, cumulative sum)

Abbreviation key: ATS = average time to signal, **cusum** = cumulative sum, **EDE** = estrous detection efficiency, **EDR** = estrous detection ratio, **SPC** = statistical process control.

INTRODUCTION

The results of most processes on dairy farms cannot be predicted with certainty. For example, milk production varies daily and inseminations may or may not result in pregnancy. Most of this variability is the result of many small causes that are not easy to identify. However, an unexpected and undesirable change in the process should be quickly identified. Then corrective action can be taken and the cost of the change can be kept at a minimum (Spahr, 1993; Quesenberry, 1997).

Detection of a change in the process is not trivial, because normal variation in the observations may obscure a real change. Also, not every observation that varies from its expectation deserves an investigation because an in-depth investigation is assumed costly and therefore only warranted if the probability of an undesirable change in the process is sufficiently high.

Statistical process control (SPC) charts are quality control methods that are widely used to monitor the consistency of production processes (Montgomery, 1997). The SPC charts are graphs of observations generated over time with a center line and control limits. The observations are either original measurements or a function of them, such as a cumulative sum. Control limits are calculated from the average variation between the observations that were generated while the process is said to be “in control.” A process that is in control is predictable and stable over time. The majority of observations will fall within the control limits when the process is in control. When an observation falls on or outside the control limits, the SPC chart is said to signal or detect a change. A signal implies that

Received June 3, 2002.

Accepted November 25, 2002.

Corresponding author: A. de Vries; e-mail: devries@animal.ufl.edu.

¹This study was a component of Multi-State Project NC-119, Management Systems for Improved Decision Making and Profitability of Dairy Herds.

²Present address: Dept. of Animal Sciences, Univ. of Florida, Bldg. 459, Shealy Dr., P.O. Box 110910, Gainesville, FL 32611.

Figure 1. Example of a Shewhart chart for estrous detection ratio assuming a binomial distribution (P-chart). Control limits for the P-chart vary per period due to a varying number of estrous days. The P-chart signals in this example in period 28. UCL = upper control limit, LCL = lower control limit. A period is a fixed number of consecutive days, such as a week.

Figure 2. Example of a cusum chart for estrous detection ratio assuming a binomial distribution. The cusum chart signals in this example in period 18. UCL = upper control limit, LCL = lower control limit. A period is a fixed number of consecutive days, such as a week.

there is strong enough evidence that the process has changed and is said to be “out of control.” Thus, SPC charts support the decision when an investigation is warranted.

Various types of SPC charts have been developed. The first SPC charts that were developed are called Shewhart charts. A P-chart is a Shewhart chart where the observations are assumed to follow a binomial distribution (Figure 1). An X-chart is a Shewhart chart for the mean of individual observations that are assumed to follow a normal distribution. The variance of normally distributed observations should be charted separately, because it is independent of the mean. Charts for monitoring the variance are not discussed here but an example can be found in Montgomery (1997).

Cumulative sum (**cusum**) charts are the most sensitive SPC charts to signal a persistent step change in a parameter (Hawkins and Olwell, 1998). Cusum charts require the calculation of both an upward and a downward cumulative sum of observations. Also, the size of the change of most interest needs to be identified when the cusum chart is designed. An example of a binomial cusum chart is shown in Figure 2.

Two types of signaling errors can be made with every observation. A Type I error (false alarm or false positive) is made when an SPC chart signals, but the process has not changed. A Type II error (false negative) is made when an SPC chart does not signal, but the process has changed. The rates of these two types of errors depend on the placement of the control limits. The objective of SPC charts, therefore, is to signal real

changes in a process as soon as possible for a given rate of Type I errors.

Reneau and Kinsel (2001) identified SPC charts as promising monitoring aids in dairy reproductive management and de Vries (2001) designed SPC charts for various variables that monitor estrous detection efficiency (**EDE**), which is defined as the fraction of true estruses that is observed. One variable to monitor EDE is the estrous detection ratio (**EDR**). The EDR is defined as the ratio of the number of observed estruses and the number of cow days where an estrus is possible in a period. The number of cow days where an estrus is possible divided by 21 is a typical measure of the number of expected estruses (Fetrow et al., 1990).

The existing literature on the application of SPC charts in dairy management does not present a formal performance analysis of the proposed SPC charts (de Vries, 2001). The goal of this study, therefore, is to show the variability in performance of various SPC chart designs when applied to EDR and to present some guidelines for designing SPC charts in practical settings.

This study used a stochastic dynamic dairy herd simulation model to measure the performance of SPC charts as a proxy for their performance in practice. Stochastic simulation has been used extensively to measure the performance of SPC charts when analytic results are difficult or impossible to obtain (e.g. Walker et al., 1991; Quesenberry, 1997). However, the use of a complex model of a production system, such as a dairy herd, to evaluate the performance of SPC charts is new.

The objectives of this study were 1) to describe the Shewhart and cusum charts for the binomial and normal model for EDR, 2) to determine what fraction of

possibly pregnant cows should be included in the calculation of EDR, 3) to quantify the performance of SPC charts for monitoring EDR with respect to variations in period length, cusum reference value, and time between Type I errors.

MATERIALS AND METHODS

Probability Distributions for EDR

EDR (p) in period j is calculated as $p_j = x_j/n_j$ where:

- x_j = Number of observed estruses in period j , and
- n_j = Number of potential estrous days in period j = number of cow days after the voluntary waiting period for first insemination of cows confirmed open + fraction $\times$ number of cow days of cows inseminated with insemination result unknown.

A period is a fixed number of consecutive days, for example a week. Every day a cow is present in the herd is counted as a cow day.

The common SPC charts discussed in this study require that a probability distribution is assumed for the EDR. The exact distribution of EDR is unknown, but two reasonable models are the binomial and normal distribution.

The assumption of the binomial distribution is a logical choice if an estrous day is considered a trial and an observed estrus is considered a success. The binomial probability distribution for the number of observed estruses (x) is given by:

$$\text{BIN}(x|n, \mu) = \binom{n}{x} \mu^x (1 - \mu)^{n-x}$$

for $x = 0, 1, 2, \dots, n$

where μ is the mean EDR. The variance of EDR, $\mu(1 - \mu)/n$, and the mean are linked through μ , the only parameter that needs to be estimated. The binomial assumption that the observed estruses per period are independent does not strictly hold because the probability to observe estrus in a cow depends on when her previous estrus occurred (Britt, 1995).

Another plausible model for EDR (p) is the normal probability distribution, which is given by:

$$N(p|\mu, \sigma^2) = 1/[\sigma \cdot (2 \cdot \pi)^{0.5}] \cdot \exp[-(p - \mu)^2/(2 \cdot \sigma^2)]$$

for $-\infty < p < \infty$

where σ^2 is the variance of EDR, which is estimated independently from the mean. The normal distribution

may provide a good fit, because the central limit theorem implies that $z = (p - \mu) \cdot n^{0.5}/\sigma \sim N(0,1)$ when $n \rightarrow \infty$ (Quesenberry, 1997) and provides the foundation that SPC charts based on the normal distribution are appropriate for proportions when the subgroup size n is sufficiently large. The normal assumption that $-\infty < p < \infty$ does not hold, because EDR cannot be negative.

Design of SPC Charts

Formulas for the design of Shewhart charts are presented in Appendix A and were used in this study. The formulas allow for a variable number of estrous days in each period. The standard deviation of the X-chart was calculated using the traditional moving range method (Montgomery, 1997). Roes et al. (1993) presented an overview of different estimators for the standard deviation of individual observations.

Formulas for the design of cusum charts are presented in Appendix B. The algorithmic or decision interval cusum used in this study is generally preferred over the older V-mask cusum (Montgomery, 1997), although they are algebraically equivalent (Hawkins and Olwell, 1998). Design of the cusum charts requires the selection of a reference value, which is a function of the average in control EDR (μ_0) and the change in EDR of most interest (μ_1).

By default, the reference values for binomial cusum charts were calculated from $\mu_1 = 0.75 \cdot \mu_0$ for a downward change in EDR and $\mu_1 = 1.25 \cdot \mu_0$ for an upward change. The change of most interest μ_1 was also used for calculation of the reference values for normal cusum charts. The change from μ_0 to μ_1 measured in units standard deviation varied with the length of the period, because longer periods result in smaller standard deviations. Consequently, the reference values for normal cusum charts were not optimal.

All SPC charts were designed with both upper and lower probability control limits. The probability control limits were set such that the probability of a signal on either the upper or lower control limit was similar or as close as possible to half the target Type I error rate ($\alpha/2$) with α the total Type I error rate per period in this study. The target $\alpha/2$ for each individual control limit was not exactly obtainable on the P-chart, because the control limits for the binomial distribution are necessarily integers. Finding the probability control limits for the P-chart required solving the equations in Appendix A for a given target Type I error rate, which was easily done by trial and error.

Calculation of the cusum control limits is not easy because the probability distribution of the cusum is generally unknown. Control limits for normal cusum charts were obtained with the program ANYGETH

(Hawkins and Olwell, 1998 [Available from <http://www.stat.umn.edu>, School of Statistics, University of Minnesota]). Control limits for binomial cusum charts were obtained with the algorithm described in Appendix C, because ANYGETH could not calculate every desired binomial reference value.

Performance Measures of SPC Charts

For practical dairy herd management purposes, SPC chart performance can be measured by the average time to signal (**ATS**), measured in this study in days. The ATS was defined as the average time until the first signal after the start of monitoring with the SPC chart. The ATS should be long when EDE has not changed and short when EDE has changed.

Three different types of ATS were distinguished: target ATS, design ATS, and observed ATS. The target ATS is the desired in control performance by the decision maker. For example, a target ATS of 365 d and a period length of 7 d implies a target Type I error rate of $7/365 = 0.0192$. The target ATS can typically not be realized exactly with binomial observations due to their discreteness.

The design ATS is the ATS obtained with the chosen control limits when the observations follow the assumed probability distribution exactly. The design ATS for Shewhart charts is easily obtained as period length/ P (signal in a period). This formula holds for both in control and out of control situations. The design ATS for the normal SPC charts equals the target ATS because the normal distribution is continuous. The design ATS for binomial cusum charts is not straightforward, but is typically obtained using Markov chain analysis (Hawkins and Olwell, 1998).

The observed ATS depends on the real distribution of the observations, sampling error, the chosen control limits, and the size of the change in EDE. The observed ATS is equal to the design ATS when there is no sampling error and the observations follow the assumed probability distribution exactly.

Stochastic Dynamic Simulation Model

Results for this study were obtained with a stochastic dynamic dairy herd simulation model based on Monte Carlo simulation (de Vries, 2001). The model simulates individual dairy youngstock and cows daily through time and contains functions for milk production, feed intake, disease, reproduction, and cow replacement. All heifers born in the herd are to be raised to enter the herd at calving. The cow with the lowest expected future profitability, calculated by the model, is culled the day after a heifer enters the herd to maintain a constant herd size.

The scheduling and outcome of events such as estrous detection is subject to random observations from appropriate probability distributions. These functions aim to produce a realistic level of the variation in output from the model that could be observed in practice. Variables such as EDR can be calculated for periods of various lengths.

The validation procedure for reproduction focused on graphical comparisons with data from five Minnesota dairy herds collected using DairyCHAMP from 1994 to 1999. The average number of cows on these farms ranged from 165 to 688. Estrous detection ratios varied from 0.015 to 0.028 for these dairy herds and was 0.0284 for the default scenario in the simulation model (Figure 3). The results showed that the distributions obtained with the simulation model were generally in agreement with results found in practice. Similar graphs with cumulative distribution functions of intervals to first insemination and conception, interval between estruses, and inseminations per conception also indicated agreement with results found in practice (not shown).

Design of Experiments

Eight scenarios were simulated with herds of 100 and 1000 cows and with changes in EDE from 0.65 to 0.55, 0.45, or 0.35, or no change. Each scenario was replicated 400 times (i.e. runs).

A run simulated a herd over 4520 d (12.4 yr). Every run was started with the same steady state herd at 0.65 EDE. In the first 2000 d of a run, the herd was simulated at 0.65 EDE to obtain a steady state herd at the end of day 2000 that was independent of the initial herd. This guaranteed that the SPC results obtained for each run were independent. No data were collected during these 2000 d.

Next, data were collected during an additional 840 d simulated at 0.65 EDE to estimate the parameters for the design of the SPC charts. Preliminary results suggested that 840 d provided enough data to estimate these parameters with sufficient accuracy. The herd remained in steady state during these 2840 d because no changes were implemented. Statistical process control charts were set up at the end of day 2840.

Finally, an additional 1680 d were simulated with either 0.65 EDE or a change to 0.55, 0.45, or 0.35 EDE. The herd was temporarily not in steady state if the level of EDE changed from 0.65 to its lower level. The number of days to first signal on the SPC chart was measured during those 1680 d (4.6 yr). The first signal occurred on the first instance after day 2840 where an observation fell outside a control limit. If an SPC chart did not signal in 1680 d beyond the initiated change,

Figure 3. Cumulative distribution functions of estrous detection ratio calculated for 36 monthly (five Minnesota dairy herds (—)) and 30-d periods (simulation model (---)).

then the number of days to signal for that run was censored at 1680 d. Preliminary results showed that nearly all SPC charts signaled within 1680 d for all runs. Additional motivation to choose intervals of 840 and 1680 d was that they are multiples of 7 and 30 d, the period lengths of initial interest.

The observed ATS was directly calculated as the average days to signal of the 400 runs if no censoring in any run occurred. If at least one run was censored, then a Weibull distribution was fitted on the 400 ordered days to signal using the method of Keats et al. (1997). The observed ATS was then calculated as the mean of the Weibull distribution. The Weibull probability distribution for days to signal (w) is defined as:

$$\text{WEI}(w|\gamma, \beta) = \gamma/\beta^\gamma \cdot w^{\gamma-1} \cdot \exp(-(w/\beta)^\gamma) \text{ for } 0 \leq w < \infty$$

where w is the days to signal, γ is the shape parameter, and β is the scale parameter. The Weibull distribution reduces to the exponential(β) distribution when $\gamma = 1$. Now the estimated observed $\text{ATS} = E(w) = \beta \cdot I([\gamma - 1]/\gamma)$ and $\text{VAR}(w) = [\beta^2 \cdot (I([\gamma + 2]/\gamma) - \{I([\gamma + 1]/\gamma)\}^2)]$.

Target ATS and observed ATS at 0.65 EDE were compared by calculating the t statistic $400^{0.5} \cdot |\text{target ATS} - \text{observed ATS}| / \text{VAR}(w)^{0.5}$ and comparing it to the t -distribution with 399 degrees of freedom.

Estimation of Estrous Days in Possibly Pregnant Cows

Because possibly pregnant cows are either pregnant or open, only a fraction of the possibly pregnant cow

days should be included in the calculation of estrous days. This fraction of possibly pregnant cow days, together with all cow days in confirmed open cows, was the number of estrous days. A fraction of 1 resulted in the maximum number of possible estrous days.

Preliminary analyses (not shown) indicated that variation of the fraction cow days in possibly pregnant cows had a large impact on the observed in control ATS. The optimal fraction to obtain the target ATS could not be observed or calculated directly and was therefore estimated empirically. A fraction 0.4 was used to calculate the number of estrous days from possibly pregnant cow days for the descriptive statistics (Table 1). This will later be shown to be a reasonable choice (Figure 5).

RESULTS

Descriptive Statistics

Steady state summary statistics for four levels of EDE with the 1000-cow herd are described in Table 1. Standard deviations between periods of 1 d are shown. A decrease in EDE from 0.65 to 0.35 both decreased the number of observed estruses per day from 6.20 to 4.45 and increased the number of possible estrous days per day from 349.3 to 468.0. This increase in possible estrous days was entirely due to an increase in the estrous days in confirmed open cows. The total number of possible estrous days in possibly pregnant cows actually decreased from 217.2 to 170.2, because fewer cows were inseminated. After adjusting for preg-

Table 1. Summary statistics for the simulated 1000-cow herd in steady state for four levels of estrous detection efficiency. Period length is 1 d.

Variable	Estrous detection efficiency							
	0.65		0.55		0.45		0.35	
	Average	SD ¹	Average	SD	Average	SD	Average	SD
Milk/cow/year (10 ³ kg)	10.676	0.157	10.659	0.154	10.633	0.146	10.640	0.137
Annual culling rate	0.358	0.354	0.348	0.348	0.330	0.340	0.309	0.326
Interval to first insemination ²	73.1	14.2	79.1	18.4	87.4	23.9	98.7	30.5
SD interval to first insemination ³	16.7	11.1	21.5	14.2	27.8	17.8	35.4	21.1
Interval to conception ²	120.3	43.6	127.0	45.9	134.8	48.4	143.4	50.7
SD interval to conception ³	51.9	29.9	54.5	30.1	56.9	30.0	58.1	29.5
Conception ratio	0.388	0.218	0.400	0.229	0.415	0.243	0.431	0.268
# observed estruses	6.20	2.48	5.79	2.39	5.24	2.28	4.45	2.09
# possible estrous days ⁴	349.3	15.6	371.2	15.7	406.0	16.0	468.0	15.9
– in open cows	132.1	11.0	162.4	12.1	211.4	13.9	297.7	15.1
– in poss. pregn. cows	217.2	13.2	208.8	12.9	194.5	12.3	170.2	11.7
# estrous days ⁵	219.1	11.3	246.1	12.1	289.4	13.5	366.0	14.3
Estrous detection ratio	0.0284	0.0114	0.0236	0.0098	0.0184	0.0079	0.0122	0.0058

¹Standard deviation between periods of 1 d.²Intervals measured in days.³Standard deviation of observations.⁴Possible estrous days = number of estrous days in confirmed open cows + number of estrous days in possibly pregnant cows.⁵Estrous days = number of estrous days in confirmed open cows + 0.4·number of estrous days in possibly pregnant cows.

nancy, the total number of estrous days increased from 219.1 per day (0.65 EDE) to 366.0 (0.35 EDE). The average EDR decreased from 0.0284 to 0.0122. Results for the 100-cow herd were similar, except that the variation between periods was larger due to the smaller number of observations per day.

Figure 4 shows the distribution of EDR for periods of 1, 7, and 30 d for the 100-cow herd at 0.65 EDE. For periods of 1 d, the distribution has two peaks because during many days no estrus was observed. The distribution became more uni-modal with longer periods and larger herd sizes.

Fraction to Estimate the Estrous Days in Possibly Pregnant Cows

Figure 5 shows the relationship between the observed in control ATS and the fraction of possibly pregnant cow days used to estimate the number of estrous days. The SPC charts were designed with a target ATS of 730 d. The standard error of observed ATS estimates in this study was approximately $ATS/400^{0.5}$. A 95% confidence interval was therefore approximately $[0.9 \cdot ATS, 1.1 \cdot ATS]$.

In general, the observed ATS increased from significantly shorter than the target ATS to significantly

Figure 4. Distribution of estrous detection ratio for period lengths of 1 (solid), 7 (striped), and 30 (empty) d for the simulated 100-cow herd.

Figure 5. Effect of fraction estrous days from possibly pregnant cows on observed in control average time to signal (ATS). Target ATS is 730 d. Estrous days = estrous days in confirmed open cows + fraction $\times$ estrous days in possibly pregnant cows. $\blacklozenge$ Shewhart P-chart (binomial distribution), $\blacktriangle$ binomial cusum chart, $\square$ Shewhart X-chart (normal distribution), $\circ$ normal cusum chart.

longer when the fraction was increased from 0 to 1. The observed ATS of the X-chart with 1 d per period remained significantly shorter than the target ATS for all fractions. For fractions close to 1, the observed ATS of the binomial SPC charts approached 2500 d in all six cases. The observed ATS of the normal SPC charts tended to remain closer to the target of 730 d for all fractions than the binomial SPC charts when the period length was 7 or 30 d.

A fraction of 0.4 resulted in an observed ATS that was typically close to the target of 730 d, provided the period length was longer than 1 d. Based on these results, a fraction of 0.4 was used to calculate estrous days from the population of possibly pregnant cow days in the remainder of this study.

Period Length

The observed in control ATS was in many cases significantly shorter or longer than the target ATS of 730 d (Table 2), depending on herd size, type of SPC chart, and period length. The X-charts with period lengths

of 60 d had an observed in control ATS of less than 400 d. The cusums for the 1000-cow herd had an observed in control ATS of more than 2100 d.

Period length had little effect on the out of control cusum chart performance, except when the period length was the limiting factor to observe a shorter ATS. In this study, the observed ATS was at least equal to one period length due to the experimental design.

The cusum charts signaled changes to 0.55 and 0.45 EDE faster than the Shewhart charts in the 100-cow herd. Shewhart and cusum chart performance was similar for the 1000-cow herd, except for period lengths of 5 d when the Shewhart charts took longer to signal the changes to 0.55 and 0.45 EDE.

The observed ATS of the X-chart for the 100-cow herd increased with a decline in EDR for 5 and 10 d in a period, as a result of lower control limits below 0. Because EDR could not be negative, a negative control limit would never trigger a signal. Thus, signals on the X-chart in this case could only be the result of signals on the upper control limits, which were unlikely to occur with a downward change in EDE.

Table 2. Average time to signal (days) for four levels of estrous detection efficiency (EDE) for various period lengths. Target ATS is 730 d. Results for 0.65 EDE indicate the average time between Type I errors.

SPC chart	EDE	EDR ¹	Period length (days)					
			5	10	20	30	40	60
100 cows								
P-chart ²	0.65	0.0284	1944*	1115*	1103*	1024*	853	634
	0.55	0.0236	4048	713	623	486	404	301
	0.45	0.0182	1816	265	156	121	102	102
	0.35	0.0122	374	92	53	46	50	62
X-chart ²	0.65	0.0284	613	879	567*	547*	448*	379*
	0.55	0.0236	1960	2779	643	389	312	263
	0.45	0.0182	40427	6308	207	130	107	103
	0.35	0.0122	— ⁴	4509	67	44	47	63
Bin. cusum ³	0.65	0.0284	639	621*	621*	628*	602*	594*
	0.55	0.0236	315	311	304	306	301	292
	0.45	0.0182	94	93	94	93	97	99
	0.35	0.0122	48	49	50	51	52	63
Nor. cusum ³	0.65	0.0284	658	546*	461*	488*	531*	538*
	0.55	0.0236	364	338	297	301	290	304
	0.45	0.0182	112	103	92	95	98	104
	0.35	0.0122	57	54	52	52	55	66
1000 cows								
P-chart	0.65	0.0284	850	906*	1082*	920*	791	629*
	0.55	0.0236	174	98	65	56	54	64
	0.45	0.0182	25	17	20	30	40	60
	0.35	0.0122	8	10	20	30	40	60
X-chart	0.65	0.0284	970*	935*	564*	469*	427*	395*
	0.55	0.0236	542	192	57	46	50	64
	0.45	0.0182	60	19	20	30	40	60
	0.35	0.0122	11	11	20	30	40	60
Bin. cusum	0.65	0.0284	956*	1014*	1006*	921*	945*	2254*
	0.55	0.0236	59	59	60	57	63	94
	0.45	0.0182	15	16	20	30	40	60
	0.35	0.0122	9	11	20	30	40	60
Nor. cusum	0.65	0.0284	1001*	829	664	683	960*	2177*
	0.55	0.0236	66	61	53	54	61	87
	0.45	0.0182	16	16	20	30	40	60
	0.35	0.0122	10	11	20	30	40	60

¹EDR = estrous detection ratio.²P-chart = Shewhart chart where observations are assumed to follow a binomial distribution; X-chart = Shewhart chart for mean of individual observations that are assumed to follow a normal distribution.³Bin. cusum = Cumulative sum chart where observations are assumed to follow a binomial distribution; Nor. cusum = Cumulative sum chart for mean of individual observations that are assumed to follow a normal distribution.⁴At least 390 runs censored, no ATS calculated.*In control ATS (for EDE = 0.65) differed ($P < 0.001$) from target ATS of 730 d.

It was of interest to know for what combinations of herd size and period length the lower control limit is at least 0 on Shewhart charts so it may trigger a signal. The lower control limit on the X-chart can signal when $\mu_0 - \Phi^{-1}(\text{period length}/\text{target ATS}/2) \cdot \sigma/(\text{number of cows} \cdot \text{period length})^{0.5} \leq 0$ where Φ^{-1} is the inverse of the standard normal distribution and σ is the standard deviation of one cow with a period length of one day. The lower control limit on the P-chart can signal when $\text{BIN}(0|n, \mu_0) = (1 - \mu_0) \cdot \exp(n \cdot \text{number of cows} \cdot \text{period length}) \leq \text{period length}/\text{target ATS}/2$, where n is the expected number of estrous days of a cow in one day. Results from the 1000-cow herd showed $\mu_0 = 0.0284$, $\sigma = 0.3605$, and $n = 0.2191$, which resulted in minimum

period lengths of about 7 to 11 d for the 100-cow herd and 1 or 2 d for the 1000-cow herd, with target ATS ranging from 365 to 1095 d (Table 3). Results for X-charts and P-charts differed slightly due to the discreteness of the binomial distribution.

Effect of Reference Value on the Performance of Cusum Charts

In practice, the actual magnitude of the change in EDE is unknown, and therefore the reference value of a cusum chart is not likely to be optimal. The effect of non-optimal reference values on cusum chart performance was investigated for various levels of EDR of

Table 3. Minimum number of days in a period for Shewhart charts to result in a lower control limit of at least 0 to be able to signal a downward change in estrous detection ratio¹.

Target ATS (days)	100 cows		1000 cows	
	P-chart	X-chart	P-chart	X-chart
365	7	8	1	2
730	8	10	1	2
1095	8	11	2	2

¹Differences between P- (binomial) and X- (normal) charts are due to the discreteness of the binomial distribution.

most interest. Out of control EDR of most interest (μ_1) was set as a fraction (F) of the in control EDR (μ_0), namely $\mu_1 = F \cdot \mu_0$ for the downward cusum, and $\mu_1 = (2 - F) \cdot \mu_0$ for the upward cusum. Optimal μ_1 for changes to 0.55, 0.45, and 0.35 EDE were 0.0236, 0.0184, and 0.0122 respectively (Table 1).

Results in Table 4 showed that the out of control ATS lengthened with reference values that were farther from optimal. The ATS varied more for small changes (e.g. from 0.65 to 0.55 EDE) than for larger changes in EDE.

The observed ATS differed in a few cases greatly from the expected ATS. The design ATS was quite different from the target ATS in some of these cases, because the target ATS was not attainable with that particular SPC chart design. The effects of the differ-

ences between design ATS and target ATS were most profound for reference values that were near optimal for signaling the change to 0.55 EDE.

Effect of Control Limits on Out of Control Performance

The observed in control ATS generally increased with a longer target ATS (Table 5). All SPC charts for the 1000-cow herd had an observed in control ATS that was longer than the target ATS. The largest relative increase (+34%) occurred with the binomial cusum with a target of 365 d. The SPC charts for the 100-cow herd signaled in several cases significantly more or less often than the target. The largest relative increase was +39% and the largest relative decrease was -43% from the target ATS.

The observed out of control ATS also increased with a longer target ATS. For example, when the target ATS increased from 365 d to 1095 d, the observed ATS for the binomial cusum for a change from 0.65 EDE to 0.45 EDE increased from 76 to 104 d (+37%) for the 100-cow herd with period lengths of 14 d. Similarly, the observed ATS increased only from 14 to 16 d (+17%) for the 1000-cow herd with period lengths of 7 d. The out of control ATS varied less, in general, when the observed ATS was shorter, as is the case with larger

Table 4. Effect of out of control estrous detection ratios (EDR) of most interest (μ_1) on cusum chart performance for various true changes in estrous detection efficiency (EDE). μ_1 for a decrease in EDR is calculated as $F \cdot 0.0284$ and for an increase as $(2 - F) \cdot 0.0284$, where 0.0284 is the in control EDR and F is a fraction. The shortest ATS is expected where μ_1 equals the observed out of control EDR. Period length is 7 d and target ATS is 730 d.

EDE	F EDR μ_1^1	0.125 0.0036	0.250 0.0071	0.375 0.0107	0.500 0.0142	0.625 0.0178	0.750 0.0213	0.875 0.0249
100 cows, normal cusum chart								
0.65	0.0284	762	834	852	785	691	583*	499*
0.55	0.0236	1389	1122	887	620	470	353	310
0.45	0.0182	497	306	200	140	116	107	113
0.35	0.0122	96	71	58	53	53	56	67
100 cows, binomial cusum chart								
0.65	0.0284	826	984*	1044*	928*	836	645	528*
0.55	0.0236	566	553	547	460	375	309	286
0.45	0.0182	218	170	135	114	97	93	102
0.35	0.0122	66	51	46	43	44	48	58
1000 cows, normal cusum chart								
0.65	0.0284	1760*	796	718	741	795	852	682
0.55	0.0236	670	240	205	177	120	61	49
0.45	0.0182	49	24	20	17	15	16	20
0.35	0.0122	10	8	8	8	9	10	14
1000 cows, binomial cusum chart								
0.65	0.0284	1423*	1375*	820	838	925*	911*	725
0.55	0.0236	59802	1110	159	125	94	59	48
0.45	0.0182	830	73	19	17	15	15	19
0.35	0.0122	47	11	8	8	8	9	14

*In control ATS (for EDE = 0.65) differed ($P < 0.001$) from target ATS of 730 d.

¹ $\mu_1 = F \cdot 0.0284$.

Table 5. Effect of target average time to signal (target ATS: 365, 730, and 1095 d) on observed ATS (days) for various true changes in estrous detection efficiency (EDE).

SPC chart	EDE	100 cows (14 d in period)			1000 cows (7 d in period)		
		365	730	1095	365	730	1095
P-chart ¹	0.65	505*	1003*	1499*	410	822	1243
	0.55	350	680	904	77	118	176
	0.45	125	213	281	14	18	21
	0.35	46	64	80	8	8	8
X-chart ¹	0.65	408	730	985	483*	888	1279
	0.55	587	1433	2020	136	299	475
	0.45	302	877	1453	19	29	41
	0.35	99	310	619	8	9	10
Bin. cusum ²	0.65	380	628*	817*	490*	908*	1336*
	0.55	223	304	361	45	59	69
	0.45	76	93	104	14	15	16
	0.35	41	49	54	8	9	10
Nor. cusum ²	0.65	285*	495*	626*	457*	852	1159
	0.55	213	302	376	47	61	72
	0.45	77	97	111	14	16	17
	0.35	42	52	58	8	10	11

¹P-chart = Shewhart chart where observations are assumed to follow a binomial distribution; X-chart = Shewhart chart for mean of individual observations that are assumed to follow a normal distribution.

²Bin. cusum = Cumulative sum chart where observations are assumed to follow a binomial distribution; Nor. cusum = Cumulative sum chart for mean of individual observations that are assumed to follow a normal distribution.

*In control ATS (for EDE = 0.65) differed ($P < 0.001$) from respective target ATS of 365, 730, or 1095 d.

changes in EDE, larger herd sizes, and using cusum charts instead of Shewhart charts for small changes.

DISCUSSION

This study focused on the design and performance of SPC charts applied to EDR in dairy herds obtained with a stochastic simulation model. Stochastic simulation is a feasible approach to obtain sufficient observations on SPC chart performance under realistic conditions. A prerequisite is that the variability in the production process is realistically modeled. Like most other simulation models (Sørensen and Enevoldsen, 1992), the simulation model used in this study simulated constant management and constant environment. However, these factors could easily be varied over time if so desired.

It is important to distinguish between phase I and II when SPC charts are discussed (Woodall and Montgomery, 1999). Phase I is an exploratory phase, where the objective is to bring the process in statistical control and estimate parameters that will be used in the (re)design of the SPC charts in phase II. In this study, EDE was in control when the parameters for the SPC charts were estimated during 840 d. Phase II is the actual monitoring phase, which was the last 1680 d in the simulation. Statistical performance analysis applies to phase II applications.

In practice, parameters for the design of SPC charts are likely to be estimated from fewer observations than

were available in this study. Quesenberry (1997) showed that differences between the true parameter value and its estimated value could have a large effect on SPC chart performance. Cusum charts are in general considered to be more sensitive (i.e. less robust) to these differences than Shewhart charts. Self-starting SPC charts have been proposed which omit the need for a large amount of in control data to calculate reliable control limits (Quesenberry, 1997; Hawkins and Olwell, 1998).

P-charts were designed with probability control limits, which is unconventional. Traditional P-charts are designed the same way as the X-charts in this study, except that the theoretical binomial standard deviation is used instead of the moving range estimate of the standard deviation. Such P-charts have control limits with unequal design Type I error rates when EDR $\neq$ 0.5 and were therefore not used.

In general, trends in observed ATS for EDR were in agreement with results obtained for observations that follow exact binomial or normal distributions (Quesenberry, 1997). Larger changes in EDE were typically detected faster than smaller changes. Cusum charts signaled small changes similarly fast or faster than Shewhart charts. The performance of X-charts and P-charts and the performance of normal and binomial cusum charts were in general similar. Changes in EDE were signaled faster in the 1000-cow herd than in the 100-cow herd.

The choice of the fraction to estimate estrous days from possibly pregnant cow days was shown to be important for SPC chart performance. Fetrow et al. (1990) suggested the use of a fraction equal to $(1 - \text{conception ratio})$. The average conception ratio in this study was approximately 0.4, which would result in a fraction of 0.6. A fraction of 0.6 typically resulted in an observed ATS that was longer than the target ATS. The results of this study suggest therefore that $(1 - \text{conception ratio})$ is not the optimal fraction for SPC chart design when monitoring EDR.

A fraction equal to $(1 - \text{conception ratio})$ seemed a logical choice because this proportion of all possible pregnant cows would remain open and thus eligible for renewed estrus. However, the number of estrous days from cows that failed to get pregnant is overestimated for two reasons. First, the average time to the next ovulation is longer after a failed insemination when no insemination takes place (Britt, 1995). This effect was included in the simulation model. Secondly, cows that were not seen in estrus within 42 d after insemination were checked in the model for pregnancy to determine if they were pregnant or not. Thus, most cows had just one opportunity to be in estrus in the 42 d when they were unknown to be pregnant or not. As a result, a smaller fraction would make the opportunity to observe estrus on a given day more similar for inseminated and not inseminated cows, which better satisfies the binomial model assumptions.

The theoretical binomial variance, $n \cdot p \cdot (1 - p)$, is a function of the fraction and was in many cases greater than the observed sample variance when the fraction was equal to $(1 - \text{conception ratio})$. This resulted in control limits for the binomial SPC charts that were too wide and consequently longer in control ATS were observed. Hence, smaller fractions resulted in in control ATS that were more similar to the target ATS.

One can calculate the fraction such that the binomial variance equals the sample variance, but this approach did not result in an acceptable in control ATS when such a fraction was estimated per period.

The chosen fraction also had an effect on the in control ATS for normal SPC charts, even though the sample variance was used to design control limits. Anderson-Darling goodness-of-fit statistics indicated a significant deviation from the normal distribution, but their relationship with the observed in control ATS was not clear. Thus the distribution of EDR is significantly different from the normal distribution.

A fraction of 0.4 was judged to be a reasonable choice for both Shewhart and cusum charts under the conditions in this study. Both types of cusum charts had similar in control performance and worked well with short period lengths, which are preferable for cusum

charts anyway (Hawkins and Olwell, 1998). Both Shewhart charts had poor in control performances for period lengths of 1 d. The in control performance was generally better for longer periods.

Whether a fraction of 0.4 is a good choice under all conditions in practice is not clear. We investigated whether the optimal fraction was dependent on the herd's conception ratio and how close the observed in control ATS was to the target ATS of 730 d when a fraction equal to $(1 - \text{conception ratio})$ was used. We repeated the simulations with conception ratios of 0.2 (study B) and 0.6 (study C) and measured the ATS of the 1000-cow herd with period lengths of 7 d for the four SPC charts and various fractions.

With the fraction increasing from 0.2 to 0.8, the average in control ATS increased from 555 to 1040 d (study B) and from 688 to 1227 d (study C). The binomial cusum chart was the most sensitive to the fractions in both studies B and C. The other three SPC charts showed similar robustness to changes in the fraction.

In all cases, the use of a fraction equal to $(1 - \text{conception ratio})$ resulted to an in control ATS > target ATS; the average in control ATS of the four charts was 1040 d (study B) and 892 d (study C). The average in control ATS for the four SPC charts with a fraction equal to 0.4 was 716 d (study B) and 892 d (study C). The average observed in control ATS with a fraction equal to 0.4 therefore was closer to the target ATS than when a fraction equal to $(1 - \text{conception ratio})$ was used. The average optimal fractions were ~ 0.41 (study B) and ~ 0.23 (study C), but this led for some charts to observed in control ATS that were sufficiently smaller than the target ATS.

Clearly, a longer in control ATS leads to a longer out of control ATS when EDE is changed, as is shown in Table 5. With the fraction varying from 0.2 to 0.8 and a change to 0.55 EDE, the average observed ATS for the four charts varied from 54 to 101 d (study B) and from 235 to 450 d (study C). For a change to 0.35 EDE, the average observed in control ATS is 8 d (study B) and 13 d (study C).

These analyses show that the ATS to detect a change in EDE was longer with higher conception ratios. Herds with high conception ratios have fewer open cows and thus fewer estrous days in a period. Consequently, the power of the SPC charts is lower and therefore it takes longer to signal a change in EDE.

The choice of a fraction equal to 0.4 seemed to be satisfactory for varying conception ratios. A fraction equal to $(1 - \text{conception ratio})$ led to longer in control ATS than the target, which is preferred to a short in control ATS. The observed in control ATS was in many cases significantly shorter or longer than the target ATS, depending on SPC chart type, design, herd size,

Figure 6. Effect of change in estrous detection ratio of most interest (Δ EDR) on observed in control average time to signal (ATS). Δ EDR is expressed in units standard deviation (SD), because SD gets smaller with longer periods. Target ATS is 730 d. ● binomial cusum chart, 100 cows, ■ binomial cusum chart, 1000 cows, ○ normal cusum chart, 100 cows, □ normal cusum chart, 1000 cows.

and estimation of estrous days. This results from the fact that EDR did not exactly follow a normal or binomial distribution. Hawkins and Olwell (1998) noted that cusum chart performance is less sensitive to deviations from the assumed normal distribution when reference values K are small and the target ATS is long. Figure 6 shows that a small reference value does not necessarily result in the best in control performance. Also, period length has a significant effect on the observed in control ATS. Similar results were obtained with a target ATS of 365 or 1095 d, indicating that the sensitivity does not necessarily reduce with a longer target ATS.

Shewhart charts are often said to be less sensitive to the assumed distribution than cusum charts, but the results in Tables 2, 4, and 5 showed also occasional significant deviations from the target ATS. Results of this study showed that the in control performance of both Shewhart and cusum charts is rather dependent on the fit of the assumed probability distribution and SPC chart design.

The choice of the length of each period affected the out of control performance of SPC charts considerably, especially for Shewhart charts with negative control limits due to small subgroup sizes. The results further indicated that the length of the period hardly affected out of control cusum chart performance, which is in agreement with for example Hawkins and Olwell (1998) and Reynolds and Stoumbos (2000).

The number of periods and period length are dependent when all observations are used in control charting.

It was chosen to maintain a constant Type I error rate per unit of time, and not per period, when the period length was varied. This implies that the probability of a Type I error per period increased with an increase in the period length. The Type I error rate may become too high with longer periods, casting doubt if the signal warrants an investigation. Alternatively, the choice of a constant Type I error probability per period will lead to a decrease in Type I errors over the same amount of time when periods are longer.

Rational sampling considerations should take precedence over power considerations for Shewhart chart design (Wheeler, 1995). A rational subgroup is a sample of observations that is generated under conditions in which only random effects cause the variation between the observations (Nelson, 1988). Out of control situations are assumed to occur primarily between rational subgroups and less within. For example, rational sampling might occur by grouping of workers responsible for estrous detection at the same work shift. Rational sampling considerations do not play a useful role for cusum charts, because the persistent change is assumed to continue until the cause is corrected (Hawkins and Olwell, 1998).

To determine the performance of SPC charts, it has traditionally been assumed that the change in the variable of interest occurs between periods, but not within. An additional simplifying assumption made in this study was that the cusum equals 0 at the start of the change in EDE on day 2841. Reynolds and Stoumbos (2000) investigated the ATS of binomial SPC charts for

the more realistic situation that a change can occur anytime, which means between and within periods. They also assumed that the cusum is in steady state at the start of the performance measure. These more realistic assumptions would lead to more complex calculations, while the ATS would differ only a little from the ATS under the simpler assumptions. The more complex calculations may be useful, however, when the ATS approaches the period length, as was the case in some of the results in this study.

The reference value of cusum charts should be chosen to quickly detect the change in EDR of most interest (μ_1). Some of the changes investigated (Table 4) would not be of interest in practice, however. For example, with period lengths of 7 d, the standard deviation of EDR for the 1000-cow herd was 0.0043 while the mean was 0.0284. Thus for $F = 0.125$, $\mu_1 = 0.125 \cdot 0.0284 = 0.0036$. This is a change of interest of $(0.0284 - 0.0036)/0.0043 = 5.8$ standard deviations from the in control mean. Similar calculations result in a change of most interest of 1.8 standard deviations for the 100-cow herd when $F = 0.125$. The change of most interest in cusum charts is typically in the order of two or less standard deviations.

The choice of the reference value mattered most when changes in EDE were small. Non-optimal reference values led to a longer ATS. On the other hand, small changes are likely to result in less economic loss per unit of time so that a longer ATS may not be problematic. More study on the economic loss because the process is out of control seems justified.

Traditionally, SPC charts are designed based on statistical considerations with no reference to the cost of decision errors. Ideally, control limits should be based on the balance of Type I and Type II error costs. Such costs are difficult to estimate. Also, the Type I error rate may become too large for the decision maker to maintain confidence in the SPC chart. Montgomery (1997) suggested using cost estimates in the design of SPC charts, but with statistical constraints on Type I and II errors.

Additional "runs tests" have been developed for the Shewhart charts to provide more sensitivity for small changes. These tests result in significantly higher Type I error rates (Walker et al., 1991). Montgomery (1997) therefore suggested that the SPC charts and tests discussed in this study provide a good set of charts in most situations.

Application

Shewhart charts are well known and provide a good start for monitoring EDR. This study showed that cusum charts deserve consideration, because they may

signal smaller changes in EDE faster. It has been suggested that Shewhart and cusum charts should be used together (Woodall, 2000).

Shewhart and cusum charts for EDR can be based on both the normal and binomial distribution. A normal cusum chart is slightly more robust to the choice of the fraction of possibly pregnant cow days and might therefore be preferred to the binomial cusum chart. A period length of 7 d might be chosen for the cusum charts to make sure that the observations are sufficiently normally distributed. The cusum reference value should be chosen for a small change in EDR that is nonetheless economically significant.

Shewhart X-charts may be preferred over P-charts with probability limits because their design is easier and they are more robust to the choice of the fraction estrous days from possible pregnant cows. Rational sampling considerations should dictate the design of the SPC charts. Otherwise, the choice of the period length should allow for lower control limits that are not negative, but the period length for the Shewhart charts should be small enough to detect the larger changes fast when a cusum chart is also used. Period lengths of 30 d for a 100-cow herd and 7 d for a 1000-cow herd for monitoring EDR are reasonable choices to obtain a short out of control ATS.

Control limits should be based on a balance of Type I and Type II error cost. Because such cost are difficult to estimate, a long in control ATS is advised. Such conservative control limits will prevent disappointment with the use of SPC charts due to too many Type I errors.

CONCLUSIONS

Statistical process control charts allow for a timely signaling of changes in EDR, but their performance is dependent on their design. Shewhart and cusum charts for EDR can be based on both the normal and binomial distribution, provided that the estrous days from possibly pregnant cows are selected well. A normal cusum chart is slightly more robust to the choice of the fraction of possibly pregnant cow days and might therefore be preferred to the binomial cusum chart. The observed in control ATS may be significantly different from the target ATS. With a target ATS of 730 d, the observed ATS for a change from 0.65 to 0.35 EDE was for many designs around 50 d for the 100-cow herd and 11 d for the 1000-cow herd. In general, the observed ATS is longer for smaller changes in EDE and with a longer target ATS. Cusum charts signal small changes faster than Shewhart charts, provided that their reference values are near optimal. The choice of the period length

deserves attention in the design of both Shewhart and cusum charts, but matters more for Shewhart charts.

REFERENCES

- Britt, J. H. 1995. The relationship between postpartum estrous cycles, estrous cycle length, and early embryonic death. *Cattle Practice* 3:85–88.
- de Vries, A. 2001. Statistical process control charts applied to dairy herd reproduction. Ph.D. Diss., Univ. of Minnesota, St. Paul. UMI Digital Dissertations AAT 3002792.
- Fetrow, J., D. McClary, R. Harman, K. Butcher, L. Weaver, E. Studer, J. Ehrlich, W. Etherington, W. Guterbock, D. Klingborg, J. Reneau, and N. Williamson. 1990. Calculating selected reproductive indices: recommendations of the American Association of Bovine Practitioners. *J. Dairy Sci.* 73:78–90.
- Hawkins, D. M., and D. H. Olwell. 1998. Cumulative Sum Charts and Charting for Quality Improvement. Springer-Verlag, New York, NY.
- Keats, J. B., F. P. Lawrence, and F. K. Wang. 1997. Weibull maximum likelihood parameter estimates with censored data. *J. Quality Technol.* 29:105–110.
- Montgomery, D. C. 1997. Introduction to Statistical Quality Control. 3rd ed., John Wiley and Sons, New York, NY.
- Nelson, L. S. 1988. Control charts: rational subgroups and effective applications. *J. Quality Technol.* 20:73–75.
- Quesenberry, C. P. 1997. SPC Methods for Quality Improvement. John Wiley and Sons, New York, NY.
- Reneau, J. K., and M. L. Kinsel. 2001. Record systems and herd monitoring in production-oriented health management programs in food producing animals. In: *Herd Health: Food Animal Production Medicine*. 3rd ed., W. D. Saunders Company, Philadelphia, PA.
- Reynolds, M. R., Jr., and Z. G. Stoumbos. 2000. A general approach to modeling cusum charts for a proportion. *IIE Trans. Quality Reliability Eng.* 32:515–535.
- Roes, K. C. B., R. J. M. M. Does, and Y. Schurink. 1993. Shewhart-type control charts for individual observations. *J. Quality Technol.* 25:188–198.
- Sørensen, J. T., and C. Enevoldsen. 1992. Modelling the dynamics of the health-production complex in livestock herds: a review. *Prev. Vet. Med.* 13:287–297.
- Spahr, S. L. 1993. New technologies and decision making in high producing herds. *J. Dairy Sci.* 76:3269–3277.
- Walker, E., J. W. Philpot, and J. Clement. 1991. False signal rates for the Shewhart control chart with supplementary runs tests. *J. Quality Technol.* 23:247–252.
- Wheeler, D. J. 1995. Advanced Topics in Statistical Process Control—the Power of Shewhart's Charts. SPC Press, Knoxville, TN.
- Woodall, W. H., and D. C. Montgomery. 1999. Research issues and ideas in statistical process control. *J. Quality Technol.* 31:376–386.
- Woodall, W. H. 2000. Controversies and contradictions in statistical process control. *J. Quality Technol.* 32:341–350.

APPENDIX A: DESIGN OF SHEWHART CHARTS

Formulas to design Shewhart charts for EDR assuming normal and binomial probability distributions are presented in Table A.

Table A. Formulas for Shewhart charts for the normal (X-chart) and binomial (P-chart) distributions for estrous detection ratio¹.

SPC chart	Statistic	Center line	Lower control limit (LCL)	Upper control limit (UCL)
X-chart	p_j	$\bar{p}$	$\bar{p} - z_{1-\alpha/2} \cdot \overline{MR}/d_2$	$\bar{p} + z_{1-\alpha/2} \cdot \overline{MR}/d_2$
P-chart	p_j	$\bar{p}$	$\sum_{x=0}^{X_{LCL}} \binom{n_j}{x} \bar{p}^x \cdot (1 - \bar{p})^{n_j-x} \approx \alpha/2$	$\sum_{x=X_{UCL}}^{n_j} \binom{n_j}{x} \bar{p}^x \cdot (1 - \bar{p})^{n_j-x} \approx \alpha/2$

¹ x_j = number of observed estruses in period j, m = number of periods, n_j = number of estrous days (subgroup size) in period j, estrous detection ratio $p_j = x_j/n_j$.

$$\bar{p} = \text{average estrous detection ratio} = \sum_{j=1}^m x_j / \sum_{j=1}^m n_j.$$

$$\overline{MR} = \text{average moving range} = \sum_{j=2}^m |p_j - p_{j-1}| / (m - 1).$$

$d_2 = 1.128$ (Montgomery, 1997, page 223).

α = probability of an observation outside the control limits in a period.

$z_{1-\alpha/2} = 1 - \alpha/2$ probability point of the standard normal distribution..

P-chart: given α , solve for X_{LCL} and X_{UCL} , then $LCL = X_{LCL}/n_j$ and $UCL = X_{UCL}/n_j$.

APPENDIX B: DESIGN OF CUSUM CHARTS

Cusum charts require the calculation of both an upward (C^+) and a downward (C^-) cumulative sum of observations. The size of the change of most interest (μ_1) needs to be identified when the cusum chart is designed

to calculate the reference value K. The binomial cusum for EDR is calculated as follows:

$$C_j^+ = \max(0, C_{j-1}^+ + x_j - n_j \cdot K) \quad \text{for a change } \mu_1 > \mu_0$$

$$C_j^- = \min(0, C_{j-1}^- + x_j - n_j \cdot K) \quad \text{for a change } \mu_1 < \mu_0$$

where x_j is the number of observed estruses in period j , n_j is the number of estrous days, and the reference value $K = -\ln([1 - \mu_1]/[1 - \mu_0])/\ln([\mu_1 \cdot (1 - \mu_0)]/[\mu_0 \cdot (1 - \mu_1)])$ where μ_0 is the average in control EDR. K is approximately the midpoint between μ_0 and μ_1 . Typically $C_0^+ = C_0^- = 0$.

The cusum of the mean of EDR assuming a normal distribution is calculated as follows:

$$\begin{aligned} C_j^+ &= \max(0, C_{j-1}^+ + p_j - K) && \text{for a change } \mu_1 > \mu_0 \\ C_j^- &= \min(0, C_{j-1}^- + p_j - K) && \text{for a change } \mu_1 < \mu_0 \end{aligned}$$

where typically $C_0^+ = C_0^- = 0$ and $K = (\mu_1 + \mu_0)/2$, which is the exact midpoint between μ_0 and μ_1 . An accessible description of the cusums presented here and many others can be found in Hawkins and Olwell (1998).

APPENDIX C: FINDING CONTROL LIMITS FOR CUSUM CHARTS

Control limits for cusum charts can be obtained by crude simulation. This may provide for situations where other methods or software are not available. The algorithm for approximating the lower control limit for the binomial cusum is as follows:

1. Define n as the sample size (average number of estrous days in a period), μ_0 as the in control probability of success (average EDR). Determine the target ATS.
2. Set arrays $C[i]$ and $TS[i]$ equal to 0 for all i with $i = 1$ to N where N the number of replicates. $C[i]$ is the array of cusum values, $TS[i]$ is array of the days to signal. Set control limit H to 0.
3. Set $H = H + \text{increment } H$.
4. For $i = 1$ to N , repeat:
 - Obtain a random deviate $X \sim \text{BIN}(n, \mu_0)$.
 - Calculate $C[i] = \min(0, C[i] + X - n \cdot K)$, with K the binomial reference value.
 - Set $TS[i] = TS[i] + 1$.
 - Until $C[i] \geq H$.
5. Calculate $ATS = (TS[1] + \dots + TS[N])/N$. If $ATS < \text{target ATS}$ then go to step 3.

This algorithm calculates acceptable binomial control limits for $N = 1000$ and an increment in H of 0.01. The accuracy of the approximation can be improved with more replicates and smaller increments. Random deviates from the binomial distribution for this study were generated using the function `IGNBIN` of the `RANLIB` library (Available from <http://odin.mdacc.tmc.edu>, Department of Biomathematics, University of Texas). The upward control limit is easily found by using $C[i] = \max(0, C[i] + X - n \cdot K)$. This algorithm will also find control limits for other distributions by replacing $\text{BIN}(n, \mu_0)$ with the desired distribution, the appropriate cusum function, and reference value K .