

## Methodology for Estimation of Days Dry Effects

M. T. Kuhn and J. L. Hutchison

Animal Improvement Programs Laboratory, Agricultural Research Service, USDA,  
Beltsville, MD 20705-2350

### ABSTRACT

The primary objective of this research was to determine if, with appropriate methodology, unbiased estimates of days dry (DD) effects on subsequent lactation milk yield can be obtained from field data, particularly when DD is correlated with cow effects. Another objective was to ascertain relevant sampling properties of designed trials for estimation of DD effects. Simulated records were used to assess methodology. Along with a model with no adjustments for cow effects, alternative models including 1) previous lactation milk yield, 2) a prior adjustment for cow effects estimated from an animal model, and 3) a combination of 1 and 2, were tested.

Estimates from the unadjusted model were biased downward; however, the 3 alternative analyses provided estimates of DD effects that were essentially unbiased, with a prior adjustment for cow effects and previous milk yield in the model providing the best results in terms of elimination of bias. Therefore, DD effects can be estimated from field data without bias from cow effects.

A designed trial with 2 groups and 10 or fewer cows/group is noninformative and has an unacceptably high probability of leading to invalid conclusions. A minimum of 30 cows/group is considerably better and should be used whenever possible. Even with 30 cows/group, however, the power is low unless the difference between DD groups for yield is at least 1130 kg. Prior correction of 305-d, mature equivalent records for cow effects, using predicted producing abilities, could be done in designed trials to improve the statistical power of tests and accuracy of estimates.

**(Key words:** days dry, methodology, simulation)

**Abbreviation key:** DD = days dry, PE = permanent environmental, P<sub>prevM</sub> = previous lactation milk yield.

### INTRODUCTION

Optimum length of dry period has recently become a topic of considerable interest as evidenced by numerous popular press articles (Annen et al., 2003; Rastani and Grummer, 2003; Bachman, 2004; Grummer, 2004). Although considerable research has supported the conclusion that 60 days dry (DD) maximizes production in the subsequent lactation (Grummer and Rastani, 2004), renewed interest in this topic has stimulated new research. Re-exploring the topic of DD is probably justified because much of the currently available research is at least 20 yr old and cows and management practices have changed considerably over the last 20 yr. Mean breeding value for lactational milk yield in US Holsteins, for example, has increased 2050 kg since 1982 (Animal Improvement Programs Laboratory, 2004). Higher potential for milk yield may mean that cows are able to sustain production for a longer period of time, which may make shorter dry periods more practical.

Although both designed trials and observational studies have been done in the past, much of the research on DD has been done using field data. However, the validity of using field data to determine DD effects has been recently questioned (Bachman and Schairer, 2003; Grummer and Rastani, 2004). Bachman and Schairer (2003) seemed to imply that only designed trials could be used to assess the impact of dry period length on performance. Their final conclusion, for example, was: "Importantly, additional animal trials that assign cows randomly to the dry period lengths that are being evaluated are needed to determine the optimal dry period length for the modern dairy cow in various management scenarios." The idea that studies based on field data are "biased" and basically noninformative has been promulgated in the popular press literature as well (e.g., Linderoth, 2003; Rastani and Grummer, 2003; Mohr, 2004). This criticism is unfortunate because it could potentially (and erroneously) erode producer confidence in results based on their data.

Certainly, some difficulties can be associated with the use of field data, but there are problems with designed trials as well and, more importantly, there are pronounced advantages of observational studies over

---

Received October 20, 2004.

Accepted December 17, 2004.

Corresponding author: Melvin Kuhn; e-mail: mkuhn@aipl.arsusda.gov.

designed trials. One primary advantage of observational studies, (beyond sample size) is that they have the potential to provide much more information than can be gleaned even from numerous designed trials. For example, one can feasibly determine the effects of DD on herd life or lifetime production in observational studies and examine interactions with other factors such as parity and previous lactation variables such as milk yield (**PrevM**), days open, or SCS. This more complete information is crucial in making appropriate recommendations for dry period length. However, few if any designed trials have done so, almost certainly because the expense, time, and effort involved would be prohibitive. The contention that field data cannot be used to determine DD effects should be examined so that this unparalleled source of information could be used to assist in the development of appropriate management recommendations and to lend guidance to further research.

The main argument put forth by Bachman and Schairer (2003), which was repeated several times in different ways throughout the paper, was that, in practice, higher producing cows tend to get shortened dry periods more frequently than do lower producing cows, which means that estimates of DD effects based on field data, at least for production, are biased by cow effects. The primary objective of this research was to determine if by using appropriate methodology, unbiased estimates of DD effects could be obtained from field data, even when higher producing cows receive shorter dry periods more frequently than do lower producing cows. Another objective was to ascertain relevant sampling properties of designed trials for estimation of DD effects.

## MATERIALS AND METHODS

### General Approach

The question of potential bias caused by cow effects is a mathematical question, not a biological question. Such questions can be answered effectively and unequivocally using simulation. The general approach is straightforward. Records are simulated so that the true DD effects are known, the aspect of interest is introduced into the data, estimates are computed using stated methodology, and then the estimated effects are compared with the true values to see whether they agree. This was the overall approach used in this study and is illustrated in Figure 1.

### Simulation of Observational Data

In practice, as well as in this simulation, DD is related to both **PrevM** and milk yield in the subsequent lacta-

tion, but in different ways. Milk yield in the lactation before the dry period partly determines the length of the dry period, but then the dry period length, in turn, has an effect on the subsequent lactation yield. In the following discussion, the generalities for simulation of records will be described first, and then the manner in which DD were simulated will be discussed. The DD effects added to simulated records will be described in more detail after simulation of DD is discussed.

**Model for simulation of records.** To address the simple question of whether partial confounding of cow effects with DD biases estimates of DD effects, the simulation need not be complicated. It is adequate to simulate only the basic aspects of the data, relevant to the question of interest. First-lactation milk records ( $y$ ) were simulated according to the equation:

$$y = \mu_M + A + PE + Parity + e \quad [1.1]$$

and second and later lactations according to the equation:

$$y = \mu_M + A + PE + Parity + DDeffect + e, \quad [1.2]$$

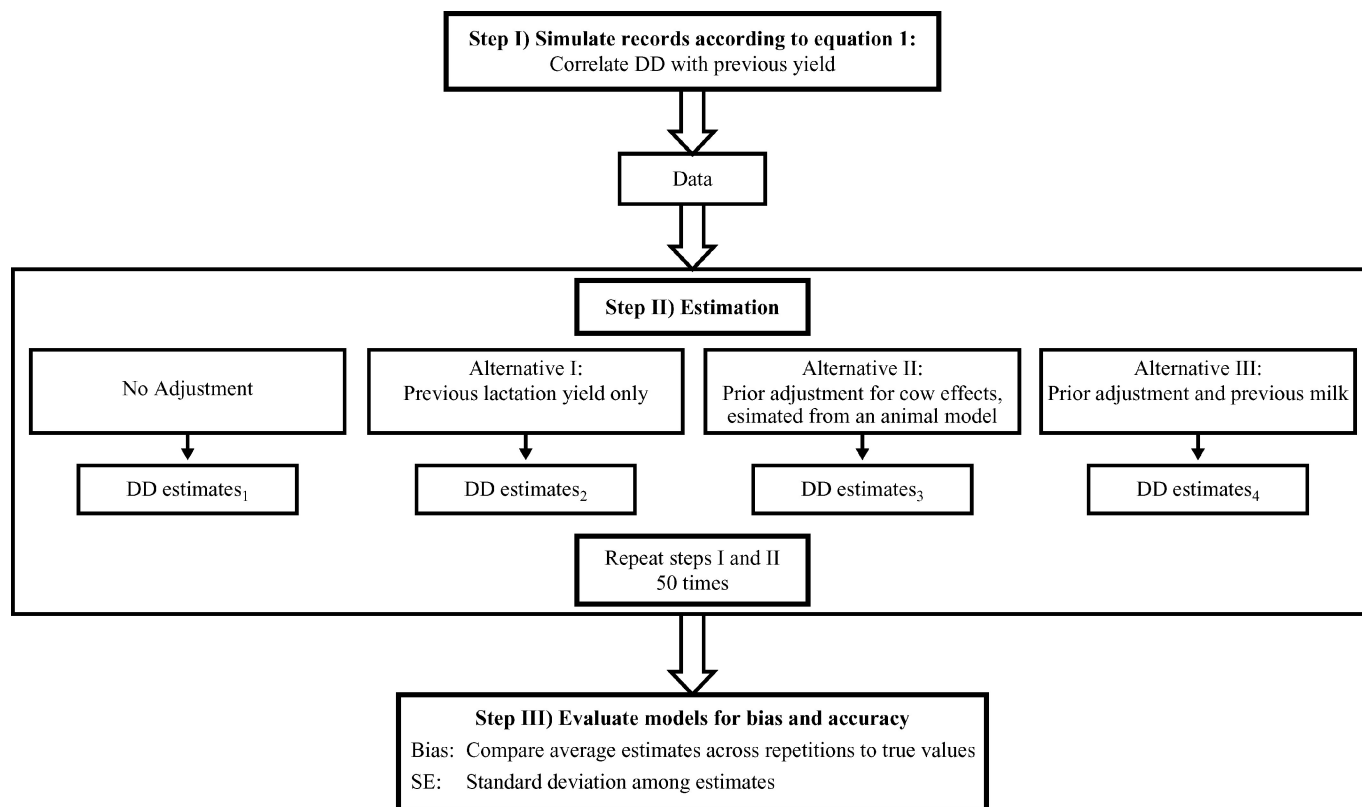
where  $\mu_M$  was an overall mean, A was an additive genetic effect, PE a permanent environmental (**PE**) effect, Parity was a parity 1 vs. second and later effect, DDeffect was the effect of DD, and e was a temporary environmental (or error) effect. Parameters used for simulation are summarized in Table 1. Breeding values for base generation animals were generated as normal random variables with mean zero. Thereafter, breeding values were generated as parent average breeding value plus a normally distributed Mendelian sampling effect. Permanent environmental effects and errors were also generated as normal random variables with mean zero. Standard deviations for A, PE, and e for simulated milk records (Table 1) were comparable to those of Kuhn et al. (2005) for kilograms of lactation milk yield.

### Simulation of DD and its Correlation with **PrevM**

Dry period length was correlated with **PrevM** by first generating a continuous DD according to the equation:

$$DD = \mu_{DD} + \beta_M \times (\text{PrevM} - \mu_{Mi}) + e_{DD}, \quad [2]$$

where  $\mu_{DD}$  was an overall mean DD,  $\beta_M$  was a linear regression of DD on the cow's milk yield deviation from the population mean in the  $i$ th yr ( $\mu_{Mi}$ ), and  $e_{DD}$  was a random normal error, independent of milk yield. The overall standard deviation of 15 d for simulated DD was similar to that of Kuhn et al. (2005). The -0.2



**Figure 1.** General approach for determining bias in estimates of days dry (DD) effects using field data.

correlation between simulated DD and PrevM used in this simulation (Table 1) was higher than the  $-0.15$  correlation reported by Kuhn et al. (2005) to ensure an adequate test of methodology. If the stronger relationship can be handled, then certainly the weaker one can be as well.

Once DD was calculated from equation [2], it was then categorized as indicated in Table 2, and the corres-

ponding DD effect, also listed in Table 2, was added to the subsequent lactation milk record in the next year, as indicated in equation [1.2]. The objective of this part of the research was simply to determine if DD effects can be estimated without bias when DD is correlated with PrevM (i.e., cow effects). Thus, the exact magnitude used for the DD effects (Table 2) is largely irrelevant. Nonetheless, the effects used in this simulation

**Table 1.** Summary of population parameters and characteristics for simulation of observational data.

Parameter <sup>1</sup>	Value	Population characteristic <sup>2</sup>	Value(s)
Mean milk ( $\mu_M$ )	9100	Number of cows/yr	100,000
Standard deviation of breeding values	750	Number of cows replaced each year	40,000
Standard deviation of permanent environmental effects	750	Number of yr (including base)	11
Standard deviation of errors for milk yield	1100	Total number of cows	500,000
Parity 1 effect	-500	Total number of records	1,100,000
Parity 2 effect	500	Total number of second lactation records	308,000
Mean DD ( $\mu_{DD}$ )	60	Number of sires/yr	1000
Regression of DD on milk ( $\beta_M$ )	-0.002	Number of sires replaced each year	200
Correlation of DD and previous lactation milk yield	-0.2	Total number of sires	2800
Standard deviation of errors for DD ( $e_{DD}$ )	14.7	Culling rates by parity	0.3, 0.4, 0.4, 0.5, 1
Standard deviation of DD	15	Culling rates by milk yield group <sup>3</sup>	0.5, 0.3, 0.2
Number of replicates	50		

<sup>1</sup>DD = Days dry.

<sup>2</sup>For each replicate.

<sup>3</sup>Low, medium, and high groups, respectively; groups were of equal size.

**Table 2.** Days dry (DD) categories, effects, and expected frequencies for simulation of observational data.

DD	DD category	DD effect	Expected frequency
<30	1	-2300	11,375
30 to 37	2	-2000	19,923
38 to 44	3	-1500	40,232
45 to 52	4	-1000	76,920
53 to 60	5	-500	101,550
61 to 67	6	0	89,815
68 to 75	7	0	80,857
76 to 82	8	0	43,711
83 to 90	9	0	24,242
>90	10	0	11,375
Total			500,000

(Table 2) are at least as large or larger than literature estimates (e.g., Funk et al., 1987; Makuza and McDaniel, 1996), which ensures at least an adequate test of methodology. Larger effects provide a more stringent test of methodology. The simulated DD effects (Table 2) for the last 4 categories were set to zero to simulate no effect of DD after the minimum dry period to maximize production is met.

**Structure of simulated populations.** Population characteristics are summarized in Table 1. One hundred thousand cows were generated in the base year. The simulation was run for 10 yr thereafter, with the 100,000-cow population being maintained each year. A replacement rate of 40,000 cows/yr resulted in 500,000 total cows [ $10 \times (40,000) + 100,000$  base cows] for each replicate. Lactation percentages for base cows were 40, 28, 17, 10, and 5 for lactations 1 through 5, respectively. Culling rates (Table 1) differed by parity and depended on level of milk yield. Dams of female replacements were selected at random.

There were 1000 sires available each year with replacement of 200 sires/yr, which resulted in 2800 total sires. For the sake of introducing some selection into the simulation, the 200 bulls culled each year were randomly selected from the bottom half of bulls based on a simulated value computed as the sum of true breeding value plus a normally distributed random error with mean zero and standard deviation 335. Sires and dams of new sires were randomly selected from the top half of their respective populations and this selection was based on a similar simulated value. Each sire and dam was allowed to produce only one son/yr.

### Estimation of DD Effects for Observational Data

**DD as a covariate vs. categorical variable.** In all analyses tested in this study, DD effects were estimated using a categorical variable for DD rather than regression. This approach has been common practice in the

past with respect to DD (e.g., Keown and Everett, 1986; Funk et al., 1987; Makuza and McDaniel, 1996) and may be justified for a number of reasons. Most research to date (Funk et al., 1987; Makuza and McDaniel, 1996) indicates that production increases with increasing DD up to a point, but after that point, additional DD add nothing to subsequent lactation yield. This relationship is somewhat difficult to model precisely with regression. The relationship is certainly not linear. A quadratic regression may be a sufficient approximation but, in principle, a quadratic is necessarily parabolic and thus would tend to "predict" a decrease in production after the maximum point. A cubic regression would allow for the flattening off after the maximum point but would predict an increase in production with long DD. Fitting DD as a covariate necessarily makes prior assumptions about the nature of the relationship between DD and milk yield and the purpose of the research should be to discover the nature of the relationship. Fitting DD as a covariate may conceal variation that would be uncovered when DD is fit as a categorical variable.

Concerning potential bias in estimates of DD effects using field data, it is clear that what needs to be corrected for are the cow effects (breeding value and PE effects). Thus, in addition to an analysis with no adjustments, 3 alternative analyses, aimed at correcting for cow effects, were tested.

**Alternative I: PrevM.** The first, and simplest, alternative was to include PrevM as a linear covariate in the model for estimation of DD effects. The model equation for analysis under this alternative, was:

$$y = \mu + \beta \times \text{PrevM} + \text{DDcatg} + e, \quad [3.1]$$

where  $y$  was a milk record and DDcatg was a categorical variable for DD as defined in Table 2. Only second and later lactations contributed to this analysis. This alternative is similar to the approach of Makuza and McDaniel (1996). Previous lactation yield has also been used in some designed trials to correct for differences in cow effects between groups (e.g., Gulay et al., 2003).

**Alternative II: Prior correction for cow effects, estimated from an animal model.** Although the properties of animal models fit using BLUP are well established (Henderson, 1984), their use in DD research presents a unique aspect. Preferably, DD would be kept in the model for estimation of cow effects to ensure that the DD effect is not partially partitioned into the animal or PE effect. However, exclusion of first-lactation records can lead to bias in predicted breeding values (Pollak and Quaas, 1981; Henderson, 1982). The dilemma, then, is that first-lactation records cannot be excluded but they also have no DD associated with them. This difficulty was circumvented by assigning all first-lacta-



tion records their own unique (dummy) DD category, category 11. This method was similar to the approach of Funk et al. (1987), although they did not use genetic relationships for estimation of cow effects. This approach to estimating the cow effects was the reason for inclusion of a parity effect in the simulation. Parity 1 and DD category 11 were completely confounded and this aspect should be tested because it would occur when working with field data.

In practice, it may be desirable to examine DD effects on subsets of data, for example, by parity. Cow effects, however, need to be estimated using all data simultaneously to maximize accuracy and to avoid biases. Both of these goals can be met using a 3-step procedure for estimation of DD effects: 1) estimate cow effects from an animal model using all data, 2) preadjust the records by subtracting the estimated cow effects, and 3) estimate DD effects from the preadjusted records. This was the procedure tested in this alternative. Cow effects were estimated from the model:

$$y = \mu + A + PE + Parity + DDcatg + e \quad [3.2a]$$

where DDcatg was a categorical variable as defined in Table 2, plus a category 11 for all first-lactation records. The effects A, PE, and e were fit as random effects with variance-covariance matrices:  $\text{Var}(A) = \mathbf{A}\sigma_A^2$ ,  $\text{Var}(PE) = I\sigma_{PE}^2$ , and  $\text{Var}(e) = I\sigma_e^2$ , where  $\mathbf{A}$  was a relationship matrix. Days dry effects were estimated from the model:

$$\text{Adjusted-}y = \mu + DDcatg + e \quad [3.2b]$$

where Adjusted- $y$  is the record with solutions for breeding value and PE from [3.2a] subtracted, and DDcatg is as defined in Table 2. First-lactation records were not included for analysis in [3.2b].

**Alternative III: Combination of I and II.** A third alternative was to preadjust for cow effects as in alternative II but to include PrevM in the model for analysis, i.e., in equation [3.2b]. Thus, alternative III was the same as II except the model for estimation of DD effects was:

$$\text{Adjusted-}y = \mu + \beta \times \text{PrevM} + DDcatg + e. \quad [3.3]$$

### Evaluation of Models/Alternatives for Analysis of Simulated Observational Data

As indicated in Figure 1, 50 replicates were run. In each replicate, DD effects were estimated using each of the 4 alternatives (no adjustment, plus alternatives I to III). The primary objective was to determine whether a given method was biased. To evaluate bias for a given alternative, estimates across the 50 repli-

cates were averaged for that alternative and simply compared with the true DD effects to see if the estimates differed from the true values.

The amount of variation (standard deviation) among replicate estimates was also calculated for each method. This reflects the precision or standard error (i.e., the standard deviation of the sampling distribution) of each approach.

### Simulation of Designed Experiments

The second part of this research was to determine the sampling properties of designed trials for estimation of DD effects. Twenty, 40, and 60 cows were generated and divided randomly and evenly between 2 DD groups. Five levels of DD effect (difference between the 2 groups) were used: 225, 450, 680, 900, and 1130. Thus, there were 3 (sample sizes)  $\times$  5 (DD effects) = 15 independent experiments, each with 1000 independent replicates.

Two records were simulated for each cow, according to the equations:

$$y_1 = \mu_M + \text{Cow} + e$$

$$y_2 = \mu_M + \text{Cow} + DD\_effect + e$$

where DD\_effect was the DD effect for the given experiment and Cow was a normally distributed cow effect with mean zero and variance equal to the sum of the variances for breeding value and PE used in the simulation of observational data. Error variance was also the same as that in Table 1. Relationships were not generated in this simulation because cows were assigned randomly to DD groups and because an animal model was not used for estimation. Two records were generated on each cow to be able to determine the usefulness of including "previous lactation" yield in the model for analysis. Two models were used for analysis in each replicate:

$$y_2 = \mu + DD + e_1$$

$$y_2 = \mu + DD + \beta * y_1 + e_2.$$

Summary statistics for each of the 15 experiments included percentage of replicates with significant differences at the level of 0.10, mean  $P$  values, the standard deviation among replicate estimates, and a description of the distribution of estimates. The difference between groups in absolute cow effects was also summarized.

**Table 3.** Mean estimates and bias across 50 replicates when cow effects were correlated with days dry (DD).

DD category	True effect	Mean estimates across replicates <sup>1,2</sup>				Bias <sup>1,2</sup>			
		No adjustment	I	II	III	No adjustment	I	II	III
1	-2300	-1945	-2286	-2282	-2294	355	14	18	6
2	-2000	-1734	-1991	-1985	-1994	266	9	15	6
3	-1500	-1299	-1493	-1490	-1497	201	7	10	3
4	-1000	-864	-995	-993	-997	136	5	7	3
5	-500	-433	-498	-496	-499	67	2	4	1
6	0	0	0	0	0	0	0	0	0
7	0	-67	-2	-4	-1	-67	-2	-4	-1
8	0	-136	-4	-8	-3	-136	-4	-8	-3
9	0	-201	-2	-10	-2	-201	-2	-10	-2
10	0	-297	-4	-14	-4	-297	-4	-14	-4

<sup>1</sup>I = Alternative I for estimation, previous milk only included in model for estimation; II = alternative II for estimation, prior correction for cow effects estimated from an animal model; III = alternative III for estimation, prior correction for cow effects and previous milk included in model for estimation.

<sup>2</sup>Estimates expressed relative to category 6.

## RESULTS

### Simulation of Observational Studies: Cow Effects Correlated with DD

Mean of estimates across replicates as well as bias (mean estimated effect – true effect) are in Table 3. Estimates were expressed relative to category 6; i.e., as DD category<sub>i</sub> – DD category<sub>6</sub>. Thus, the category 6 estimate, as well as its standard deviation, is necessarily zero. Estimates are clearly biased downward with no adjustment. As expected, when there is no adjustment for cow effects, biases are positive for the shorter dry periods simply because the cow effects associated with shorter dry periods are positive and the converse is true for longer dry periods for the same reason: poorer cows received longer dry periods.

The estimates in Table 3 clearly show that the magnitude of the negative impact of shortened dry period is underestimated when there is no correction for cow effects. However, all alternative analyses provided estimates of DD effects that were essentially unbiased. Use of a prior adjustment for cow effects, along with inclusion of PrevM in the model for estimation of DD effects (alternative III), provided the best results in terms of elimination of bias. However, it was only slightly better than using PrevM alone. This result is probably because with a large number of records, the regression on PrevM is well estimated. The main point of Table 3 is simply that DD effects can be estimated from field data without bias from cow effects.

Standard deviations among estimates (Table 4) were small for all analyses and there was surprisingly little difference among them. Use of PrevM was as good as adjustment for cow effects from an animal model. This, too, probably reflects large sample size. If sample size

for estimation of DD effects was small but sample size for estimation of cow effects was large (e.g., if cow effects in a designed trial were corrected for by using animal model estimates from the national database), then prior correction for cow effects would likely show an advantage in standard error over the use of PrevM alone. An analysis using only second-lactation records (number of records per replicate = 308,000) was done and results (not shown) were similar to those in Tables 3 and 4.

### Simulation of Designed Trials

In contrast with the simulation of observational studies, the concern with designed studies is not “bias” per se because cows are (presumably) assigned at random to DD groups. Rather, due to sample sizes that are generally small, the concern with designed experiments

**Table 4.** Standard deviations among estimates when cow effects were correlated with days dry (DD).

DD Category	No adjustment	Alternative <sup>1</sup>		
		I	II	III
1	14	12	13	13
2	10	9	9	9
3	8	7	6	6
4	6	6	6	5
5	6	6	5	5
6	0	0	0	0
7	7	6	6	6
8	7	6	6	6
9	10	9	9	9
10	19	16	15	15

<sup>1</sup>I = Alternative I for estimation, previous milk only included in model for estimation; II = alternative II for estimation, prior correction for cow effects estimated from an animal model; III = alternative III for estimation, prior correction for cow effects and previous milk included in model for estimation.

**Table 5.** Percentage of replicates with significant differences ( $P \leq 0.10$ ) for designed trials.

Cows/group <sup>1</sup>		Days dry effect				
		225	450	680	900	1130
10:	No PrevM	12	16	24	38	47
	PrevM	12	19	28	41	54
20:	No PrevM	15	25	39	58	72
	PrevM	15	28	49	66	81
30:	No PrevM	18	29	51	69	88
	PrevM	19	35	61	80	93

<sup>1</sup>No PrevM = previous lactation milk yield not included in the model for analysis; PrevM = previous lactation milk yield included in the model for analysis.

is power or the ability to detect real differences given that they exist.

The probabilities of detecting real differences, for different magnitudes of DD effects and the 3 different sample sizes used in this simulation, are given in Table 5. If DD effects truly exist, it would be better to flip a coin to decide whether dry period length affects milk yield than to run an experiment with only 10 cows/group in 2 groups. The 10-cow groups did not detect a DD effect even 50% of the time until the magnitude of the effect was 1130 kg, and even then only with the enhanced analysis using PrevM. With a milk price of \$0.26/kg (\$12.00/cwt), the 10-cow groups would fail to detect an effect on subsequent lactation yield (1130 kg) worth \$294.00/cow per lactation, 46% of the time. The lowest mean  $P$  value (Table 6) for the 10-cow groups was only 0.19. With 20- and 30-cow groups, better results were obtained, as expected, but even with 30 cows/group, DD effects were not detected more than 50% of the time until the magnitude of the DD effect was at least 680 kg. Mean  $P$  values dropped to about 0.20 for both 20- and 30-cow groups when the DD effect reached 680 kg. However, 30-cow groups did not reach a mean

**Table 6.** Mean  $P$  values for a 2-tailed  $t$ -test from the simulation of designed trials.

Cows/group <sup>1</sup>		Days dry effect				
		225	450	680	900	1130
10:	No PrevM	0.47	0.44	0.37	0.30	0.24
	PrevM	0.47	0.42	0.33	0.26	0.19
20:	No PrevM	0.46	0.39	0.26	0.17	0.10
	PrevM	0.45	0.36	0.21	0.13	0.07
30:	No PrevM	0.43	0.33	0.20	0.11	0.04
	PrevM	0.42	0.29	0.15	0.07	0.02

<sup>1</sup>No PrevM = previous lactation milk yield not included in the model for analysis; PrevM = previous lactation milk yield included in the model for analysis.

**Table 7.** Standard deviation among estimates for simulation of designed trials.<sup>1</sup>

Cows/group	No PrevM	PrevM
10	695	620
20	491	433
30	396	351

<sup>1</sup>No PrevM = previous lactation milk yield not included in the model for analysis; PrevM = previous lactation milk yield included in the model for analysis.

$P$  value less than 0.10 until 900 kg, and 20-cow groups not until 1130 kg.

The standard deviations among estimates, across replicates, are in Table 7. Because these do not depend on the magnitude of the DD effect, values presented in Table 7 are averages across the DD effect categories. Even though the regression of milk yield on PrevM may not be well estimated with such small sample sizes, clearly inclusion of PrevM is advantageous and should be done in the analysis of data from designed trials. Some researchers may argue that they have obtained standard errors smaller than those indicated in Table 7. If that is the case, then a truly random sample from the population to which inferences were to be made was most likely not taken because the standard deviations for animal, PE, and error used in this simulation were realistic, at least for the US Holstein population.

The distribution of estimates, for the 680-kg DD effect, is shown in Table 8. For 10-cow groups, estimates in the wrong direction were obtained 16% of the time. For example, if the true effect of a 30-d dry period were to cause a reduction (relative to 60 DD) of 680 kg in the subsequent lactation, an experiment with only 10 cows/group would actually find a higher mean for cows with 30 DD 16% of the time. Fortunately, this probability is considerably lower for 20- and, especially, 30-cow groups. For 30-cow groups, there was a 71% chance of

**Table 8.** Distribution of estimates for designed trials with a days dry effect of 680: percentage of estimates in each range.<sup>1</sup>

Estimate range	Cows/group		
	10	20	30
<0	16	6	2
0 to 250	9	10	7
251 to 500	15	19	20
501 to 600	8	10	12
601 to 650	2	4	5
651 to 700	3	5	6
701 to 800	6	8	12
801 to 1000	9	15	19
1001 to 1200	10	12	10
1201 to 1500	11	8	6
>1500	11	3	1

<sup>1</sup>With previous lactation milk yield in the model for analysis.

**Table 9.** Differences between days dry (DD) groups in mean cow effects: percentage in each range.

Mean difference between DD groups in absolute cow effects	Cows/group		
	10	20	30
0 to 50	8	13	14
51 to 100	8	12	14
101 to 200	16	21	25
201 to 300	15	17	19
301 to 400	12	13	13
401 to 500	11	10	8
501 to 600	8	6	4
601 to 700	7	4	2
>700	15	4	1

obtaining an estimated difference between groups of at least +500 kg when the true difference was 680 kg. Perhaps most disappointing is the percentage of estimates in the 500 to 800 kg range. Only 35% of the estimates fell in this range, even for the 30-cow groups. The 20-cow groups had 27% of their estimates in this range and the 10-cow groups had only 19% in this range.

Bachman and Schairer (2003) argued that the advantage of designed trials was that cows were randomly assigned to groups and, indeed, truly random assignment of cows to groups is desirable. However, random assignment does not mean equal assignment, especially when group sizes are small. This is illustrated in Table 9. Even with 30 cows/group, the mean difference in cow effects between the 2 groups was less than 50 kg only 14% of the time. With only 10 cows/group, the mean difference in cow effects was  $\geq 300$  kg 53% of the time. Days dry groups differed in mean cow effects by  $\geq 300$  kg 37% of the time for 20-cow groups and 28% of the time for 30-cow groups. Use of mature-equivalent records for estimation of DD effects (in contrast to use of actual yields) may cloud differences in lactation yield caused by DD. Nonetheless, they may still be the best choice of production variable for designed studies because the predicted producing abilities, computed by Animal Improvement Programs Laboratory and based on a large national database, could be used to preadjust the mature-equivalent records. This, in turn, would reduce differences between groups in cow effects and, subsequently, increase power of tests and precision of estimates.

## DISCUSSION

In a recent review, Bachman and Schairer (2003) criticized DD studies based on field data, with the overall conclusion being that only designed trials could be used for determining the effect of DD on subsequent lactation performance. Their main concern (that esti-

mates from field data are biased by cow effects) was addressed in this research. Ironically, advocates of a shortened dry period seem to ignore the fact that such a bias, if it existed, would favor shorter, not longer, DD which is shown in Table 3. The arguments made by Bachman and Schairer (2003) concerning cow effects were not consistent throughout their paper. In criticizing a study by Klein and Woodward (1943) because of lack of correction for year and season effects, they categorize repeatability of milk yield as being "low." Low repeatability, however, implies that cow effects are of minor importance and yet it was argued repeatedly that cow effects are a major source of bias in studies utilizing field data. In spite of the inconsistencies in Bachman and Schairer's (2003) discussion, cows with higher producing abilities do tend to have shorter dry periods more frequently than do cows with lower producing abilities, as was clearly shown in a recent study by Kuhn et al. (2005). However, this research demonstrated that such effects could be readily accounted for with a proper statistical analysis. Because the correlation between DD and cow effects can be readily accounted for in analysis of field data, these cow effects certainly present no reason to abandon the use of field data in helping to determine recommendations for appropriate length of dry period.

Bachman and Schairer (2003) listed 15 factors affecting milk yield and seemed to indicate that these factors were not or could not be accounted for in the analysis of field data. In fact, with the exception of bST usage, virtually all of the factors (and certainly all of the major factors) listed by Bachman and Schairer (2003) can be readily accounted for in the analysis of field data and perhaps better so than what can or has been done in designed trials. Season of calving effects, for example, would be estimated much more accurately with the larger field data sets than with the smaller experimental data sets.

One last point of contention brought up by Bachman and Schairer (2003), concerning the use of field data, was quality of data. One argument was that, in field data, "short dry-period categories are composed primarily of a nonrandom population of cows that freshened earlier than expected for various reasons." Ignoring the fact that they also argued that high-producing cows were the ones with short DD, this concern can be easily dealt with using an edit, as can abortions, which were pointed to as another possible source of corruption in field data.

The topic of "planned vs. unplanned shortened dry period" is rather questionable. First, this argument is largely speculative; it is by no means well documented that most short dry periods occurring in practice are "unplanned." Furthermore, from a certain perspective,



it does not matter why a cow was dried off or how the cow was managed, only how long the dry period was. The “planned vs. unplanned” mantra necessarily implies that 1) shortened dry periods do cause a loss in subsequent lactation milk yield and 2) that there are (special?) management practices that can offset this loss, either partially or completely. The viewpoint taken in making this argument, then, is that the only correct comparison to make is the one where these management practices are used. It would be useful if these management practices were clearly specified. One could then evaluate whether such practices are likely to be implemented in practice and, subsequently, what should be estimated: the effect of DD with or without this management. The “unplanned” dry periods can be readily identified with an edit if one desires to use “planned” dry periods for comparison. Both days open and calving dates are reported to DHI. For a given cow, the reported days open from the previous lactation can be compared with the calving date in the next lactation. If they match (say, within 10 d), that would suggest that the producer knew (or should have known) when the cow was going to calve because it was, in effect, reported to DHI; if they do not match, then the record can be excluded. If a given herd has many records that do not match, the whole herd can be deleted. Furthermore, if the field data indicate that shorter DD are associated with decreases in yield, and management practices do exist that can offset these losses, then the conclusion is either that 1) farmers do not know about these management practices, or 2) they are not feasible for implementation. Such a result, in turn, would suggest that these practices (which putatively offset the losses) need to be better publicized or need to be re-worked because the farmer cannot implement them in a cost-effective way.

In spite of an overstated criticism of the use of field data for determining the effects of DD on performance, Bachman and Schairer (2003) are correct in their overall implication that well-designed experiments, followed up by appropriate statistical analysis, have certain potential advantages over the use of field data. The converse, however, is also very much true. A major limitation of designed experiments will almost invariably be sample size and the ensuing lack of statistical power (ability) to detect real differences, which was clearly demonstrated in Tables 5 through 9 of this paper. Another concern related to small sample sizes is scope of inference. With only a small number of cows, typically in a single herd (and possibly a specialized research herd), located in one part of the country, it would be difficult to make confident inferences for all cows across the United States. To overcome the shortcomings of designed trials related to sample size, the

experiments need to be independently repeated numerous times, in different herds, across different regions of the country.

Another serious limitation of most designed trials for determining the consequences of shortened dry periods is that they usually involve short-term investigation only. The long-term consequences of shortened dry periods need to be investigated before being recommended to producers. Even if shortened DD resulted in no loss in production in the subsequent lactation, there could be other consequences that result in shortened herd life. Designed experiments aimed at addressing this aspect, although feasible, would be costly both financially and in terms of time. Furthermore, it seems likely that the degree of control in a long-term experiment, one of the potential advantages of designed trials, would deteriorate over time.

Bachman and Schairer (2003) bring up a number of good points related to use of field data for examining the effect of DD, although most if not all of these can be handled using appropriate edits and data analysis. Designed experiments and observational studies both have advantages and disadvantages. Furthermore, the weaknesses of one are the strengths of the other. Designed trials have the potential advantage of better control over some extraneous variables and, perhaps most importantly, direct observation of cows involved. They are limited, however, in their ability to detect real differences, in their ability to ensure equal representation or proper correction for cow effects across groups, in the scope of inference that can be made, and in their ability to examine long-term consequences or interactions with other factors, all of which are the strengths of observational studies.

The objective should be to present dairy farmers, as accurately as possible, with all the consequences of varying dry period length. Both types of studies should be done (and done well) to meet that goal. An extreme viewpoint in either direction, that designed studies are too small and therefore noninformative or that observational studies are “biased” and merit little if any consideration, is contrary to accomplishing that goal.

## CONCLUSIONS

The effects of varying dry period lengths can be estimated from field data, free of bias by cow effects. The best alternative for estimation is the use of an animal model combined with inclusion of PrevM as a covariate in the model for estimation. Prior correction of records with estimates of cow effects from an animal model is acceptable. A considerable advantage of estimates from field data, compared with designed trials, is that they

have the potential for much higher precision due to large sample sizes.

A designed experiment with 10 or fewer cows/group is noninformative and has an unacceptably high probability of leading to erroneous conclusions. Although the use of 20 cows/group reduces this probability somewhat, an absolute minimum of 30 cows/group is considerably better and should be used whenever possible. Even with 30 cows/group, however, standard errors are large and, thus, group sizes larger than 30 would be preferable whenever feasible. With 30-cow groups, repeated trials would still be needed to obtain results in which confidence could be placed.

Previous-lactation milk yield should be included as a covariate in the analysis of data from designed experiments. Mature-equivalent production records corrected using the predicted producing abilities computed quarterly by Animal Improvement Programs Laboratory may be the best choice of trait for analysis in designed trials. Correction of records using predicted producing abilities would help eliminate cow effects and increase the accuracy of estimates, which is the greatest weakness of designed trials. If mature-equivalent yields are used, however, scrutiny of the records involved would be necessary because these records are extended to a 305-d basis, even for cows milking less than 305-d. Use of 305-d records could conceal differences in lactation yield if cows in either group failed to milk the full 305-d, as would be the case if short dry periods led to earlier dry off or earlier culling in the subsequent lactation.

#### ACKNOWLEDGMENTS

Ignacy Misztal is gratefully acknowledged for the use of his BLUPF90 program, used to compute animal model solutions.

#### REFERENCES

- Animal Improvement Programs Laboratory. 2004. Trend in milk breeding values for Holstein calculated August, 2004. Online. Available: <http://aipl.arsusda.gov/dynamic/trend/current/trndx.html>. Accessed Sept. 30, 2004.
- Annen, E. L., R. J. Collier, and M. A. McGuire. 2003. Older cows fared well without a dry period. *Hoard's Dairyman* 148:650.
- Bachman, K. C. 2004. Will shorter dry periods pay for you? *Hoard's Dairyman* 149:321.
- Bachman, K. C., and M. L. Schairer. 2003. Invited review: Bovine studies on optimal lengths of dry periods. *J. Dairy Sci.* 86:3027–3037.
- Funk, D. A., A. E. Freeman, and P. J. Berger. 1987. Effects of previous days open, previous days dry, and present days open on lactation yield. *J. Dairy Sci.* 70:2366–2373.
- Grummer, R. 2004. Here's more on short dry periods. *Hoard's Dairyman* 149:142.
- Grummer, R. R., and R. R. Rastani. 2004. Why reevaluate dry period length? *J. Dairy Sci.* 87:E77–E85.
- Gulay, M. S., M. J. Hayen, K. C. Bachman, T. Belloso, M. Liboni, and H. H. Head. 2003. Milk production and feed intake of Holstein cows given short (30-d) or normal (60-d) dry periods. *J. Dairy Sci.* 86:2030–2038.
- Henderson, C. R. 1982. Best linear unbiased prediction in populations that have undergone selection. Page 191 in *Proc. World Congr. Sheep and Beef Cattle Breeding*, Palmerston North, New Zealand.
- Henderson, C. R. 1984. Applications of linear models in animal breeding. University of Guelph, Guelph, ON, Canada.
- Keown, J. F., and R. W. Everett. 1986. Effect of days carried calf, days dry, and weight of first calf heifers on yield. *J. Dairy Sci.* 69:1891–1896.
- Klein, J. W., and T. E. Woodward. 1943. Influence of length of dry period upon the quantity of milk produced in the subsequent lactation. *J. Dairy Sci.* 26:705–713.
- Kuhn, M. T., J. L. Hutchison, and H. D. Norman. 2005. Characterization of days dry in US Holsteins. *J. Dairy Sci.* 88:1147–1155.
- Linderoth, S. 2003. Decrease dry periods by 20 days. *Dairy Herd Management*. Online. Available: <http://www.dairyherd.com>. Accessed Oct. 12, 2004.
- Makuza, S. M., and B. T. McDaniel. 1996. Effects of days dry, previous days open, and current days open on milk yields of cows in Zimbabwe and North Carolina. *J. Dairy Sci.* 79:702–709.
- Mohr, P. 2004. 60 days dry – says who? *Dairy Today*. Online. Available: [http://www.agweb.com/pub\\_get\\_article.asp?sigcat=dairy&pageid=106373](http://www.agweb.com/pub_get_article.asp?sigcat=dairy&pageid=106373). Accessed Oct. 12, 2004.
- Pollak, E. J., and R. L. Quaas. 1981. Monte Carlo study of genetic evaluations using sequentially selected records. *J. Anim. Sci.* 52:257–264.
- Rastani, R., and R. Grummer. 2003. Shorter dry periods look good. *Hoard's Dairyman* 148:599.