Short communication: Test for nonpregnancy in dairy cows based on plasma progesterone concentrations before and after timed artificial insemination

L. J. Wilsdorf, D. H. Keisler, S. E. Poock, W. R. Lamberson, R. C. Escalante, and M. C. Lucy
Division of Animal Sciences, University of Missouri, Columbia 65211

ABSTRACT

Timed artificial insemination (AI) programs have increased reproductive efficiency in dairy herds. A low timed AI pregnancy per AI is partially explained by cows that fail to respond optimally to the series of treatments that are designed to synchronize ovulation for AI. We hypothesized that testing cows for plasma progesterone concentrations during a timed AI protocol could be used as an early diagnostic test for nonpregnancy. Lactating Holstein cows (n = 160) in 2 confinement-style dairies were used. Cows were treated with Presynch Ovsynch 56 for timed AI. Concentrations of progesterone in plasma were measured at −3, 0, 7, and 25 d relative to timed AI. Progesterone data were analyzed and receiver operating characteristic curves were generated by using logistic regression. The area under the receiver operating curves for a progesterone test for nonpregnancy on d −3 (PGF2α), 0 (AI), 7, and 25 d relative to timed AI were 0.68, 0.52, 0.55, and 0.89, respectively. The cutpoints and sensitivity (respectively) for the progesterone test were 0.51 ng/mL (lower = nonpregnant) and 28.2% for the day of PGF2α, 0.43 ng/mL (greater = nonpregnant) and 17.9% for the day of AI, 1.82 ng/mL (lower = nonpregnant) and 23.1% for 7 d after AI, and 2.67 ng/mL (lower = nonpregnant) and 76.0% for 25 d after AI. The false positive rate was less than 5% for all tests. Analysis of a second data set from a published study gave approximately the same cutpoints and sensitivity. When both studies were combined, approximately 20% of nonpregnant cows could be identified with a single test that was done before or shortly after AI with a false positive rate of less than 5%. When 2 and 3 tests were applied sequentially, the sensitivity for identifying nonpregnant cows increased from 38.4 to 50.5%. The pregnancy per AI for those cows that met the established progesterone criteria was approximately 3 to 4 times greater than those that failed to meet the criteria. The conclusions were that cows destined to be nonpregnant after timed AI can be identified before or shortly after AI. Testing for nonpregnancy before or shortly after AI may have utility with respect to eliminating a nonproductive AI (cows identified before AI) or shortening the time to reinsemination (cows identified by 1 wk after AI).

Key words: progesterone, pregnancy, timed AI
Lactating Holstein cows (n = 160) in 2 confinement-style dairies (University of Missouri, Foremost Dairy, Midway, MO, and Heartland Dairy, La Belle, MO) were used. Cows were treated with Presynch Ovsynch 56 (PGF2α, 14 d, PGF2α, 14 d, GnRH, 7 d, PGF2α, 56 h, GnRH, 16 h, timed AI) so that first timed AI was 70 to 76 d postpartum. The PGF2α was Lutalyse (5 mL; 25 mg; Zoetis, Florham Park, NJ). The GnRH was Factrel (gonadorelin hydrochloride; 2 mL; 100 μg; Zoetis). All inseminations were performed at timed AI. The cows were diagnosed for pregnancy by using a blood test for pregnancy associated glycoproteins (PAG; Idexx Bovine Pregnancy Test; Idexx Laboratories, Westbrook, MA) at 25 d after AI and then by using ultrasound by the herd veterinarian on 38 d after AI (Heartland Dairy) or by a theriogenologist from the University of Missouri College of Veterinary Medicine at 32 d after AI (Foremost Dairy). Initially, 177 cows were assigned to the trial, but 17 cows had a PAG result that did not agree with the ultrasound result and were excluded from the analysis so that 160 cows were used.

Blood samples were collected from the coccygeal vein into vacutainer tubes containing EDTA. Samples were put on ice until they reached the laboratory where they were centrifuged at 1,500 × g for 15 min and the plasma was stored at −20°C until analysis. Plasma progesterone concentrations were measured at −3, 0, 7, and 25 d relative to timed AI using the MP Biomedical Double Antibody 125I Kit for progesterone (MP Biomedical, Santa Ana, CA). A complete validation for this assay kit was recently published (Pohler et al., 2016). The samples were analyzed in a single assay with an intra-assay coefficient of variation of 7.7%.

Progesterone data were analyzed and receiver operating characteristic (ROC) curves were generated by using logistic regression in SAS (PROC LOGISTIC, SAS ver. 9.4; SAS Institute Inc., Cary, NC). The event was the result of the pregnancy diagnosis. For the purpose of the analysis, plasma progesterone concentration was evaluated as a test for nonpregnancy. The sensitivity (true positive rate) was defined as the proportion of nonpregnant cows that were correctly identified as nonpregnant for a given progesterone concentration (cutpoint). The specificity (true negative rate) was defined as the proportion of pregnant cows that were correctly identified as pregnant for a given progesterone concentration; 1-specificity is the false positive rate (pregnant cows incorrectly diagnosed as nonpregnant).

Progesterone concentration for each of the days was tested independently in the analysis (−3, 0, 7, and 25 d relative to timed AI). Herd was included in the statistical model but later removed when found not significant. The ROC curve plots themselves were created using a macro as described in the SAS Knowledge Base (Sample 25018: Plot ROC curve with cutpoint labeling and optimal cutpoint analysis; http://support.sas.com/kb/25/018.html). Cutpoints were selected that had the greatest sensitivity (true positive rate) with a false positive rate of less than 5%.

We noted 82/160 (51.3%) cows that were pregnant at the time of diagnosis by ultrasound. The area under the ROC curve for a progesterone test for nonpregnancy on d −3 (PGF2α treatment) was 0.68 (Figure 1A; P < 0.003). A cutpoint at 0.51 ng/mL of progesterone on the day of PGF2α had a sensitivity of 28.2% with a false positive rate of 3.7%, meaning that 28.2% of all nonpregnant cows would be correctly diagnosed (cows with ≤0.51 ng/mL diagnosed as not pregnant; Figure 1A). The same cutpoint would incorrectly diagnose 3.7% of pregnant cows as nonpregnant (false positive). The area under the ROC curve for a progesterone test for nonpregnancy on d 0 (day of AI) was 0.52 (P < 0.05; Figure 1B). A cutpoint of 0.43 ng/mL had a sensitivity of 17.9% with a false positive rate of 4.9%. The area under the ROC curve for a progesterone test for nonpregnancy on d 7 (7 d after AI) was 0.55 (P > 0.10; Figure 1C) and was not different from chance (area = 0.50). A cutpoint of 1.82 ng/mL had a sensitivity of 23.1% with a false positive rate of 4.9%. The area under the ROC curve for a progesterone test for nonpregnancy on d 25 after AI was 0.89 (P < 0.001; Figure 1D). A cutpoint of 2.67 ng/mL had a sensitivity of 76.0% with a false positive rate of 0%.

Progesterone concentrations immediately before AI are associated with improved fertility (Pursley and Martins, 2011; Wiltbank et al., 2011, 2014), and the favorable ROC curve area for the day of PGF2α treatment (0.68) supports this observation. The cutpoint established for the day of timed AI (0.43 ng/mL) was similar to the 0.3 to 0.5 ng/mL range proposed by Wiltbank et al. (2014) as detrimental to fertility. Wiltbank et al. (2015) reported that administering a second PGF2α treatment 24 h after the first decreased the percentage of cows with progesterone greater than 0.5 ng/mL from 17% (approximately equal to what we observed in our study) to 2.4%. The second PGF2α treatment increased P/AI in the Wiltbank et al. (2015) study as would be expected based on the ROC curve that we present. A cutpoint of approximately 1.5 ng/mL was established 7 d after AI. This appears to be a suitable cutpoint for identifying cows that did not ovulate within 1 to 2 d after GnRH treatment. A non-
pregnancy test based on progesterone on d 25 after AI had the greatest area under the ROC curve (0.89), the greatest sensitivity (76%), and lowest false positive rate (0%). The sensitivity and specificity for the d-25 progesterone test was similar to that reported in a review of progesterone-based pregnancy tests (Nebel, 1988). At 25 d after AI, nonpregnant cows were presumably returning to estrus and had low progesterone, whereas pregnant cows maintained the CL and had elevated blood progesterone concentrations (greater than 2.67 ng/mL). With respect to the 4 tests, the false positive rate was less than 5% as defined by our criteria. Lower-

**Figure 1.** Receiver operating characteristic (ROC) curves for a bovine test for nonpregnancy in lactating dairy cows undergoing an Ovsynch program for timed AI based on plasma progesterone concentration on the day of PGF$_{2\alpha}$ treatment (A), the day of AI (B), 7 d after AI (C), and 25 d after AI (D). The labeled points on the ROC curve are the cutpoint (ng/mL of progesterone) for the test. The points were identified by using the default values in SAS (SAS Institute Inc.) for Youden index (labeled points exceed 50% of the maximum of all cutpoints) with a separation of at least 0.10 for sensitivity. The arrow indicates the cutpoint described in the text and also in the analysis in Table 1.
ing the cutpoint for a test would reduce the number of false positives and also decrease the number of true positives. For example, lowering the cutpoint on the day of PGF\(_{2\alpha}\) treatment to 0.14 ng/mL would decrease to false positive rate to 0%, but at the same time decrease the number of true positives from 28.2 (cutpoint = 0.51) to 15.4% (cutpoint = 0.14; Figure 1A).

Given the possibility that plasma progesterone could be used to identify nonpregnant cows, we retrospectively examined a second data set (Escalante et al., 2013) to determine if the cutpoints that we described in the previous 2 paragraphs would apply. Data from the second data set were for cows (n = 198) undergoing 2 different presynchronization programs followed by an Ovsynch 56 program (Escalante et al., 2013). A different progesterone RIA, which is no longer available from the manufacturer, was used. Progesterone on the day of PGF\(_{2\alpha}\) (3 d before AI), on the day of GnRH (1 d before AI), and 7 d after GnRH (6 d after AI) were tested. The presynchronization treatment was initially included in the model but was removed because it was not significant (P > 0.10). We observed 78/198 (39.4%) cows pregnant at the time of diagnosis by ultrasound. The area under the ROC curve for progesterone on the day of PGF\(_{2\alpha}\) was 0.57 (P < 0.09; Figure 2A). A cutpoint of 0.54 ng/mL had a sensitivity of 15.8% and a false positive rate of 3.8%. The area under the ROC curve for a progesterone test on the day of GnRH (0.52) was not significant (P > 0.10; Figure 2B). A cutpoint of 0.54 ng/mL had a sensitivity of 21.7% and a false negative rate of 2.6%. The area under the curve for 7 d after GnRH (0.54) was not significant (P > 0.10; Figure 2C). A cutpoint of 1.49 ng/mL had a sensitivity of 23.3% and a false negative rate of 3.8%.

Across both studies, approximately 20% of nonpregnant cows could be identified before or shortly after AI with a false positive rate of less than 5%. The ability to identify nonpregnant cows based on progesterone can be explained by the underlying biology of ovulation synchronization. Cows with low progesterone at the time of PGF\(_{2\alpha}\) do not have a functional CL and therefore do not respond to PGF\(_{2\alpha}\) treatment. Cows with elevated progesterone 2 to 3 d after PGF\(_{2\alpha}\) have either a developing CL or an incompletely regressed CL after PGF\(_{2\alpha}\) treatment. Either scenario is known to compromise fertility (Pursley and Martins, 2011; Wiltbank et al., 2011, 2014). Finally, cows that have not ovulated after GnRH treatment do not have an increase in progesterone after 6 or 7 d. Despite the fact that the 2 studies employed different assays and also had slightly different sampling times, the cutpoints for the progesterone-based nonpregnancy test were remarkably similar (below approximately 0.5 ng/mL on the day of PGF\(_{2\alpha}\); above approximately 0.5 ng/mL 2 to 3 d later; and below 1.5 ng/mL 6 to 7 d after AI). Also similar was the general shape of the ROC curve and the progesterone concentration at which the sensitivity and 1-specificity were approximately equal (point at which the line crossed the diagonal; the cutpoint would not distinguish nonpregnant from pregnant cows).

The nonpregnancy tests based on progesterone do not necessarily achieve statistical significance in the ROC curve. The test only identifies a subpopulation of cows that have a very low likelihood of pregnancy based on the known biology of ovulation synchronization. When cows that either met or failed to meet the criteria were analyzed, the P/AI for those cows that met the criteria were approximately 3 to 4 times greater than those that failed to meet the criteria (Table 1). We further examined the ROC curve by combining 2 or 3 cutpoints together. Cows that fell below the cutpoint at PGF\(_{2\alpha}\) and those that fell above the cutpoint 2 to 3 d after PGF\(_{2\alpha}\) are not necessarily the same cows. Thus, a cumulative effect of applying 2 cutpoints to identify nonresponding cows exists. When a combination of criteria was used and both studies were combined, the true positive rate was 38.4% and the false positive rate was 6.3% for identifying nonpregnant cows before insemination (cows failing the test at either PGF\(_{2\alpha}\) or GnRH/AI; Table 1). If a third cutpoint at 7 d after GnRH/AI was added, 50.5% of cows that were eventually identified nonpregnant could be identified with a false positive rate of 10.6% (Table 1).

Using a combination of 3 cutpoints identified 2 populations of cows. Cows that failed 1 of 3 tests (117/358; 32.7%) had a P/AI of 14.5% and cows that passed all 3 tests (241/358; 67.3%) had a P/AI of 59.3% (Table 1). The P/AI for responding cows (59.3%) demonstrates that lactating cows in confinement dairies with the correct response to timed AI treatment have P/AI that approaches 60%. This level of fertility is similar to or better than what is reported for Holstein heifers that were AI after estrus (Lopes et al., 2013) as well as post-partum beef cattle (Stevenson et al., 2015). The result suggests that a major cause of infertility in lactating dairy cows is abnormal ovarian function within a few days of AI.

Diagnosing nonpregnancy before or shortly after AI will not increase reproductive efficiency unless cows known to be not pregnant are managed for rebreeding. Cows that are diagnosed as nonpregnant before AI could be resynchronized for AI the following week by using an Ovsynch program (Thatcher and Santos, 2007; Wiltbank and Pursley, 2014). This will theoretically improve their P/AI from approximately 5 to 10% to the herd average for timed AI on the following week. Considering both studies combined (Table 1), we identified 117 cows that failed the cutpoint on either
the day of PGF$_{2\alpha}$ or day of GnRH/AI or 7 d after GnRH/AI. A total of 17 pregnant cows (false positives) in the group of 117 (Table 1). Assuming a 44.7% P/AI (based on both studies combined), resynchronizing the 117 cows failing the test would give 52 pregnancies (35 greater than the original number of pregnancies in the group of 117 cows). Adding these 35 pregnancies to the original total (160 pregnancies) would yield a P/AI of 54.5% (195 pregnancies out of 358 cows), which is a 9.8 percentage point increase in P/AI from the original 44.7%. Achieving herd average assumes that the cows identified as nonpregnant by testing will have normal
Table 1. True positive (TP) and false positive (FP) rate for a plasma progesterone test used to diagnose nonpregnancy; and pregnancies per AI (P/AI) for cows diagnosed as nonpregnant (NP) or pregnant (P) based on a plasma progesterone test for cows tested at PGF<sub>2α</sub> injection (3 d before GnRH/AI), on the day of GnRH/AI, or 7 d after GnRH/AI.

<table>
<thead>
<tr>
<th>Test day&lt;sup&gt;1&lt;/sup&gt;</th>
<th>Study 1&lt;sup&gt;2&lt;/sup&gt;</th>
<th>Study 2&lt;sup&gt;3&lt;/sup&gt;</th>
<th>Combined&lt;sup&gt;4&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>TP&lt;sup&gt;5&lt;/sup&gt;</td>
<td>FP&lt;sup&gt;6&lt;/sup&gt;</td>
<td>P/AI NP&lt;sup&gt;7&lt;/sup&gt;</td>
</tr>
<tr>
<td>(28.2) (3.7)</td>
<td>(12.0) (0.36)</td>
<td>(58.5&lt;sup&gt;**&lt;/sup&gt;)</td>
<td>(13.6) (0.44)</td>
</tr>
<tr>
<td>(17.9) (3.7)</td>
<td>(17.6) (0.28)</td>
<td>(55.2&lt;sup&gt;**&lt;/sup&gt;)</td>
<td>(7.1) (0.14)</td>
</tr>
<tr>
<td>(23.1) (4.9)</td>
<td>(18.2) (0.28)</td>
<td>(56.5&lt;sup&gt;**&lt;/sup&gt;)</td>
<td>(9.7) (0.14)</td>
</tr>
<tr>
<td>PGF&lt;sub&gt;2α&lt;/sub&gt; or GnRH/AI&lt;sup&gt;6&lt;/sup&gt;</td>
<td>34/78</td>
<td>6/82</td>
<td>6/40</td>
</tr>
<tr>
<td>(43.6) (7.3)</td>
<td>(15.0) (0.36)</td>
<td>(63.3&lt;sup&gt;**&lt;/sup&gt;)</td>
<td>(8.7) (0.14)</td>
</tr>
<tr>
<td>PGF&lt;sub&gt;2α&lt;/sub&gt; or GnRH/AI or 7 d after GnRH/AI&lt;sup&gt;7&lt;/sup&gt;</td>
<td>42/78</td>
<td>10/82</td>
<td>10/52</td>
</tr>
<tr>
<td>(53.8) (12.2)</td>
<td>(19.2) (0.36)</td>
<td>(66.7&lt;sup&gt;***&lt;/sup&gt;)</td>
<td>(10.8) (0.14)</td>
</tr>
</tbody>
</table>

<sup>1</sup>Day of Ovsynch program when blood sample was collected for progesterone analysis for non-pregnancy test.
<sup>2</sup>Present study. Pregnancies per AI (P/AI; all cows) = 82/160 (51.3%).
<sup>3</sup>Reanalysis of Escalante et al. (2013). P/AI (all cows) = 78/198 (39.4%).
<sup>4</sup>Study 1 and 2 combined. P/AI (all cows) = 160/358 (44.7%).
<sup>5</sup>Number of NP cows below (PGF<sub>2α</sub> or 7 d after GnRH/AI) or above (GnRH/AI) selected cutpoint for nonpregnancy divided by total number of NP cows. The selected cutpoints for nonpregnancy were 0.51 and 0.54 for the day of PGF<sub>2α</sub>, 0.43 and 0.54 for the day of GnRH/AI, and 1.84 and 1.49 for 7 d after GnRH/AI (study 1 and 2, respectively).
<sup>6</sup>Number of P cows below (PGF<sub>2α</sub> or 7 d after GnRH/AI) or above (GnRH/AI) selected cutpoint for nonpregnancy divided by total number of P cows. The selected cutpoints for non-pregnancy were 0.51 and 0.54 for the day of PGF<sub>2α</sub>, 0.43 and 0.54 for the day of GnRH/AI, and 1.84 and 1.49 for 7 d after GnRH/AI (study 1 and 2, respectively).
<sup>7</sup>Pregnancies per AI (P/AI) for cows diagnosed NP based on selected cutpoint for nonpregnancy.
<sup>8</sup>Pregnancies per AI (P/AI) for cows diagnosed P based on selected cutpoint for nonpregnancy.
<sup>9</sup>Cumulative diagnosis of nonpregnancy based on 2 tests (PGF<sub>2α</sub> and GnRH/AI) where the cow was diagnosed NP if she tested nonpregnant for either test.
<sup>10</sup>Cumulative diagnosis of nonpregnancy based on 3 tests (PGF<sub>2α</sub>, GnRH/AI and 7 d after GnRH/AI) where the cow was diagnosed NP if she tested nonpregnant for any test.

**P < 0.01, ***P < 0.001 versus NP.
fertility. This may or may not be the case. For example, cows that have low progesterone at PGF<sub>2α</sub> may not be cycling and have inherently lower fertility. The application of a progesterone-releasing insert to cows identified as nonpregnant by using a CIDR_Ovsynch program may overcome some of the limitations in noncycling cows, as demonstrated in recent studies (Bisinotto et al., 2013, 2015).

In practice, the third test at 7 d would only be applied to cows that passed the first 2 tests. In the combined studies, 22 of the 53 cows (41.5%) that failed the d-7 test also failed at either PGF<sub>2α</sub> or GnRH/AI (Table 1), leaving 31 cows that were newly identified with the third test. The value of the d-7 test is decreased because about 40% of d-7 failures would have already been identified by an earlier test. The third test is also applied to cows that have completed the AI program and is unlike earlier testing that can potentially eliminate unproductive treatments or AI. The economic value of the third test should be considered relative to earlier testing with respect to the reduced number of newly identified cows and the possible interventions that can be applied.

In 2 of our previous studies, we noted that the response to presynchronization before the timed AI had an effect on the P/AI after insemination (Escalante et al., 2012, 2013). Specifically, cows that came into estrus or had reduced progesterone by 5 d after presynchronization had greater P/AI. We hypothesized that the estrus event or low progesterone after presynchronization indicated that the cow was cycling and the presynchronization was successful. For example, in the study of Escalante et al. (2012), cows with a paint score = 0 (indicative of estrus) after presynchronization had a P/AI at timed AI of 48%, whereas cows with a paint score = 5 (indicative of no estrus) had a P/AI at timed AI of 28%. Likewise, in a second paper (Escalante et al., 2013), cows that had high progesterone 5 d after presynchronization (indicative of presynchronization failure) had P/AI of 10.5%, whereas cows with a correct response had a P/AI of 52.2%. In either case, the response to the presynchronization was recorded approximately 2.5 wk before the timed AI. In a reanalysis of the data from Escalante et al. (2013), we found that progesterone at 5 d after the final PGF<sub>2α</sub> injection in the presynchronization treatment of Presynch Ovsynch could be used to identify nonpregnant cows to timed AI (Figure 2D). The area under the ROC curve for progesterone sampling on d 5 for the nonpregnancy test was 0.58 (P < 0.015). A cutpoint of 1.88 ng/mL (cows with greater than 1.88 diagnosed as nonpregnant) would correctly diagnose 25.8% of nonpregnant cows and not misdiagnose any pregnant cows (false positive rate = 0) by 5 d after the completion of the presynchronization. This diagnosis is 7 d before the start of Ovsynch and 17 d before AI.

The collective interpretation of these data is that a single progesterone test before or shortly after AI can be used to identify approximately one-fifth of the cows that will not become pregnant with a false positive rate of less than 5% (Table 1; combined data). If 2 or more time points are combined then the percentage of nonpregnant cows that can be identified increases to approximately one-third (2 time points) to one-half (3 time points; Table 1). Correctly identifying nonpregnant cows either before or shortly after insemination has utility with respect to eliminating a nonproductive insemination (cows identified before AI) or shortening the time to reinsemination (cows identified by one week after AI). Testing 5 d after presynchronization is an additional opportunity to identify approximately 25% of the cows that will fail to become pregnant after timed AI (approximately 2.5 wk later). Although cows could be blood sampled and tested individually, an automated method for progesterone testing is a more likely approach on farms. Although blood could be used, it is difficult to automate the sampling. A possible scenario could include milk as the biological sample for progesterone analysis. Fully automated milk progesterone sampling and analysis is available commercially on farm (Herd Navigator System; DeLaval, Tumba, Sweden), although this system is not typically used in combination with synchronization systems. Regardless, concentrations of milk progesterone are typically greater than plasma so appropriate cutpoints for milk progesterone remain to be established.

ACKNOWLEDGMENTS

The study was funded by the Food for the 21st Century Program of the University of Missouri.

REFERENCES


