Sample frequency and reactivity of farmers influences correct insemination window on farms using on-farm milk progesterone for fertility management.

Dyan Meuwissen[1], Kim Brosens[1], Ben Aernouts[1], Ines Adriaens[1,2]

[1]KU Leuven, Livestock Technology, Biosystems department, Kleinhoefstraat 4, 2440, Geel, Belgium

[2]Wageningen University and Research, Animal Breeding and Genomics, 6700 AH Wageningen, the Netherlands

Corresponding author: ines.adriaens@kuleuven.be

ABSTRACT

Improving fertility of dairy cows on farm remains difficult, as it is a complex trait which depends on physiological, management and external factors. Even small improvements in fertility (e.g. in terms of conception rate) potentially make a big difference for cow longevity, lifetime production and as such, sustainability. The availability of 1) new sensor data to get a better grip on physiological status of cows and 2) new systems to accurately predict the best insemination window based on milk progesterone (P4), offers new opportunities to analyse and understand fertility performance, and subsequently improve this trait. Before external and physiological factors that impact conception success can be investigated, it is crucial that we have certainty on whether the correct insemination time was applied. This study uses an extensive dataset from 5 modern dairy farms with a Herd Navigator system (HN), which farmers use to determine the optimal moment of insemination upon a drop in milk P4 preceding luteolysis. Despite that in many cases the estimation of the insemination moment goes well, inconsistencies in detection of luteolysis arise when sampling rate or milking frequency deviate from optimal conditions, which subsequently impacts fertility performance. More specifically, the use of an algorithm for which a time lag is introduced that depends on both absolute P4 concentration, rate of change of P4 during luteolysis and a fixed threshold, leads to inconsistencies in detection of luteolysis when sampling rate or milking frequency deviate. In this study, we analysed the factors influencing

correct identification of the insemination window, to better separate cases in which conception failed due to human or system errors and cases in which the cow physiology caused poor fertility performance. To this end, we analysed different traits related to insemination and sampling frequency of the Herd Navigator and looked at their effect on success of conception across the 5 farms. We found that, in farms with worse fertility performance, the time between the HN alarm and the true luteolysis (i.e. the first raw P4 measurement below 5 ng/mL) was greater and less consistent. When this time was used to distinct between high- and low P4 sampling rate cycles, successful insemination that followed a low P4 sampling rate has significantly less time between an insemination and the preceding HN alarm compared to unsuccessful inseminations following a low P4 sampling rate. Moreover, successful insemination that followed a low P4 sampling rate, are associated with less time between an insemination and the preceding HN alarm compared to successful inseminations that followed a high P4 sampling rate. These finding confirm that when the time between the HN alarm and the first raw P4 measurement below 5 ng/mL is greater, farmers should inseminate sooner. From this analysis, we can make a better selection of the cycles in which insemination time was correct, and thus depended on other factors. This is a firs step towards improved understanding of fertility performance on farm, and as such, of a more sustainable dairy sector.

Key words: dairy cow fertility, milk progesterone, online milk sampler, performance evaluation

INTRODUCTION

Optimal fertility management is paramount to the profitability and sustainability of the dairy sector. Poor reproduction performance not only leads to increased calving intervals and a high number of inseminations per conception, thereby decreasing the profit per cow, but also has significant indirect effects such as reduced genetic gain, increased veterinary costs and involuntary culling (Galvão *et al*., 2013; Krpálková *et al.,* 2020). Accordingly, already the slightest improvement in fertility can have great economic and environmental impact, for example through the optimization of the profitable fraction of the herd and lowering total herd green-house gas emissions. Improving reproductive performance, however, has proven difficult despite previous efforts (Friggens & Chagunda, 2005; Hansen, 2011).

Fertility is a complex trait and has been shown to depend on e.g., how a cow transitions from non-lactating to lactating upon calving, and the concurrent negative energy balance and depressed immunity are shown to have a negative effect on reproduction performance (Leroy *et al.*, 2008; Leroy *et al.*, 2011). Hence, fertility is a multi-factorial and complex trait which is influenced by a cow's physiological status, the farm's environmental and management conditions and their interactions. With the plethora of data collected by sensors on modern farms, opportunities for data-based exploration and understanding of these factors rise. A crucial starting point when looking into factors influencing reproduction success is having identified the correct moment of insemination. When this condition is fulfilled and the insemination was done correctly at the right moment in the cycle (i.e., 12h before ovulation), other factors that determine success or non-success of the service can be studied.

In recent years, the use of milk progesterone (P4) has become the gold standard for identifying correct insemination time, as it allows to detect luteolysis preceding ovulation, resulting in a steep drop in the P4 concentration. With the on-farm measurement device 'Herd Navigator^{TM} (HN)', milk P4 can be measured in an automated and continuous way. To limit analysis costs, this device combines a smart sampling scheme based on the calving date, the time in cycle and the P4 concentration with an algorithm

to interpret the P4 dynamics and predict the ideal insemination window. The performance of the latter algorithm depends on how often the cows are milked, on the raw P4 concentration measured, the reactivity of the farmer upon the alarm and how meticulously they inseminate the cows, as also pointed out by Adriaens et al. (2019). In non-optimal conditions, for example when cows are not regularly milked, the fixed thresholds and predictions lack flexibility to deal with the variability across different cycles from the same cow and between different cows. While for many cases the system works well, in some cases it does not, which restricts the reliability of the predicted insemination window that was based upon it, and therefore, its role as a gold standard for evaluating fertility success is compromised. Based on historical data collected at different farms with different reproduction performance levels, however, we can increase our understanding of the system, identify when it does or does not work well and how it potentially can be improved. With these insights, a selection of cases in which the moment of insemination was probably correct, can be made. This can in its turn help to unravel other factors that affect fertility success. Therefore, the HN system and its application on different commercial dairy farms will be the focus of this study. This leads to the research question: Are there differences in the application of the HN system on different dairy farms and what are its consequences for insemination success?

In this study we aim to improve the reliability of the usage of P4 as a gold standard for evaluating fertility success. With the acquired insights, decision support towards identification of the correct insemination window can be achieved, which can, additional to the P4 concentration, cow's intrinsic value, health and management, help the farmer to improve fertility management. To answer the research question, this research answers following questions: (1) what are the key farm-, fertility- and cyclicity characteristics on five commercial dairy farms with a HN system; (2) what are the differences in the HN bio-model between the five farms; (3) what are the differences between farms in their insemination timing; (4) what are the differences in the HN bio-model between successful and unsuccessful inseminations. This will ultimately allow us to combine the application of the HN bio-model and the insemination reactivity of the farmer, thereby explaining conception success across farms.

MATERIAL AND METHODS

Data collection and selection

Available data. Software backups of the farm management system were collected for 4 Belgian and 1 Dutch farm that use an automated milking system (AMS) from Delaval, Tumba, Sweden and a HN system (Lattec, Hillerod, Denmark). The time period covered with these data varied between 2016 and 2021. On average, 1.8 lactations were included per animal. All farms had intensive production systems, with cows kept indoors and fed with both forage and concentrates. Raw data tables containing milking and milk P4 data, together with ancillary cow information, were extracted from the backup files of the AMS software system using SQL Server Management Studio (Microsoft cor., Redmond, WA). The further data editing of the data and the rest of the analyses described below were performed in Python (3.7).

Selection of the cycles. Only complete P4 estrus cycles were retained for the analysis. A complete cycle starts with a HN alarm, indicating that luteolysis has happened. This is followed by a period of low P4 concentration, indicating the follicular phase, and subsequently an increase in P4 concentration, representing the luteal phase. Next, the cycle is completed by a decrease in P4 concentration, generating the next HN alarm. After extraction of the data tables, milk P4 data was selected based on the availability of the complete P4 estrus cycle and the availability of all milk records within the same estrus cycle.

Insemination selection. Next, a distinction was made between unsuccessful and successful inseminations. An insemination was classified as unsuccessful if the P4 level dropped indicating a new luteolysis after insemination. Inseminations were classified as successful when no drop in P4 level was detected and when a next calving was found between 8 and 10 months after the respective insemination.

Calculation of key farm-, fertility- and cyclicity characteristics

The following cyclicity characteristics were calculated: (1) the cycle length, which was the time between two HN alarms and normally lasts for 18 to 24 days (Crowe, 2011); (2) the length of the postpartum anestrus period, which is the period after calving when the P4-profile shows no cyclicity yet and is calculated as the period of low P4 concentration until the first measurement above 5 ng/mL; (3) the length of the follicular phase, i.e. the period in which the P4 concentration is low and where the dominant follicle can develop towards a preovulatory stage and consecutively ovulate (typically 4 to 6 days (Adriaens et al., 2017)); (4) the length of the luteal phase, i.e. the period in which the P4 concentration is slowly increasing from basal to maximal luteal P4 concentration which is associated with the growth of the CL after ovulation (typically 14 to 18 days (Ranasinghe et al., 2011); and (5) the percentage of cycles with prolonged follicular phase (i.e. exceeding 15 days) and (6) the percentage of cycles with prolonged luteal phase (i.e. exceeding 20 days). The duration of the follicular and luteal phase are calculated using a model that combines a Hill function to describe the increasing part of the cycle during luteal development, and a decreasing Gompertz function to model the drop in P4 during luteolysis (Adriaens et al. 2017). This combined model describes the P4 concentration throughout a single estrus cycle, thereby characterizing the follicular phase, luteal phase and following luteolysis (Figure 1). The length of the follicular phase is calculated as the period from the start of the cycle until this model exceeds 5 ng/mL. The length of the luteal phase is calculated from the moment the model exceeds 5 ng/mL until the moment the model P4 concentration drops below 5 ng/mL again minus 12 hours, which is assumed to be the average duration of the luteolysis.

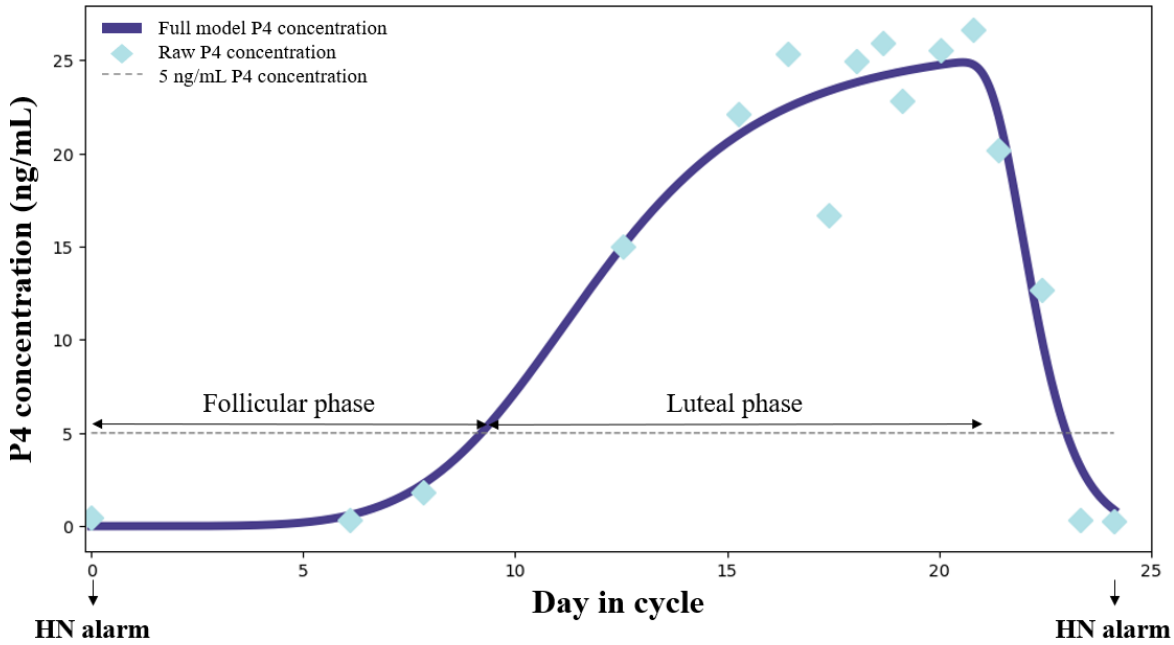

Figure 1. Progesterone (P4) concentration throughout an estrus cycle (i.e HN alarm to HN alarm) described with a combination of an increasing Hill function and a decreasing Gompertz function. The length of the follicular and luteal phase can be calculated using the full model P4 concentration.

Herd Navigator bio-model

Smooth progesterone values. To reduce random noise and allow interpretation of the P4 time series, the raw P4 time series are smoothed. To generate these smooth values (ng/mL), a multiprocess "extended" Kalman filter is used. This algorithm takes the probability and direction and change of the P4 measurements into account as described in Adriaens et al. (2018a, 2019); Friggens and Chagunda (2005); Friggens et al., (2008); Løvendahl and Chagunda, (2010), thereby reacting conservatively to outliers and quickly to measurements that rapidly change level or slope (e.g., during luteolysis). These are the estimated P4 level at time *t* from samples up to and including time *t*. The optimum trade-off between time lag in availability of samples with a higher accuracy of smooth value estimation may vary according to local farm conditions (e.g. cow's milk interval and farmer's insemination timing). This means that the ability of the HN bio-model to distinguish between cycle phases and detect luteolysis depends on the accuracy of the smooth P4 (SmP4) values and on the sampling rate of the raw P4 (RaP4) values.

Progesterone sampling rate and interpretation. The HN system determines the sampling rate such that the maximum amount of information is extracted from the P4 data with as little as possible measurements (Adriaens *et al.*, 2019). To this end, a low P4 sampling rate of about one measurement per 3 á 4 days is maintained in the first 15 days of a cycle, after which it increases to one measurement per 0.5 á 1 day when luteolysis is expected. As soon as the RaP4 concentration drops below 5 ng/mL, a next P4 sample is taken from the successive milking. Once the SmP4 concentration drops below 5 ng/mL, a HN alarm is generated and luteolysis is detected. The RaP4 and SmP4 concentration throughout an individual estrus cycle and the hours between current and next P4 measurement (i.e. the sampling rate) are presented in Figure 2.

Exploring traits related to sampling frequency of the Herd Navigator and insemination timing

This study aims at improving the reliability of the usage of P4 as a gold standard for evaluating fertility success. To study the factors that influence the reliability of the predicted insemination window and subsequently conception success, several P4 sampling rate-, milking frequency and insemination timing related traits were derived and explored across the 5 farms and between different estrus cycles, between unsuccessful and successful inseminations within farms, and between high and low P4 sampling rate cases within and across farms.

1st RaP4 and 1st SmP4 concentration below 5 ng/mL. In this study, we focus on two P4 sample frequency related traits that might influence the correct identification of luteolysis and subsequently correct estimation of the insemination window. The first trait is the time between the moment the RaP4 concentration drops below 5 ng/mL preceding a HN alarm, which serves as a proxy for the time between the true moment of luteolysis, and the previous observation (i.e., the last observation above 5 ng/mL). Trait 1 is indicated by the 1st arrow in the upper panel of Figure 2. This trait corresponds to the measurement time between the RaP4 above and below 5 ng/mL during luteolysis. When this time increases, the uncertainty on when luteolysis has taken place increases, resulting in a higher probability

that the insemination window was inaccurately determined and ovulation was missed. The second trait is the time between the moment the SmP4 concentration drops below 5 ng/mL, which generates a HN alarm, and the moment the RaP4 concentration drops below 5 ng/mL. This is indicated by the 2nd arrow in the upper plot of Figure 2. When the time between these two observations increases, the delay between the true luteolysis and the HN alarm becomes larger, thereby potentially not identifying luteolysis and subsequent ovulation on time.

As a first step to evaluate the usage of P4 as a gold standard for evaluating fertility success, trait 1 and 2 were compared across different estrus cycles. Since the low sampling rate in the first 15 days of the cycle and the followed increase when luteolysis is expected, depends on several key points the P4 estrus cycle (Friggens and Chagunda, 2005), it was expected that the HN bio-model deviates from the general sampling rate when the P4 cycle profile gets disturbed. It was explored whether the time between the 1st RaP4 below 5 ng/mL and the previous observation and the time between the 1st SmP4 below 5 ng/mL and the 1st RaP4 below 5 ng/mL differed for 'normal' cycles compared to cycles that either (1) have a prolonged follicular phase (>15 days); (2) have a deviating luteal phase (< 5 days or > 20 days); or (3) have both a prolonged follicular and deviating luteal phase. Cycles in which the time between 1st RaP4 below 5 ng/mL and the previous observation was > 120h or the time between 1st SmP4 below 5 ng/mL and 1st RaP4 below 5 ng/mL was > 72h were removed. Finally, trait 1 and trait 2 calculated for "normal" P4 estrus cycles were compared between the 5 farms.

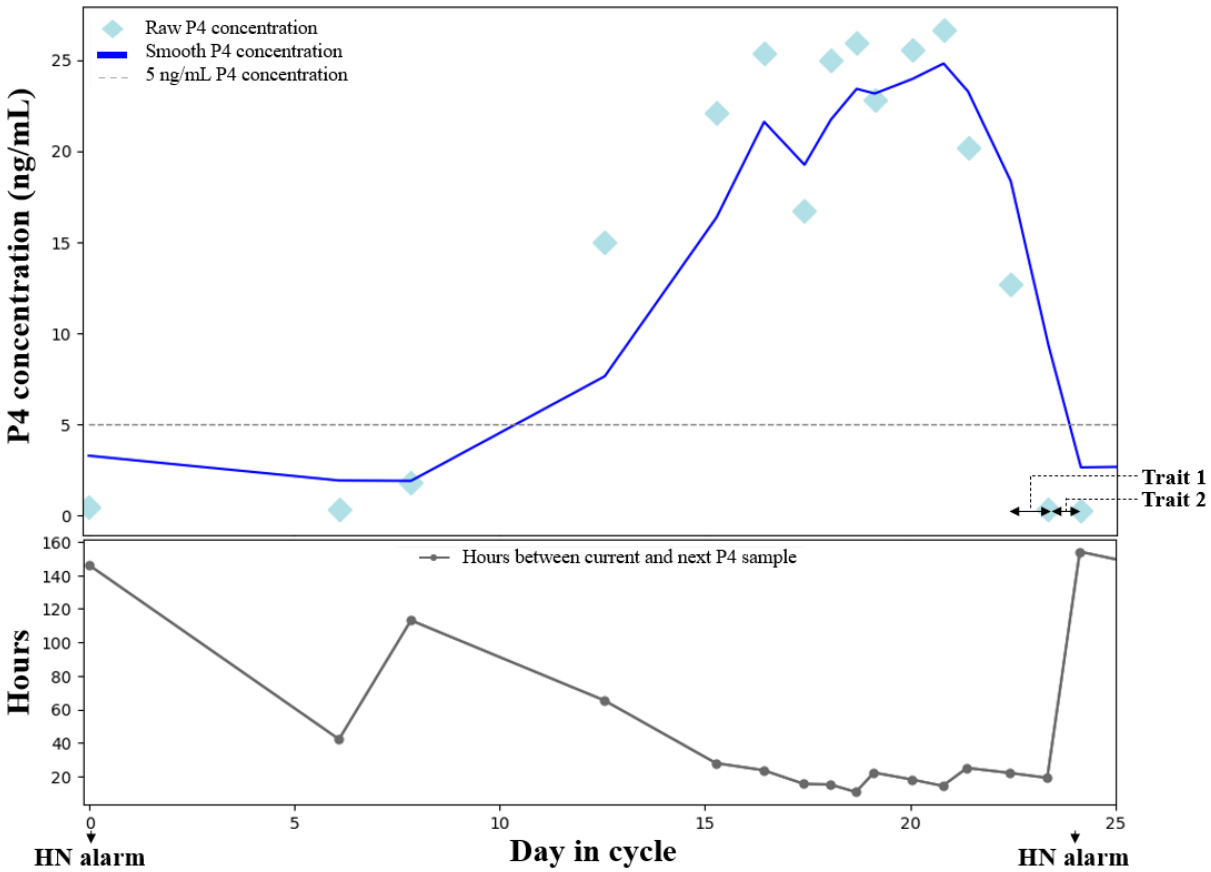

Figure 2. Raw and smooth progesterone (P4) concentration and the time between current and next P4 sample throughout an estrus cycle. To evaluate correct timing of luteolysis, two traits are focused on. (1) the time between the 1st raw P4 measurement below 5 ng/mL and the previous observation; (2) the time between 1st smooth P4 measurement below 5 ng/mL (i.e., the HN alarm) and the 1st raw P4 measurement below 5 ng/mL.

Milking interval. When cows are milked by AMS, the milking interval (MI) depends on multiple factors including the milk yield, parity, settings of the robot, the lactation stage and the health of the cows. For example, lame cows will visit the AMS less frequently. The maximal sampling rate for P4 analysis corresponds with the milking frequency (one sample/milking). Therefore, the milking frequency can impact the performance of the HN system and the traits related to the P4 sampling rate discussed above. More specifically, a larger MI causes a larger difference between the 1st RaP4 measurement below 5 ng/mL and the previous observation, thereby causing a higher uncertainty on when luteolysis has taken place and increasing the inaccuracy of the determined insemination window. Likewise, a larger MI causes a larger difference between the 1st SmP4 below 5 ng/mL and the 1st RaP4 below 5 ng/mL, thereby increasing the delay between true luteolysis and the HN alarm which causes

an incorrect estimation of the insemination window. Next, the MI was calculated for early and late lactation cows and outliers were removed if the difference between two consecutive milkings was > 2 days. As a milk sample for P4 analysis can only be taken at most each milking, it was expected that the MI can be an explanation for potential differences found for the two P4 sample frequency related traits.

Expected insemination moment. For one farm (Farm 2), the exact moment of insemination was available. For the other 4 farms, we had only access to the insemination date. The farmers of these farms indicated that they inseminate either in the morning between 7h00 and 8h00 or in the afternoon between 16h00 and 17h00. To get the best possible estimation of the exact insemination time, we therefore assumed that insemination took place at 12h00. With this approach, an error of at most 5 hours was introduced (Adriaens et al., 2018b). Moreover, the farmers indicated that HN advices to inseminate between 30 and 38 hours for primiparous cows and between 38 and 46 hours for multiparous cows. In this study, we focus upon two insemination timing related traits that might influence conception success. The first trait is the difference between the HN alarm and the insemination moment, which reflects the reactivity and consistency of the farmers inseminating upon a HN alarm. The second trait is the difference between the 1st RaP4 below 5 ng/mL and the insemination moment which represents the insemination timing upon the true luteolysis. These two traits related to insemination timing are important to study since not only conception can fail due to system errors, also due to human errors. Outliers were removed (time between insemination and HN alarm > 150h and time between insemination and 1st RaP4 below 5 ng/mL > 200h) and the two insemination timing traits were compared across the 5 farms and between unsuccessful and successful inseminations within farms.

Separation between high- and low P4 sampling rate. To obtain insights in the combined effect of P4 sampling rate and the reactivity of the farmer upon a HN alarm, a distinction was made between high- and low P4 sampling rate cases based on the difference between the 1st SmP4 below 5 ng/mL and the 1st RaP4 below 5 ng/mL. For this, the median difference between 1st SmP4 and RaP4 below 5 ng/mL is assumed to make the distinction between high P4 sampling rate (time between 1st SmP4 and RaP4

below 5 ng/mL < median) and low P4 sampling rate (time between 1st SmP4 and RaP4 below 5 ng/mL ≥ median) for each farm separately. Next, to identify whether the insemination advice to the farmer should be different between high- and low sampling rate cases, the time between insemination and the HN alarm was compared between unsuccessful and successful inseminations for high- and low sampling rate separately and between successful inseminations of high- and low sampling rate cases within farm.

RESULTS AND DISCUSSION

Key farm-, fertility- and cyclicity characteristics

An overview of the key farm-, fertility- and cyclicity characteristics for the 5 farms is given in Table 1. The period that information was collected and included in this study differed across the farms, but was within the period from March 2016 to June 2021. Farm 1 had the largest contribution to the number of cows and lactations in the dataset (682 cows and 1185 unique lactations). Next was farm 3 with 301 cows and 644 unique lactations, farm 2 with 187 cows and 376 unique lactations, farm 4 with 212 cows and 352 unique lactations, and last farm 5 with 170 cows and 261 unique lactations. Farm 1 and 4 have the lowest average 50-d, 100-d, 150-d and 305-d milk yield. Also, farm 1 has the worst fertility performance in terms of the lowest 1st service conception rate, highest number of inseminations per conception, and largest calving interval and days open. Farm 4 and 5 have the best numbers for these fertility characteristics. Farm 2 has the highest percentage of cycles with a prolonged follicular phase (> 15 days) and farm 3 with a prolonged luteal phase (> 20 days).

Table 1. Key farm-, fertility and cyclicity characteristics for the 5 commercial dairy farms.

	Farm 1 (s)	**Farm 2 (h)**	**Farm 3 (v)**	**Farm 4 (l)**	**Farm 5 (g)**
Farm characteristics					
Period	2016/03/31 – 2021/03/17	2016/12/07 – 2021/06/20	2017/10/10 – 2021/06/21	2018/02/14 – 2020/10/12	2018/12/15 – 2021/03/15

No. of cows	682	187	301	212	170
No. of lactations	1185	376	644	352	261
No. of primiparous lactations	459	142	169	116	104
No. of multiparous lactations	726	234	475	236	157
Average 50-d milk yield (kg)	1,785 ± 461 [507 – 3,002]	1,827 ± 431 [718 – 2,962]	1,959 ± 454 [684 – 2,808]	1,803 ± 379 [865 – 2,720]	1,860 ± 448 [941 – 2,774]
Average 100-d milk yield (kg)	3,745 ± 850 [1,492 – 5,844]	3,848 ± 822 [1,727 – 6,037]	4,047 ± 867 [1,535 – 5,805]	3,779 ± 712 [1,928 – 5,559]	3,971 ± 832 [2,307 – 5,883]
Average 150-d milk yield (kg)	5,546 ± 1,161 [2,241 – 8,373]	5,728 ± 1,157 [2,724 – 8,799]	5,915 ± 1,213 [2,376 – 8,464]	5,582 ± 996 [3,260 – 8,294]	5,942 ± 1,148 [3,548 – 8,558]
Average 305-d milk yield (kg)	10,155 ± 1,849 [4,275 - 14,834]	10,572 ± 1,924 [5,858 - 15,990]	10,685 ± 2,056 [4,869 - 16,006]	10,144 ± 1,700 [5,464 - 15,181]	10,918 ± 1,811 [6,961 - 14,758]
Fertility characteristics					
Average age at 1st calving (yr)	2.1 ± 0.4 [1.6 - 4.8]	2.0 ± 0.2 [1.7 - 2.8]	2.0 ± 0.1 [1.8 - 2.6]	2.2 ± 0.4 [1.7 - 6.4]	2.0 ± 0.1 [1.7 - 2.5]
1st service conception (%)	12.6	24.7	22.6	21.5	18.1
No. of inseminations per conception	2.9 ± 1.7 [1 - 9]	2.5 ± 1.6 [1 - 8]	2.3 ± 1.6 [1 - 9]	2.4 ± 1.5 [1 - 7]	1.9 ± 1.2 [1 - 6]
DIM at 1st insemination	62 ± 15 [24 - 101]	67 ± 15 [34 - 102]	69 ± 13 [46 - 97]	61 ± 16 [31 - 98]	64 ± 14 [48 - 94]
Calving interval (d)	412 ± 68 [299 - 691]	407 ± 67 [306 - 725]	395 ± 59 [191 - 666]	390 ± 60 [288 - 804]	389 ± 69 [313 - 699]
Days open	134 ± 68 [41 - 407]	132 ± 60 [45 - 426]	119 ± 62 [50 - 369]	113 ± 57 [40 - 332]	118 ± 58 [51 - 333]

Cyclicity characteristics					
Average cycle length (d)	28 ± 10 [12 - 90]	29 ± 11 [13 - 90]	30 ± 13 [13 - 93]	28 ± 10 [11 - 99]	28 ± 9 [15 - 98]
Average postpartum anestrus length	44 ± 18 [27 - 174]	44 ± 18 [27 - 135]	39 ± 16 [27 - 135]	50 ± 25 [27 - 209]	38 ± 14 [27 - 99]
Average follicular phase length (d)	9 ± 6 [0.5 - 86]	10 ± 7 [0.5 - 57]	10 ± 7 [1 - 63]	10 ± 8 [0.5 - 84]	10 ± 6 [5 - 74]
Average luteal phase length (d)	16 ± 9 [5 - 100]	16 ± 9 [5 - 80]	17 ± 12 [5 - 96]	16 ± 8 [5 - 84]	15 ± 9 [5 - 87]
% of cycles with prolonged follicular phase (>15 days)	6.7	9.8	9.2	5.6	7.2
% of cycles with prolonged luteal phase (>20 days)	16.8	20.3	22.4	16.6	17.8

Traits related to P4 sampling rate across different P4 estrus cycles

As a first step to evaluate the P4 sampling rate and how it affects fertility success, the time between the 1st RaP4 below 5 ng/mL and the previous observation was explored. Clear differences were found between "normal" and "deviating" P4 estrus cycles (Table 2). Across the 5 farms, the average time between the 1st RaP4 measurement below 5 ng/mL and the previous observation for "normal" cycles was 27.0 ± 22.5 h, for the cycles with only a deviating follicular phase this was 47.3 ± 35.7 h, for the cycles with only a deviating luteal phase this was 57.0 ± 34.1 h, and for the cycles with both a deviating follicular and luteal phase this was 34.1 ± 43.3 h. These numbers confirm that when a P4 estrus profile deviates from "normal", a greater time and deviation between the 1st RaP4 measurement below 5 ng/mL and the previous observation is found. This trend is also found within the individual farms between "normal" and "deviating" P4 estrus cycles (Table 2). This suggests that the HN bio-model functions less well when follicular- and luteal phase lengths are suboptimal, thereby causing an increase in the

time. When the time between the 1st RaP4 below 5 ng/mL and the previous observation increases, it is more likely that true luteolysis and subsequent ovulation and insemination window is not reliably estimated.

Next, the time between the 1st SmP4 and 1st RaP4 measurement below 5 ng/mL was explored. Less clear differences were found between "normal" and "deviating" P4 estrus cycles compared to the previous P4 sampling rate related trait (Table 2). Across the 5 farms, the average time between the 1st SmP4 measurement below 5 ng/mL and the 1st RaP4 below 5 ng/mL for "normal" cycles is 16.6 ± 9.2 h, for the cycles with only a deviating follicular phase 15.4 ± 10.4 h, for the cycles with only a deviating luteal phase 14.3 ± 9.5 h, and for the cycles with both a deviating follicular and luteal phase it is 15.6 ± 9.6 h. Also within individual farms, less concrete differences exist between "normal" and "deviating" P4 estrus cycles. These numbers are as expected since a next P4 sample is taken from the successive milking as soon as the RaP4 concentration drops below 5 ng/mL, which causes the SmP4 concentration to drop below 5 ng/mL thereby generating a HN alarm. Therefore, the time between the 1st SmP4 and 1st RaP4 measurement below 5 ng/mL is independently of the length of the follicular or luteal phase.

Nevertheless, there is an increase in time and variability between the 1st RaP4 below 5 ng/mL and the previous observation for P4 estrus cycles with a "deviating" profile compared to cycles with a "normal" profile. To explain the variability in fertility success, it is important to start from reliable information that reflects typical farm-specific HN related characteristics. Moreover, to understand the effect of P4 sampling rate on conception success, the inseminations that resulted in embryonic mortality should not be included. By assuming that a luteal phase length > 20 is deviating, (at least the late) embryonic mortality cases are excluded. Therefore, only estrus cycles with a "normal" P4 profile were selected for further analysis.

Table 2. Mean and standard deviation (std) for the two traits: (1) hours and number of milkings between 1st RaP4 below 5 ng/mL and the previous observation and (2) hours and number of milkings between 1st SmP4 below 5 ng/mL and the previous observation for "normal" P4 estrus cycles and cyles with a deviating follicular (> 15 days), luteal (< 5 or > 20), or follicular and luteal phase for the 5 farms

		Farm 1	Farm 2	Farm 3	Farm 4	Farm 5
Between 1st RaP4 measurement below 5 ng/mL and the previous observation (mean ± std)						
Normal phases	Hours	26.1 ± 20.9	26.4 ± 22.4	31.1 ± 26.9	26.1 ± 21.2	26.1 ± 22.2
	#Milkings	3.1 ± 2.6	3.7 ± 3.4	4.0 ± 3.7	2.9 ± 2.4	3.5 ± 3.0
Deviating follicular phase	Hours	45.5 ± 35.1	53.1 ± 34.7	44.0 ± 35.9	37.8 ± 33.1	64.5 ± 35.6
	#Milkings	5.3 ± 4.3	7.3 ± 4.6	5.5 ± 4.9	4.0 ± 3.7	8.3 ± 4.7
Deviating luteal phase	Hours	55.0 ± 35.1	55.6 ± 35.9	60.9 ± 35.1	57.0 ± 31.3	58.7 ± 25.4
	#Milkings	6.7 ± 4.6	7.3 ± 4.9	8.1 ± 5.0	6.2 ± 3.8	7.6 ± 3.7
Deviating follicular and luteal phase	Hours	69.7 ± 34.0	70.8 ± 30.2	78.3 ± 29.2	65.6 ± 33.2	72.8 ± 35.6
	#Milkings	5.0 ± 3.7	6.9 ± 3.8	5.1 ± 4.5	4.5 ± 2.5	5.7 ± 5.6
Between 1st SmP4 measurement below 5 ng/mL and the 1st RaP4 measurement below 5 ng/mL (mean ± std)						
Normal phases	Hours	17.2 ± 9.8	15.5 ± 7.6	15.7 ± 8.4	17.1 ± 9.0	15.9 ± 8.9
	#Milkings	1.9 ± 1.0	2.0 ± 0.9	1.9 ± 0.8	1.7 ± 0.7	2.0 ± 1.0
Deviating follicular phase	Hours	16.9 ± 11.2	12.2 ± 8.0	14.9 ± 9.6	18.5 ± 12.1	12.1 ± 8.2
	#Milkings	1.8 ± 1.1	1.5 ± 0.9	1.7 ± 1.0	1.7 ± 0.8	1.4 ± 0.9
Deviating luteal phase	Hours	15.8 ± 11.0	14.0 ± 9.2	12.7 ± 8.1	13.3 ± 7.7	12.3 ± 4.4
	#Milkings	1.7 ± 1.1	1.8 ± 1.1	1.4 ± 0.7	1.3 ± 0.6	1.6 ± 0.6
Deviating follicular and luteal phase	Hours	17.4 ± 11.8	14.7 ± 4.2	13.8 ± 6.8	18.8 ± 8.5	9.9 ± 7.0
	#Milkings	1.9 ± 1.4	1.9 ± 0.5	1.7 ± 0.9	2.0 ± 0.5	1.1 ± 0.7

Traits related to P4 sampling rate between "normal" estrus cycles

316 "Normal" estrus cycles were compared between the five farms for the two P4 sample frequency related
317 traits (Table 3). A Wilcoxon signed rank test revealed that the time between the 1st RaP4 below 5 ng/mL
318 and the previous observation was significantly smaller for farm 2 (*Md* = 16.7, *IQR* = (13.6-25.6), *n* =
319 402) and farm 5 (*Md* = 16.8, *IQR* = (13.4-25.7), *n* = 343) compared to farm 3 (*Md* = 18.2, *IQR* = (13.7-
320 37.2), *n* = 681), *p* = 0.016 and *p* = 0.017 respectively. There was no significant difference with farm
321 1 (*Md* = 17.5, *IQR* = (14.1-27.1), *n* = 1918), or farm 4 (*Md* = 18.4, *IQR* = (14.5-26.4), *n* = 384). Since
322 in general, it is a larger time period between the 1st RaP4 below 5 ng/mL and the previous observation,
323 only small significant differences were found between the 5 farms which can be the result of outliers
324 (but still smaller than 120 h) causing a larger time period. These outliers are possibly the result of factors
325 such as prior insemination resulting in embryonic mortality, or prior deviating P4 estrus profiles,
326 thereby causing the HN bio-model to work less well in a following "normal" cycle. Although these
327 "outliers" are interesting, they are too rare to explore (especially when putting them in perspective of
328 successful and unsuccessful inseminations later on). In addition, since the time between 1st RaP4 below
329 5 ng/mL and the previous is observation is larger compared to the time between 1st SmP4 below 5
330 ng/mL and the 1st RaP4 below 5 ng/mL (on average 27.0 ± 22.5 h versus 16.6 ± 9.2 h respectively), this
331 small significant difference is expected to have less of an impact on fertility success compared to time
332 between 1st SmP4 below 5 ng/mL and the 1st RaP4 below 5 ng/mL. For this second trait, a Wilcoxon
333 signed rank test revealed that the time was significantly more for farm 1 (*Md* = 14.8, *IQR* = (11.8-18.5),
334 *n* = 1911) compared to farm 2 (*Md* = 14.1, *IQR* = (11.5-16.9), *n* = 399), *p* = 0.005, farm 3 (*Md* = 13.7,
335 *IQR* = (11.5-17.2), *n* = 681), *p* = 0.000, and farm 5 (*Md* = 13.8, *IQR* = (11.6-16.6), *n* = 343), *p* =
336 0.005. Also, farm 4 (*Md* = 15.2, *IQR* = (11.7-19.1), *n* = 384) had significantly more time between 1st
337 SmP4 below 5 ng/mL and the 1st RaP4 below 5 ng/mL compared to farm 2, *p* = 0.004, farm 3, *p* =
338 0.001, and farm 5, *p* = 0.003. From this, it can be concluded that farm 1 and 4 have a longer time period
339 between 1st SmP4 below 5 ng/mL and the 1st RaP4 below 5 ng/mL. However, the number of milkings
340 within this same time period is on average similar for the 5 farms (Table 2: 1.9, 2.0, 1.9, 1.7 and 1.4
341 milkings for farm 1 to 5 respectively). This suggests that cows from farm 1 and 4 tend to have a lower
342 milking frequency, thereby causing a significant greater time between the 1st SmP4 below 5 ng/mL and

the 1st RaP4 below 5 ng/mL. This results in an increased delay between the true luteolysis and the estimated luteolysis (the 1st SmP4 below 5 ng/mL thereby generating a HN alarm), which could have serious effect the reliability of the estimated insemination window and subsequently on insemination success.

Farm 1	Farm 2	Farm 3	Farm 4	Farm 5
Time between 1st RaP4 below 5 ng/mL and the previous observation (median (IQR))				
17.5 (14.1-27.1)	16.7 (13.6-25.6)	18.2 (13.7-37.2)	18.4 (14.5-26.4)	16.8 (13.4-25.7)
Time between 1st SmP4 below 5 ng/mL and the 1st RaP4 below 5 ng/mL (median, IQR)				
14.8 (11.8-18.5)	14.1 (11.5-16.9)	13.7 (11.5-17.2)	15.2 (11.7-19.1)	13.8 (11.6-16.6)

Milk interval

The comparable number of milkings and increased time period between the 1st SmP4 below 5 ng/mL and the 1st RaP4 below 5 ng/mL for farm 1 and 4 compared to the other farms suggests that cows from farm 1 and 4 tend to have a lower milking frequency. This was investigated by calculating the MIs for cows from the 5 farms. The MIs are presented in Figure 3. A distinction is made between early and late lactation milkings since cows in late lactation are less milked compared to cows in early lactation. Comparisons were made between the MIs during early lactation between the farms and between the MIs during late lactation between the farms. The MIs in early and late lactation were normally distributed for each farm and a two-sided independent sample t-test revealed that the mean MI for both early and late lactation was significantly different between all 5 farms. It should be noted that even small differences in mean MI, such as the difference between farm 2 (early; late: 7.1 ± 2.2 h; 7.7 ± 2.4 h), farm 3 (7.6 ± 2.5 h; 8.6 ± 2.5 h) and farm 5 (7.8 ± 2.4 h; 8.2 ± 2.3 h), and the difference between farm 1 (8.3 ± 2.9 h; 9.4 ± 2.9 h) and farm 4 (8.4 ± 2.6 h; 9.9 ± 2.8 h), differed significantly from each other. This is explained by the large amount of data, i.e. every milk robot visit, that was used to calculate

these mean MIs. Remarkably, the MI for both early and late lactation was significantly the greatest for farm 1 and farm 4. These findings support the previously found increased time between the 1st SmP4 below 5 ng/mL and the 1st RaP4 below 5 ng/mL which suggested that cows from farm 1 and 4 have a lower milking frequency. It could be expected that the greater MI of farm 1 and 4 might also cause a significant larger time period between the 1st RaP4 and the previous observation for these farms compared to the other farms. Especially since there are on average 1.8 more milkings between the 1st RaP4 below 5 ng/mL and the previous observation compared to between the 1st SmP4 below 5 ng/mL and the 1st RaP4 below 5 ng/mL (Table 2). However, this was not the case. In addition, farm 1 and 4 have fewer milkings between the 1st RaP4 below 5 ng/mL and the previous observation (3.1 ± 2.6 and 2.9 ± 2.4) compared to farm 2 (3.7 ± 3.4), farm 3 (4.0 ± 3.7) and farm 5 (3.5 ± 3.0). This suggests that the HN bio-model tries to correct for the increased MI and therefore increased time between 1st RaP4 below 5 ng/mL and previous observation of farm 1 and 4. In contrast, after the 1st RaP4 drops below 5 ng/mL the next P4 sample is taken from the successive milking, the HN bio-model does not have time to correct for the greater MI for the time period between 1st SmP4 below 5 ng/mL and the 1st RaP4 below 5 ng/mL. Therefore, the MI has a greater significant effect on this time period for farm 1 and 4. Moreover, an increased MI could have an effect on the farm's average milk yield (ref). This is agreement with the 50-d, 100-d, 150-d and 305-d milk yield that was found to be lowest for farm 1

(305-d milk yield: 10,155 ± 1,849 kg) and 4 (10,144 ± 1,700 kg) compared to farm 2 (10,572 ± 1,924 kg), 3 (10,685 ± 2,056 kg) and 5 (10,918 ± 1,811 kg) (Table 1).

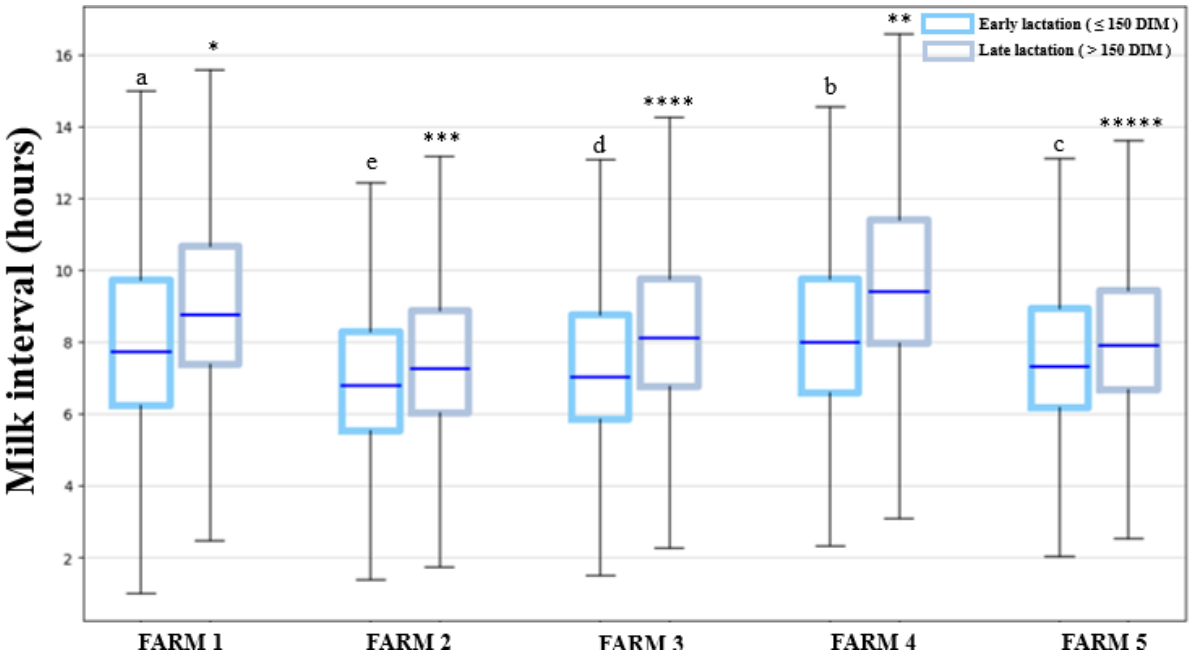

Figure 3. Milking interval in early (≤ 150 DIM) and late lactation (> 150 DIM).

Time between insemination and (1) the HN alarm and (2) the 1st RaP4 below 5 ng/mL

Optimal insemination timing should happen 24 to 12 hours before ovulation (Roelofs *et al*., 2005). In the study of Roelofs et al. (2006), however, it was shown that the milk P4 concentration dropped below a 5 ng/mL threshold at least 54 and at maximum 98 hours before ovulation. This means that an insemination should follow on average between 52 and 64 hours after a drop in the P4 concentration. HN advices to inseminate multiparous cows between 38 and 45 hours following an HN alarm as it takes the delay between the 1st RaP4 and 1st SmP4 measurement below 5 ng/mL into account. In Table 4, the time between insemination and the HN alarm and between insemination and the 1st RaP4 measurement below 5 ng/mL for multiparous cows are presented. A Wilcoxon signed rank test revealed that Farm 1 (*Md* = 48.4, *IQR* = (40.3-55.3), *n* = 645) inseminates significantly later on a HN alarm compared to Farm 2 (*Md* = 42.4, *IQR* = (36.2-49.8), *n* = 157), *p* = 0.000, Farm 4 (*Md* = 44.0, *IQR* = (37.8-50.4), *n* = 138), *p* = 0.000, and Farm 5 (*Md* = 41.7, *IQR* = (35.5-48.7), *n* = 63), *p* = 0.000. In addition, Farm 3 (*Md* = 48.4, *IQR* = (40.4-56.7), *n* = 242) also inseminates significantly later on a HN alarm compared to Farm 2, *p* = 0.000, Farm 4, *p* = 0.000, and Farm 5, *p* = 0.000. For the second trait, a Wilcoxon signed rank test revealed that Farm 1 (*Md* = 63.5, *IQR* = (55.4-72.7)) has significant more time between an

399 insemination and the 1st RaP4 below 5 ng/mL compared to Farm 2 (*Md* = 57.8, *IQR* = (49.9-66.6)), *p*
400 = 0.000, Farm 4 (*Md* = 60.5, *IQR* = (53.2-67.1)), *p* = 0.006 and Farm 5 (*Md* = 57.5, *IQR* = (49.2-
401 63.3)), *p* = 0.000. In addition, Farm 3 (*Md* = 62.8, *IQR* = (55.7-70.0)) has significant more time between
402 an insemination and the 1st RaP4 below 5 ng/mL compared to Farm 2, *p* = 0.001 and Farm 5, *p* = 0.000.
403 Moreover, Farm 4 has significant more time between an insemination and the 1st RaP4 below 5 ng/mL
404 compared to Farm 5, *p* = 0.015. For Farm 1 and Farm 3 this increased time can be explained by the fact
405 that these farms inseminate later on a HN alarm compared to the other 3 farms. On top of this, Farm 1
406 also has a significant larger MI compared to the other farms, thereby causing this farm to have the most
407 time between an insemination and the 1st RaP4 below 5 ng/mL. Farm 4 also had a significant larger MI
408 compared to the other farms, however, this farm tends to inseminate relatively fast on the HN alarms.
409 The larger MI explains that Farm 4 only has slightly more time between an insemination and the 1st
410 RaP4 below 5 ng/mL compared to Farm 5. Although the effect of insemination timing on conception
411 success has not been considered yet, based on the time between insemination and the 1st RaP4 below 5
412 ng/mL, which is largest for Farm 1 due to an increased MI and in general more time between
413 insemination and HN alarm, it could be suggested that this farm has the worst fertility characteristics.
414 This is in agreement with the farm fertility characteristics presented in Table 1. Farm 1 performs worst
415 for the 1st service conception rate, the number of inseminations per conception, the calving interval and
416 the number of days open compared to the other 4 farms. However, it should be noted that these results
417 are subjected to some bias since for 4 out of 5 farms the insemination timing was simplified to 12:00.

418 **Table 4**. Median and inter quartile range (IQR) for the two traits: time between insemination and (1) the HN alarm
419 and (2) the 1st RaP4 below 5 ng/mL for "normal" P4 estrus cycles. Total number of inseminations were compared
420

Farm 1	Farm 2	Farm 3	Farm 4	Farm 5
Time between insemination and the HN alarm (median (IQR))				
TOTAL	***TOTAL***	***TOTAL***	***TOTAL***	***TOTAL***
48.4 (40.3-55.3)	42.4 (36.2-49.8)	48.4 (40.4-56.7)	44.0 (37.8-50.4)	41.7 (35.5-48.7)

UNS	*SUC*	*UNS*	*SUC*	*UNS*	*SUC*	*UNS*	*SUC*	*UNS*	*SUC*
48.8 (41.2 - 55.4)	47.4 (38.0 - 54.4)	41.9 (34.0 - 49.5)	42.8 (37.2 - 50.0)	47.2 (40.0 - 54.9)	50.1 (40.8 - 57.3)	45.4 (38.1 - 50.8)	42.3 (37.6 - 48.8)	42.1 (34.9 - 48.7)	40.7 (36.2 - 47.2)
Time between insemination and the 1st RaP4 below 5 ng/mL (median IQR)									
TOTAL		*TOTAL*		*TOTAL*		*TOTAL*		*TOTAL*	
63.5 (55.4-72.7)		57.8 (49.9-66.6)		62.8 (55.7-70.0)		60.5 (53.2-67.1)		57.5 (49.2-63.3)	
UNS	*SUC*	*UNS*	*SUC*	*UNS*	*SUC*	*UNS*	*SUC*	*UNS*	*SUC*
64.0 (56.0-73.0)	61.6 (53.5-71.7)	55.8 (48.6 - 64.6)	60.0 (51.2 - 67.1)	62.1 (55.6 - 69.1)	63.2 (55.9 - 70.8)	62.4 (56.5-69.8)	58.4 (51.7-64.5)	57.8 (49.2 - 63.6)	52.0 (48.9 - 62.2)

Next, to understand the effect of the insemination reactivity of farmers for conception success, a distinction between unsuccessful and successful inseminations was made. In Table 4, the time between insemination and the HN alarm and the 1st RaP4 below 5 ng/mL is presented between unsuccessful and successful inseminations. Although there is a visible trend for Farm 1, 4 and 5 with successful inseminations having a less time between an insemination and the HN alarm, which is what you would expect especially for Farm 1 due to the increased MI and increased time between insemination and HN alarm and in some degree for Farm 4 due to the increased MI, no significant differences were found between unsuccessful and successful inseminations for the time between insemination and the HN alarm. In contrast, a Wilcoxon signed rank test revealed that the difference between an insemination and the 1st RaP4 below 5 ng/mL is significantly smaller for successful inseminations compared to unsuccessful inseminations for Farm 4 (*Md* = 58.4, *IQR* = (51.7-64.5), *n* = 62 vs *Md* = 62.4, *IQR* = (56.5-69.8), *n* = 76, *p* = 0.005)). For Farm 1, a smaller difference in time, however not significant, was found for successful inseminations (*Md* = 61.6, *IQR* = (53.5-71.7), *n* = 193) compared to unsuccessful ones (*Md* = 64.0, *IQR* = (56.0-73.0), *n* = 452). For the other farms, no significant difference was found between unsuccessful and successful inseminations. These results are again supported by the increased

437 significant time between HN alarm and the 1st RaP4 below 5 ng/mL, as a result of an increased MI, for
438 Farm 4 compared to Farm 2, 3 and 5. The not significant difference between insemination and HN alarm
439 and the significant difference between insemination and the 1st RaP4 below 5 ng/mL for unsuccessful
440 and successful inseminations for Farm 4, suggests that the increased time between HN alarm (i.e. the
441 1st SmP4 below 5 ng/mL) and the 1st RaP4 below 5 ng/mL has a significant effect on the insemination
442 outcome.
443
444 *Time between insemination and the HN alarm for high- and low P4 sampling rate*
445 The significant effect on conception success of (1) the time between HN alarm and the 1st RaP4 below
446 5 ng/mL and (2) the time between an insemination and the HN alarm, was explored by comparing these
447 two traits between unsuccessful and successful inseminations and between high- and low P4 sampling
448 rate cases. The distinction between high- and low P4 sampling rate cases was made for each farm
449 separately, thereby using the median time between HN alarm and the 1st RaP4 below 5 ng/mL as the
450 threshold (Table 3). In Table 5, the time between an insemination and the HN alarm are presented for
451 the unsuccessful and successful inseminations of high- and low P4 sampling rate cases. A Wilcoxon
452 signed rank test revealed that the time between an insemination and the HN alarm is significantly less
453 for successful inseminations (*Md* = 44.4, *IQR* = (36.1-51.9), *n* = 82) compared to unsuccessful ones
454 (*Md* = 48.6, *IQR* = (40.5-55.9), *n* = 220), *p* = 0.030, for the inseminations that followed a low P4
455 sampling rate on Farm 1. In addition, the time between an insemination and the HN alarm is
456 significantly less for the successful inseminations that followed a low P4 sampling rate compared to the
457 successful inseminations that followed a high P4 sampling rate (*Md* = 48.4, *IQR* = (40.6-56.2), *n* =
458 111), *p* = 0.022. Likewise, for Farm 4, the time between an insemination and the HN alarm is
459 significantly less for successful inseminations (*Md* = 41.0, *IQR* = (34.5-43.7), *n* = 19) compared to
460 unsuccessful ones (*Md* = 45.2, *IQR* = (37.9-50.7), *n* = 43), *p* = 0.048, for the inseminations that
461 followed a low P4 sampling rate. Also for Farm 4, the time between an insemination and the HN alarm
462 was significantly less for the successful inseminations that followed a low P4 sampling rate compared
463 to the successful inseminations that followed a high P4 sampling rate (*Md* = 44.0, *IQR* = (38.2-50.7),

464 n = 43), p = 0.045. This indicates that, for both Farm 1 and 4, when there is more time between a HN
465 alarm and the 1st RaP4 below 5 ng/mL (i.e. a low sampling rate case), successful inseminations have a
466 shorter time between an insemination and the HN alarm. Moreover, successful inseminations that
467 followed a low P4 sampling rate, are associated with less time between an insemination and the HN
468 alarm compared to successful inseminations that followed a high P4 sampling rate. These finding
469 suggest that when the time between the HN alarm and the 1st RaP4 is greater, farmers should inseminate
470 sooner. No significant differences were found between unsuccessful and successful inseminations of
471 the high P4 sampling rate cases for Farm 1 and 4, or between any other comparisons for Farm 2, 3 and
472 5. For the inseminations that followed a high sampling rate, other factors than sampling rate and
473 insemination reactivity are likely to influence conception success, such as oocyte quality or the cow's
474 health. For the other three farms, there is no significant difference for the unsuccessful and successful
475 inseminations that followed a low sampling rate. For these farms, there was no increased time between
476 the HN alarm and the 1st RaP4 below 5 ng/mL in general, which makes the high- and low P4 sampling
477 rate cases rather similar. Nevertheless, the same trend as for Farm 1 and 4, which entails that successful
478 inseminations that followed a low P4 sampling rate are associated with less time between an
479 insemination and the HN alarm compared to successful inseminations that followed a high P4 sampling
480 rate, is present for Farm 2 and 3. Only Farm 5 did not follow this trend. A possible explanation for this
481 could be the amount of available data, which is rather less for Farm 5 (only 19 and 44 successful and
482 unsuccessful inseminations in total) compared to the other farms.

483 **Table 5**. Median and inter quartile range (IQR) for the time between insemination and the HN alarm. A
484 distinction is made between high- and low P4 sampling rate cases based on the time between HN alarm and the
485

CONCLUSION

The reliability of the usage of P4 as a gold standard for evaluating fertility success was studied. Based on historical data collected at 5 different farms, we analysed different traits related to P4 sample frequency of the HN system, milking frequency and insemination timing of the farmer and looked at their effect on success of conception across the 5 farms. We found that, in farms with worse fertility performance, the time between the HN alarm and the true luteolysis (i.e. the 1st RaP4 below 5 ng/mL) was greater and less consistent. This increase in time is explained by an increased MI which was found for the same farms. When the time between the HN alarm and the 1st RaP4 below 5 ng/mL is used to distinct between high- and low P4 sampling rate cases, successful inseminations that followed a low P4 sampling rate had significantly less time between an insemination and the preceding HN alarm compared to unsuccessful inseminations following a low P4 sampling rate. Moreover, successful inseminations that followed a low P4 sampling rate, are associated with less time between an insemination and the preceding HN alarm compared to successful inseminations that followed a high P4 sampling rate. These finding confirm that when the time between the HN alarm and the 1st RaP4 is greater, farmers should inseminate sooner. From this analysis, we can make a better selection of the cycles in which insemination time was correct, and thus depended on other factors. This is a first step

	High P4 sampling rate		**Low P4 sampling rate**	
	UNS (median (IQR))	***SUC*** (median (IQR))	***UNS*** (median (IQR))	***SUC*** (median (IQR))
Farm 1	49.3 (41.7-54.6)	48.4 (40.6-56.2)	48.6 (40.5-55.9)	44.4 (36.1-51.9)
Farm 2	39.9 (32.6-48.0)	43.7 (39.0-50.7)	42.9 (36.0-49.8)	42.1 (36.3-49.9)
Farm 3	50.4 (41.6-57.4)	50.6 (43.5-57.2)	44.2 (39.7-49.7)	46.8 (40.8-57.5)
Farm 4	46.7 (38.2-51.4)	44.0 (38.2-50.7)	45.2 (37.9-50.7)	41.0 (34.5-43.7)
Farm 5	43.9 (36.6-53.1)	38.1 (36.0-43.9)	40.2 (34.9-44.2)	43.1 (38.1-49.4)

504 towards improved understanding of fertility performance on farm, and as such, of a more sustainable
505 dairy sector.
506
507 **REFERENCES**
508
509 Adriaens, I., Huybrechts, T., Geerinckx, K., Daems, D., Lammertyn, J., De Ketelaere, B., Saeys, W.,
510 Aernouts, B. (2017). Mathematical characterization of the milk progesterone profile as a leg up to
511 individualized monitoring of reproduction status in dairy cows. *Theriogenology*, *103*, 44-51.
512
513 Adriaens, I., Martin, O., Saeys, W., De Ketelaere, B., Friggens, N.C., Aernouts, B. (2019). Validation
514 of a novel milk progesterone-based tool to monitor luteolysis in dairy cows: Timing of the alerts and
515 robustness against missing values. *Journal of dairy science*, *102*(12), 11491-11503.
516
517 Adriaens, I., Saeys, W. (sup.), De Ketelaere, B. (cosup.), Piepers, S. (cosup.) (2018a). Data-based
518 monitoring of dairy cows-milk progesterone as a mirror of fertility. KU Leuven.
519
520 Adriaens, I., Saeys, W., Lamberigts, C., Berth, M., Geerinckx, K., Leroy, J., De Ketelaere, B., Aernouts,
521 B. (2018b). Short communication: Sensitivity of estrus alerts and relationship with timing of the
522 luteinizing hormone surge. *Journal of dairy science*, *102*(2), 1775-1779.
523
524 Crowe, M.A. (2011). Reproduction, Events and Management, Estrous Cycles: Characteristics.
525 Encyclopedia of Dairy Sciences, 428-433
526
527 Friggens, N.C., Bjerring, M., Ridder, C., Højsgaard, S., Larsen, T. (2008). Improved detection of
528 reproductive status in dairy cows using milk progesterone measurements. *Reproduction in Domestic*
529 *Animals*, *43*, 113-121.
530

531 Friggens, N.C., & Chagunda, M.G. (2005). Prediction of the reproductive status of cattle on the basis
532 of milk progesterone measures: model description. *Theriogenology*, *64*(1), 155-190.
533
534 Galvão, K.N., Federico, P., De Vries, A., Schuenemann, G.M. (2013). Economic comparison of
535 reproductive programs for dairy herds using estrus detection, timed artificial insemination, or a
536 combination. *Journal of dairy science*, *96*(4), 2681-2693.
537
538 Hansen, P.J. (2011). Challenges to fertility in dairy cattle: from ovulation to the fetal stage of
539 pregnancy. *Rev Brasil Reprod Anim*, *35*, 229-238.
540
541 Krpálková, L., O'Mahony, N., Carvalho, A., Campbell, S., Walsh, J. (2020). Evaluating the economic
542 profit of reproductive performance through the integration of a dynamic programming model on a
543 specific dairy farm. *Czech Journal of Animal Science*, *65*(4), 124-134.
544
545 Leroy, J. L. M.R., Opsomer, G., Van Soom, A., Goovaerts, I.G.F., & Bols, P.E.J. (2008). Reduced
546 fertility in High-yielding dairy cows: Are the oocyte and embryo in danger? Part I the importance of
547 negative energy balance and altered corpus luteum function to the reduction of oocyte and embryo
548 quality in High-yielding dairy cows. *Reproduction in domestic animals*, *43*(5), 612-622.
549
550 Leroy, J. L.M.R., Rizos, D., Sturmey, R., Bossaert, P., Gutierrez-Adan, A., Van Hoeck, V., Valckx, S.,
551 Bols, P. E. J. (2011). Intrafollicular conditions as a major link between maternal metabolism and oocyte
552 quality: a focus on dairy cow fertility. *Reproduction, Fertility and Development*, *24*(1), 1-12.
553
554 Løvendahl, P., & Chagunda, M.G. (2010). On the use of physical activity monitoring for estrus
555 detection in dairy cows. *Journal of Dairy Science*, *93*(1), 249-259.
556

Ranasinghe, R.M.S.B.K., Nakao, T., Yamada, K., Koike, K., Hayashi, A., Dematawewa, C.M.B. (2011). Characteristics of prolonged luteal phase identified by milk progesterone concentrations and its effects on reproductive performance in Holstein cows. *Journal of Dairy Science*, *94*(1), 116-127.

Roelofs, J.B., Van Eerdenburg, F.J.C.M., Hazeleger, W., Soede, N.M., & Kemp, B. (2006). Relationship between progesterone concentrations in milk and blood and time of ovulation in dairy cattle. *Animal reproduction science*, *91*(3-4), 337-343.

Roelofs, J.B., Van Eerdenburg, F.J.CM., Soede, N. M., Kemp, B. (2005). Various behavioral signs of estrous and their relationship with time of ovulation in dairy cattle. *Theriogenology*, *63*(5), 1366-1377.