Use of Big Data to Monitor Herd Health

Ryan Pralle and Heather White

University of Wisconsin-Madison, Madison, WI, USA
Email: heather.white@wisc.edu

- Take Home Messages

- Increased collection of data on farms from the cow, pen, and farm level may provide opportunities to improve management of herd and animal health, especially during the transition to lactation period.
- Data-based diagnostics may be a valuable and less invasive way to improve management and disease detection systems compared with blood-based diagnostics.
- Broad implementation of data-based diagnostics can allow for the generation of large-scale field data to identify patterns of risk or success.
- Improved large datasets for research may uncover novel genomic risk factors or metabolic pathways of interest.
- Integration of data streams available on farm (milk production and components, genomics, etc.) may improve the utility of management resources.

- Introduction

Recent and ongoing advances in dairy cattle management, nutrition, genetics, and reproduction have led to improvements in our understanding and knowledge of dairy cattle biology. Additionally, improvements in data availability and timeliness of data from farms on an individual cow, herd, and farm system level have increased. For example, the routine analysis of milk samples started as a way to monitor milk fat and protein but has progressed to other complex traits and monitoring of herd health indicators. Other data are now available to farms from numerous sources that range from feeding systems, rumination sensors, activity monitors, robotic milking systems, and weather stations to name a few (Wathes et al., 2008; Wolfert et al., 2017). Furthermore, management systems incorporating advanced technology can be accessed by off-farm management team members and send real-time alerts.

There is no doubt that there is an immense amount of information nested within data that is available from farms; however, there is also a need to sort, validate, and collate these data to format them in a way that is useful to the producer. For those that focus on animal health—producers, veterinarians, and researchers—a clear potential for this data is the potential use for detection of health events. Health diagnostics serve many different roles and range from confirming a clinical case of a suspected disorder to a screening for sub-clinical disorders that have non-visible symptoms. These tools can also be at a herd-level, which provides information about how well we are doing at preventing or managing a disorder, or at an individual cow-level, which can give information regarding the necessity of treatment. These tools can be integrated into management programs that allow them to complement each other.

- What Makes a Good Diagnostic Tool?

Before we dive into progress of diagnostic tools based on big data, we must first understand what makes a good diagnostic tool and what special or additional considerations must be made for 'big data' tools. Generally speaking, accuracy is used to describe how well the diagnostic or model fits the actual data. More specifically, we often need to know the ability of the diagnostic tool to appropriately identify a case, which we can calculate as sensitivity and specificity. Sensitivity is the ability of the diagnostic to identify
positive cases as positive. Specificity is the ability of the diagnostic to correctly identify a negative case as a negative. In general, our goal for these are 90%, but it is important to realize that when we optimize to improve one metric (i.e., sensitivity), the other metric (i.e., specificity) tends to decrease. In research, we try to maximize sensitivity and specificity simultaneously when developing a diagnostic tool. In reality, which metric we want to maximize also depends on the consequences of misdiagnosis. For example, low specificity results in more false positives (cows that do not have the disorder but are identified as positive). These cases could inappropriately receive pharmaceutical treatment, so in a scenario such as mastitis where a treatment entails antibiotic use and discarded milk, this can be a costly misdiagnosis. Conversely, low sensitivity results in more false negatives (cows that do have the disorder but are not identified) and could result in the disorder getting more severe before identification and treatment.

Additional considerations are needed when we are building diagnostics on ‘big data’. Data generation needs sufficient controls and validation and should come from representative data sets. Diagnostic tools should be built on data sets that are sufficiently large and robust to remain accurate when applied to novel data. To this last point, prediction models or diagnostics must be validated on data sets that the tools weren’t built on to ensure that the models can be applied to other data sets.

- **Sub-Clinical Ketosis Detection as an Example of Data Integration**

Of particular interest to us and others is monitoring animal health during the transition to lactation period. The transition to lactation period is a time of increased risk for metabolic disorders such as ketosis (hyperketonemia), milk fever (hypocalcemia), and fatty liver. Commonly categorized as sub-clinical ketosis (SCK) or clinical ketosis, hyperketonemia is defined as elevated blood ketone body concentrations. The reference test for SCK diagnosis is the quantification of blood β-hydroxybutyrate (BHB) concentration via enzymatic assay (Iwersen et al., 2009; Mcart et al., 2013). The establishment of current blood BHB thresholds based on the increased risk of unfavourable animal health and performance outcomes has been thoroughly reviewed (Mcart et al., 2013; Overton et al., 2017), with thresholds of BHB ≥ 1.2 mmol/L or BHB ≥ 1.4 mmol/L being the most frequently used.

Detection of SCK on farm has largely moved to using hand-held BHB meters that can be used cow side, such as the Precision Xtra and PortaCheck meters (Iersen et al., 2009; Sailer et al., 2018). Diagnostic protocols may involve screening cows between 4 and 18 days in milk (DIM) either twice within the first week of the postpartum period, or once a week, but in either scenario the goal is to test each cow twice. These detection protocols are based on our understanding of the risk period and causes for SCK and are customizable from farm to farm. Blood BHB testing is costly and labour-intensive, making it difficult to apply on farm (Denis-Robichaud et al., 2014; McArt et al., 2014). Furthermore, there are real on-farm challenges with dedicating skilled labour and resources to weekly blood analysis, and thus, many farms struggle to implement a blood-based detection protocol consistently.

If one challenge of appropriate SCK detection is the intensiveness of blood collection, the obvious question is: what readily available data or samples could allow for SCK detection? One potential answer involves closely examining milk samples using techniques that are non-invasive and can be incorporated into milking routines, and realizing that concentrations of milk ketone bodies are correlated to blood concentrations (Marstorp et al., 1983; Andersson, 1984; Denis-Robichaud et al., 2014). Additionally, many farms routinely collect milk samples for analysis as a part of dairy herd improvement. Early uses of milk data to monitor postpartum herd-health were through use of the milk fat:protein as an indicator of SCK. At a herd-level, the proportion of fresh cows above a ratio threshold was used as an indicator but this is not tightly correlated to blood BHB (Figure 1). Despite this tool not being as accurate as others available now, it represents our progress in using milk data to predict or detect postpartum health events.
Although milk analysis historically only provided information on major milk components (fat, protein, lactose) or somatic cell count, advances in milk analysis have provided recent opportunities to learn more from these samples. Many milk analysis labs now incorporate Fourier transform infrared (FTIR) spectrometry (Rutten et al., 2009; 2011) that provides a practical method of predicting milk ketone body concentrations. This method is currently not as accurate as laboratory analysis, but it can be implemented on a larger scale.

Predicted concentrations of milk acetone or BHB from FTIR are correlated ($r = 0.80$) to chemically analyzed milk concentrations (de Roos et al., 2007); however, the prediction of milk acetone or BHB alone may not be sensitive or specific enough to detect individual cases of SCK. Across a range of cutoffs determined for milk ketone body concentrations for SCK diagnosis, milk BHB and acetone alone lack sufficient accuracy, sensitivity, and specificity for individual-cow diagnosis (Enjalbert et al. 2001; van Knegsel et al., 2010; van der Drift et al., 2012b). Predicting blood BHB from FTIR-predicted milk BHB ($\text{BHB} \geq 1.2 \text{ mmol/L as the diagnostic threshold}$) resulted in an overall sensitivity and specificity of 81% and 92%, respectively (Renaud et al., 2018). Similar sensitivity and specificity were achieved when FTIR-predicted BHB and acetone were analyzed in primiparous, but not multiparous, Holstein and Jersey cows (Chandler et al., 2018).

Based on understanding risk factors for SCK onset, we may find benefits for incorporating some of these factors into the prediction models, especially since these data are often collected during routine milk analysis and would be readily available. Incorporation of performance variables, including dry period length, gestation length, and lactation number, into milk analysis results (fat, protein, FTIR-predicted acetone and BHB) improved accuracy of models to predict SCK, but still lacked the sensitivity to provide an individual cow diagnostic tool (Chandler et al., 2018). Regardless of the lack of sensitivity of these models compared with enzymatic blood BHB quantification, there is a role for these models in herd-level diagnostics and monitoring (Denis-Robichaud et al., 2014; Chandler et al., 2018). Routine (i.e., monthly) implementation of these herd-level SCK predictions can aid farms in monitoring prevalence and identifying patterns and farm-specific risk factors. This is key since we know monthly prevalence of SCK changes across the year within a farm. For example, based on the economics of blood diagnostics and farm goals, a farm may only be able to justify postpartum blood sampling when prevalence is greater than 10%. If the farm’s current prevalence is less than 10% and the farm stops sampling, the farm may miss an increase in prevalence to 30% because it doesn’t have an indicator of the change. Figure 2 shows a real example of SCK prevalence over 12 months and demonstrates that while there are several months the farm may not want to do postpartum blood sampling, it could easily miss an increase in SCK events if it didn’t have a monitoring protocol or tool in place. This is one role of a milk-based herd-level diagnostic.
tool that can give a monthly indicator of prevalence to monitor these changes. There can also be additional benefits of observing herd health consequences from intentional or unintentional management and environment changes such as changes in forage, new or changing personnel, or implementation of a new protocol related to prepartum or postpartum cow management or nutrition.

Figure 2. Example of real variance in sub-clinical ketosis (SCK) prevalence on a dairy farm over a 12-month period. Use of a herd-level milk-based diagnostic tool can give the farm an indicator of what the monthly prevalence is, whether the farm is doing blood-based diagnostics or not. Using this information, the farm could choose which months to do blood-based diagnostics based on economics, labour availability, and farm goals.

Ongoing attempts to increase reliability of model predictions have continued and have involved the use of more advanced statistical modelling techniques, larger sample sets, and use of raw FTIR data. Using a dataset of more than 3,600 samples, artificial neural network models based on milk analysis (i.e., milk fat, protein, FTIR-predicted acetone and BHB) and raw FTIR data resulted in improved SCK prediction models compared with models using either FTIR data or milk analysis data alone, or models using other statistical methods (Pralle et al., 2018). Despite failing to reach the diagnostic accuracy of the cowside enzymatic blood tests such as the Precision Xtra and PortaCheck meters (Iwersen et al., 2009; Sailer et al., 2018), these models are considered to have ‘very good’ diagnostic capacity based on general thresholds of diagnostic accuracy (Šimundić, 2009). When we recall that the goal of using FTIR milk analysis to diagnose SCK is to reduce the cost and labour necessary to employ blood-based diagnostics on farm, we may be willing to accept slightly lower diagnostic capacity in exchange for the reduced intensity and cost of FTIR-based diagnostics. A positive case of SCK is likely to be treated with oral propylene glycol, which represents a low-cost, no milk withdrawal treatment (McArt et al., 2012; 2014), resulting in minimal consequence to treating a false-positive case. Given this, future work could also optimize the prediction models in favour of sensitivity to increase the likelihood of identifying all positive cases, even if some negative cases are identified as false positives (decreased model specificity). These ‘real-world’ adjustments to research models could maintain cost-effectiveness and promote animal welfare and should be further considered. Additionally, as datasets from milk-based diagnostics increase in size, we can use this information to identify epidemiological type patterns regarding SCK. Growing data sets from milk analysis labs will be able to give information about SCK risk factors, patterns over time (within and across years), and patterns that are unique to region or management system. These data will never indicate causes, but if explored could provide insight that could guide future research into nutrition or management influences that might decrease SCK prevalence.
Integration of Genetic and Genomics into Ketosis Management and Detection

Examining genetic selection for SCK susceptibility has predominately relied on voluntary records provided by dairy producers. Each farm has its own system of detection and recording for health events including SCK and therefore, this database includes innate variance. If a farm is actively implementing a blood detection method, it may appear to have more SCK recorded compared with a farm that is only recording clinical cases that are obvious. This can result in misclassification of SCK cases as controls (or non-events) when they are actually just un-diagnosed positive cases; this presents an issue in using voluntary records, especially when considering subclinical cases and the variation in SCK diagnosis criteria (Oetzel, 2004; Pryce et al., 2016). Regardless, ranges of heritability can be learned from this data. Threshold model heritability estimates have ranged from 0.02 to 0.17 in Holstein cows (Kadarmideen et al., 2000; Zwald et al., 2004; Gaddis et al., 2014; Klein et al., 2019). Weigel et al. (2017) explored the heritability of SCK based on intensive sampling of blood BHB concentration (four timepoints between 5 and 18 DIM). The estimated heritability for the diagnosis of SCK (positive or negative diagnosis) was 0.07, which was within the range reported for voluntary records.

Incorporation of SCK susceptibility genetic and genomic evaluations into management through animal breeding programs is intuitive. A potential concern for directly breeding against SCK susceptibility is the unintentional elimination of compensatory metabolism (i.e., ketogenesis), at least with respect to the subclinical phenotype. An alternative, unexplored use of genomics is as a tool for precision health management. Genome-guided management of SCK would be like personalized medicine, where estimated breeding values for SCK alone or coupled with other on-farm data streams, such as test-day prediction models (Chandler et al., 2018; Pralle et al., 2018), could be used to stratify cows into risk groups for group-specific management (Weigel et al., 2017).

Opportunities to Further Integrate Data-Based Diagnostics On-Farm

Although the collective body of research has generated diagnostic tools, on-farm integration and implementation remains a challenge. A few key challenges present hurdles and deserve our vested interest. One key challenge for consideration from the farm perspective is the nature of ‘monthly’ DHI milk testing on farm. Historically, DHI milk sampling has been done about every 4 weeks on privately-owned dairy farms (Weigel et al., 2017). Based on the at-risk period and duration for SCK, a farm would only test half of the fresh cows within the peak risk period (3 to 18 DIM; McArt et al. 2012b) and those cows would only be tested once. The two intuitive ways to address this challenge are to: 1) produce equations based on in-line measurement systems and 2) implement weekly milk testing of fresh cows. Current in-line systems include Herd Navigator (DeLaval International AB, Tumba, Sweden) and AfiliLab (Afimilk, Kibbutz Afikim, Israel); however, only the former provides milk BHB predictions for monitoring SCK (Blom et al., 2015). For farms that do not have in-line measurement systems, the second option provides a viable option. In this system, all early postpartum cows would be milk sampled weekly, which results in fresh cows being sampled twice during the at-risk period. Use of prediction models on these samples would allow for either treatment of cows predicted positive or generation of a list of cows that warrant blood BHB testing. Beyond detection of SCK, these weekly milk samples could have additional benefits. Reporting back SCC, or differential SCC, could provide valuable information in this postpartum period to help identify cases of mastitis earlier (Fourdraine et al., 2019). As the technology continues to improve (e.g., differential SCC), it may also be a means to flag cows with increased inflammatory markers. Furthermore, if a group of cows are being milk tested weekly, it is not beyond reason that other sub-groups would be selectively milk sampled that day, which opens the door further for using milk FTIR to determine energy balance (McParland et al., 2015; Grelet et al., 2016), evaluate likelihood of conception (Ho et al., 2019), test for pregnancy in mid-lactation cows (Lainé et al., 2017; Toledo-Alvarado et al., 2018), or monitor cows within treatment pens. Milk analysis does not occur without cost, but analysis from postpartum cows may provide more valuable information than from cows in late lactation, allowing us to envision that DHI programs of the future may shift to more frequent testing of cows in periods of interest.
Acknowledgements

We would like to acknowledge and thank the collaborators of this research (namely Drs. Tawny Chandler, Robert Fourdraine, Gary Oetzel, and Kent Weigel), the countless undergraduate and graduate students, farm and research staff, and collaborators that make our research contributions to this field possible. Gratitude is extended to AgSource Coop. for their extensive collaboration. We would like to recognize funding for the projects completed by the H. M. White research group referenced here: USDA AFRI Critical Agriculture Research and Extension (CARE; 2015-67028-23572); USDA Hatch (WIS01877 and WISC01878) from the Wisconsin Agricultural Experiment Station (Madison, WI); Cooperative Research Program for Agriculture Science and Technology Development (Project No. PJ012078) Rural Development Administration, Jeonju-si, Jeollabuk-do, Republic of Korea; AgSource Cooperative Services (research funding support and collaborative efforts; Verona, WI); Purina Animal Nutrition, LLC (St. Louis, MO) graduate student fellowships; VitaPlus Corporation (Madison, WI) graduate student fellowship; and the Wisconsin Alumni Research Foundation (Madison, WI).

References


COMFORT SOLUTIONS TO HELP MEET YOUR DAIRY PERFORMANCE GOALS

PROMAT’S INDUSTRY LEADING COW MATTRESS SOLUTIONS

PASTURE MAT™ enjoyed by over 3 million cows worldwide

COMFORT MAT™ the undisputed best foam mattress on the market

PACK MAT™ all the benefits of deep bedding using up to 70% less sand

RUBBER FLOORING SOLUTIONS

Natural & Recycled Interlocking & Straight Cut Mats

Rolled Alley Flooring

Dairy Parlor Anti-Fatigue Flooring

COWCOON - THE HYBRID COW STALL

Utilizes a flexible polyethylene tube across the bottom portion of the stall loop that acts as a shock absorber

Proud Sponsor of the Western Canada Dairy Seminar

1.888.337.6265 | sales@promatinc.com
www.Promatinc.com  facebook.com/promatinc  @Promat_inc

The Leader in Animal Comfort
The AM 1000s was designed to solve the challenges farmers face when feeding cattle in extreme conditions. It is solidly built with a 300 HP engine, high ground clearance, large off road tires, crab steering for tight spaces, brand name proven hydraulic components, a spacious cab and hydraulic suspension. While other machines are built for ideal conditions, this machine is truly ... *Built For Here™*.

The *Hired Hand™* self loading arm eliminates the need for multiple operators and machines. It is designed to peel the silage face smooth and solid which results in less spoilage.

The 150 HP variable speed milling head is simple and easy to service with all drive components being external to the milling head. Ribbed rubber belting quickly conveys the ingredients into the mixing tub.