Norwegian University of Life Sciences
Faculty of Biosciences
Department of Animal and Aquacultural Sciences

Philosophiae Doctor (PhD)
Thesis 2017:65

# Multivariate Analyses of Milk Infrared Spectra and Predictions from it in Dairy Cattle Populations

## Fleireigenskapsanalyse av infraraude mjølkespekter og prediksjonar frå dei i mjølkekupopulasjonar

Tesfaye Kebede Belay

# Multivariate Analyses of Milk Infrared Spectra and Predictions from it in Dairy Cattle Populations

Fleireigenskapsanalyse av infraraude mjølkespekter og prediksjonar frå dei i mjølkekupopulasjonar

Philosophiae Doctor (PhD) Thesis

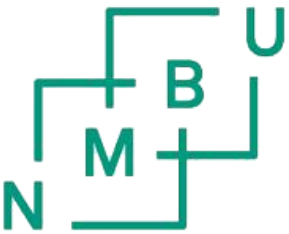
Tesfaye Kebede Belay

Department of Animal and Aquacultural Sciences

Faculty of Biosciences

Norwegian University of Life Sciences

Ås 2017




Thesis number 2017: 65

ISSN: 1894-6402

ISBN: 978-82-575-1462-4

**PhD Supervisors**

**Assoc. Prof. Tormod Ådnøy**
Department of Animal and Aquacultural Sciences
Norwegian University of Life Sciences
P. O. Box 5003, N-1432 Ås
Norway

**Dr. Morten Svendsen**
GENO Breeding and A.I Association
P. O. Box 5025, N-1432 Ås
Norway

**Dr. Binyam Sime Dagnachew**
Department of Animal and Aquacultural Sciences
Norwegian University of Life Sciences
P. O. Box 5003, N-1432 Ås
Norway

**Prof. Achim Kohler**
Faculty of Sciences and Technology
Norwegian University of Life Sciences
P. O. Box 5003, N-1432 Ås
Norway

**Assoc. Prof. Erling Thuen**
Department of Animal and Aquacultural Sciences
Norwegian University of Life Sciences
P. O. Box 5003, N-1432 Ås
Norway

**PhD Evaluation Committee**

**Prof. D Just Jensen**
Center for Quantitative Genetics and Genomics
Department of Molecular Biology and Genetics, AU Foulum, Aarhus University
P.O. Box 50, 8830 Tjele
Denmark

**Prof. Donagh Berry**
Animal & Bioscience Research Department
Teagasc, Moorepark, Fermoy, Co. Cork
Ireland

**Dr. Bjørg Heringstad**
Department of Animal and Aquacultural Sciences
Norwegian University of Life Sciences
P. O. Box 5003, N-1432 Ås
Norway

**ACKNOWLEDGMENTS**

Data used for this thesis work was from TINE SA, the Norwegian Dairy Association, and Polish Federation of Cattle Breeders and Dairy Farmers in Poland. They are acknowledged for allowing me to use their data.

I want to express my sincere appreciation and gratitude to my supervisor Assoc. Prof. Tormod Ådnøy for your esteemed guidance and advice from the beginning up to the end of the PhD study. I thank you for your follow-ups, invaluable comments and suggestions. I also thank you for being their when I needed you and for translating the summary of the thesis. I greatly indebted to you since without your encouragement, moral and technical support and professional expertise, the completion of this work would have not been possible.

I would like to extend my special thanks to my co-supervisors Dr. Morten Svendsen and Dr. Binyam Sime Dagnachew for their guidance, encouragement, valuable comments and discussions and for the unreserved support whenever needed. I also gratefully thanks my co-supervisor Prof. Achim Kohler for your support in multivariate calibration and spectral pre-processing and for your interesting discussions and suggestions. I would like also to thank my co-supervisor Assoc. Prof. Erling Thuen for your encouragements.

I am also thankful to Prof. Z.M. Kowalski for your hospitality when visiting you in University of Agriculture in Krakow, Poland. I am also very grateful for all your inputs, thoughtful discussions and co-authorship. I also thank Mr. Artur and Prof. Wojciech Jagusiak for helping me to understand the Polish data and construct the pedigree file, respectively. I am thankful to Mr. K. Słoniewski for transferring the data and providing explanation on the data when needed. I also acknowledge Dr.Valeria Tafintseva for your help in model calibration. I would like also to extend my special gratitude to Dr. Solomon Antwi Boison for writing simulation program, and for your encouragement, support and co-authorship.

My sincere appreciation goes to Mr. Teshome Dagne Mulugeta for your unreserved help on how to use CIGENE server and to work on Linux/Unix environment as well as for your encouragement and advice throughout the course of my study. Thank you for making my stay as enjoyable as possible and for your support whenever needed. I would like to covey my special thanks to Oscar Okechukwu Michael Iheshiulor, Gebreyohans Gebregiwergis and Kahsay Gebremariam Nirea for your technical support and encouragement.

I am grateful to the encouragement and support of the Breeding and Quantitative Genetics group at the Department of Animal and Aquaculture Sciences (IHA). It has been great to work and study with you. Thank you for your helpful comments and suggestions. I would also like to express my gratitude to all the administrative staff at IHA for their assistance whenever I needed you. Thank you made my stay easier!

I wish to express my appreciation to my father Mr. Kebede Belay and my mother Mrs. Tejitu Oljira for their precious care from childhood, advice, support and encouragement, and for their self-sacrifices in allowing me to pursue my education. I also appreciate encouragement and support from my late mother in-law Mrs. Bote Akassa. You wished me best luck when I leaved home for my study, but you have gone forever before I complete it. The God gave, and the God has taken away. I would also like to extend my gratitude to my brothers, sisters and relatives for your assistance and encouragements.

A special thanks to my beautiful wife Almaz Mengistu. Thank you for your love, motivation, understanding and support. You have carried the entire burden in managing our children. You have been my pillar behind this achievement. I am also greatly indebted to our lovely son Firaol and daughters Koket and Yomift for their love and patience

Finally, I would like to thank all those people who have made this thesis possible. I express my apology for that I could not mention your name one by one.

God is great!

*Ås, June 2017*

*Tesfaye Kebede Belay*

**DEDICATION**

*To the innocent Oromo people and other Ethiopians who have been killed while struggling for freedom and justice in the last few years (March 2014 to June 2017).*

**TABLE OF CONTENTS**

**SUMMARY**

Belay, T. K. 2017. ***Multivariate Analyses of Milk Infrared Spectra and Predictions from it in Dairy Cattle Populations.*** Norwegian University of Life Sciences, Philosophiae Doctor Thesis, 2017: 65, ISSN: 1894-6402, ISBN: 978-82-575-1462-4.

Fourier-transform mid-infrared (FT-MIR) spectra of milk is one of the multivariate information routinely recorded by many milk-recording organizations in the world. Use of such information is becoming central to research in dairy sciences. This is because the FT-MIR spectra and phenotypes predicted from those spectra could be useful for better estimation of parameters related to breeding, feeding and health. The focus of this PhD study has been to verify methods for exploiting milk FT-MIR spectral information for prediction of breeding values and phenotypes.

In paper I, we compared the conventional single-trait (ST) and multi-trait (MT) animal models for genetic evaluations using test-day data from Norwegian milk recording. Results show that estimates of heritability were very similar in both analyses. The MT analyses improved accuracies of estimated breeding values (EBV) for cows (e.g., improvement from 2.5 % for milk yield to 9.83 % protein yield) and sires with < 50 daughters (e.g., 3.25% improvement for protein yield), but they were similar for sires with >50 daughters. Estimated genetic trends were slightly higher under MT for protein and fat contents, and for fat yield, but slightly lower for the remaining traits. With comparison of MT with ST rank correlations for EBV, sires were less re-ranked than cows. In paper II, we compared two prediction approaches using mixed models for their ability to predict blood β-hydroxybutyrate (BHB) from milk FT-MIR spectra in Polish cows. One approach (indirect prediction – IP) transforms spectra to a single-trait before genetic analysis, while the other (direct prediction – DP) uses a multi-trait mixed model on (dimension reduced) spectral variables to obtain multi-trait predictions of random effects. Both approaches involve genetic analyses for ultimate phenotypic and EBV prediction. Performances of the IP and DP approaches were similar for phenotypic prediction of blood BHB. A slightly more accurate prediction of BHB was found when univariate variance structure (IP) was used compared to when multivariate covariance structures were used. Accuracies ($R^2$) were low, 0.28-0.30 for the IP, and 0.26-0.30 for the DP approach. For partial least square (PLS) regression with untransformed blood BHB, the $R^2$ was 0.29 to 0.37. In paper III, an established connection between milk FT-MIR spectra and blood BHB

in Polish dairy cattle was used to identify Norwegian Red cows treated for ketosis. Genetic parameters for FT-MIR predicted blood BHB and for clinical ketosis (KET) were estimated. Genetic associations of predicted blood BHB with KET and milk production traits were also examined. Heritability estimates for predicted blood BHB at different stages of lactation were moderate, ranging from 0.250 to 0.365. Genetic correlations between BHB traits were higher for adjacent lactation stages. Predicted blood BHB at 11-30 DIM was moderately genetically correlated with KET (0.469) and milk traits (ranged from -0.367 with protein content to 0.277 with milk yield). In paper IV, we simulated three traits and compared the IP and DP approaches for predictions of EBV and phenotypes under different genetic (low: 0.10 to high: 0.90) and residual (zero to high: ± 0.90) correlation scenarios of the traits. Relationships between performances of the two approaches and the accuracy of calibration equations were evaluated. Moreover, the effect of using different PLS regression coefficients estimated from simulated phenotypes ($\beta_p$), true breeding values ($\beta_g$) and residuals ($\beta_r$) on performance of the two approaches were evaluated. Aaccuracies of EBV predictions were higher in the DP than in the IP approach. The reverse was true for accuracy of phenotypic prediction when using $\beta_p$, but not when using $\beta_g$ and $\beta_r$. Within the DP approach, accuracies of EBV when using $\beta_g$ were higher than when using $\beta_p$, especially at the low genetic correlation scenario. However, there were no differences in EBV prediction accuracy between the $\beta_p$ and $\beta_g$ in the IP approach. Performance of both approaches increased with increase in accuracy of the calibration model, which increased with increase in genetic or residual structures between traits.

In conclusion, MT analyses would be useful when number of observations are small, for example for genetic evaluation of cows and sires with < 50 daughters. Use of the DP approach for prediction of EBV seems useful while the IP or PLS regression based prediction equations are a method of choice for phenotypic prediction. There is a direct relationship between performance of the two approaches and accuracy of the calibration model. Performance of the DP approach is sensitive to the type of PLS regression coefficients used. Milk predicted blood BHB is heritable and has moderate positive genetic correlations with ketosis. Therefore, predicted blood BHB can be an alternative for breeding cows to have lower susceptibility to ketosis.

## SAMANDRAG

Belay, T. K. 2017. ***Fleireigenskapsanalyse av infraraude mjølkespekter og prediksjonar frå dei i mjølkekupopulasjonar.*** Norges miljø- og biovitenskapelige universitetet, Philosophiae Doctor (Ph.d.) avhandling, 2017: 65, ISSN: 1894-6402, ISBN: 978-82-575-1462-4.

Fourier-transformerte midt-infraraude (FT-MIR) spekter frå mjølkeprøver er eitt slag fleireigenskapsinformasjon som blir registrert rutinemessig av mange mjølkekontrollorganisasjonar. Bruk av slike spekter er i ferd med å bli viktig for mjølkeforsking fordi eigenskapar som kan predikerast frå spekter kan vera nyttige for avl, fôringsrettleiing og helsekontroll. Fokus i denne PhD-oppgåva har vore å verifisera metodar for utnytting av FT-MIR-mjølkespekter til prediksjon av avlsverdiar og fenotypar.

I artikkel 1 blei testdagsobservasjonar av mjølk analyserte med den vanlege ein-eigenskaps dyremodellen (ST) og samanlikna med ein fleireigenskaps-modell (MT) til avlsformål. Arvegradsestimat blei svært like i begge modellane. MT-modellen ga betre sikkerhet for predikerte avlsverdiar for kyr (for eksempel 2,5% betre for mjølkemengde og 9,83% for proteinmengde) og for oksar med mindre enn 50 døtrer. Estimert avlsframgang var litt høgare med MT for protein- og feitt-mengde, men litt mindre for dei andre eigenskapane. Ved samanlikning av avlsverdiar frå MT og ST hadde oksar høgare rangkorrelasjon enn kyr. I artikkel 2 samanlikna vi to metodar til å predikera blod-β-hydroxybutyrat (BHB) frå FT-MIR-mjølkespekter hos polske kyr. Den eine metoden (indirekte prediksjon – IP) gjer spekteret om til éin eigenskap før analyse med ein blanda modell, den andre (direkte prediksjon – DP) bruker fleireigenskaps- blanda modell på (dimensjonsreduserte) spekterdata. For begge predikerer ein dei tilfeldige effektane i modellane, og predikerer til slutt fenotypar og avlsverdiar for BHB. IP og DP ga omtrent like resultat for prediksjon av fenotypisk blod-BHB. Univariat variansstruktur (IP) ga litt meir nøyaktige prediksjonar. Sikkerhetane for modellane var låge: 0,28-0,30 for IP, 0,26-0,30 for DP. For PLS med utransformert BHB var sikkerheten 0,29-0,37. I artikkel 3 blei samanhengen mellom FT-MIR-mjølkespekter og blod-BHB funnen for polske kyr brukt til å identifisera kyr i norsk kukontroll og sjekka om dei var behandla for ketose. Genetiske parameter for FT-MIR-predikert BHB og for registrert klinisk ketose (KET) blei estimerte. Genetisk samanheng mellom predikert BHB og KET og med mjølkeproduksjonseigenskapar blei òg estimerte. Arvegradar for predikert

blod-BHB for ulike laktasjonsstadiar var moderate: frå 0,250 til 0,365. Genetiske korrelasjonar var høgare for nære stadiar. Predikert blod-BHB for 11-30 dagar i laktasjonen hadde moderat genetisk korrelasjon med KET (0,469) og med mjølkeeigenskapar (-0,367 med proteininnhald og 0,277 med mjølkemengde). I artikkel 4 simulerte vi tre eigenskapar og samanlikna IP- og DP-metoden med ulike genetiske (låg: 0,1 til høg: 0,9) og residual- (null til høg: ± 0,9) korrelasjonar for eigenskapane. Resultat for dei to metodane, og for PLS, blei studerte. Dessutan såg ein på bruk av ulike regresjonskoeffisientar som blei estimerte frå simulerte fenotypar ($\beta_p$), sanne avlsverdiar ($\beta_g$) eller residualar ($\beta_r$) og kva effekt dette hadde. Sikkerhet for prediksjon av avlsverdiar (EBV) blei høgare med DP enn med IP. Når $\beta_p$ blei brukt var IP betre til fenotypisk prediksjon, men ikkje når $\beta_g$ eller $\beta_r$ blei brukte. Med DP blei sikkerheten høgare når ein brukte $\beta_g$ enn når ein brukte $\beta_p$, spesielt med låg genetisk korrelasjonssenariet. Med IP var det ikkje forskjell på bruk av $\beta_g$ eller $\beta_p$. For både DP og IP auka sikkerheten når sikkerheten med PLS auka. Dette skjedde når enten genetisk eller residual-korrelasjon auka.

Til konklusjon: MT kan vera nyttig når det er få observasjonar, for eksempel for kyr eller for oksar med >50 døtrer. DP-metoden synest å vera nyttig for avlsverdiutrekning, mens IP eller PLS-regresjon er betre for fenotypeprediksjon. Der er ein direkte samanheng mellom kor gode IP- og DP-metodane er og sikkerheten til PLS. DP-metoden er følsom for kva slag PLS-regresjonskoeffisient som blir brukt. Blod-BHB predikert med mjølkespekter er arveleg og har moderat positiv korrelasjon med ketose. Difor kan det vera eit alternativ for å avla for kyr som er mindre utsette for ketose.

**ABBREVIATIONS**

BHB – β-hydroxybutyrate

BLUP – Best Linear Unbiased Prediction

DIM – Days in Milk

DP – Direct Prediction

EBV – Estimated Breeding Values

EMSC – Extended Multiplicative Signal Correction

FT-MIR – Fourier Transform Mid-Infrared

HTD – Herd Test-Day

IP – Indirect Prediction

IR – Infrared

KET – Clinical Ketosis

MT – Multi-Trait

NRF – Norwegian Red

PCA – Principal Component Analysis

PE – Permanent Environment

PLS – Partial Least Square

PNE – Phenotype without Error

PWE – Phenotype with Error

REML – Restricted Maximum Likelihood

RMSE – Root Mean Square Error

SCK – Sub-Clinical Ketosis

SG – Savitzky-Golay

ST – Single-Trait

TBV – True Breeding Values

TD – Test-Day

TPV – True Phenotypic Values

## LIST OF PAPERS

I. T. K. Belay, M. Svendsen, T. Ådnøy. **Comparison of single-trait and multi-trait animal models for genetic evaluation of milk production traits predicted from milk infrared spectra in Norwegian dairy cattle.** *(Under resubmission to Acta Agriculturae Scandinavica, Section A - Animal Science)*

II. T. K. Belay, B. S. Dagnachew, Z. M. Kowalski, T. Ådnøy. **An attempt at predicting blood β-hydroxybutyrate from Fourier-transform mid-infrared spectra of milk using multivariate mixed models in Polish dairy cattle.**
*Journal of Dairy Science (In Press)*

III. T. K. Belay, M. Svendsen, Z. M. Kowalski, T. Ådnøy. **Genetic parameters of blood β-hydroxybutyrate predicted from milk infrared spectra and clinical ketosis, and their associations with milk production traits in Norwegian Red cows.**
*Journal of Dairy Science (In Press)*

IV. T. K. Belay, B. S. Dagnachew, S. A. Boison, T. Ådnøy. **Prediction accuracy of direct and indirect approaches, and their relationships with accuracy of calibration models.** *(Submitted to Journal of Dairy Science)*

# 1. GENERAL INTRODUCTION

## 1.1. Background

There is much unused information in animal husbandry. For example in Norway, about 45 % of the milk now comes from automatic milking systems (AMS). Additional analyses of milk, based on the light spectrum, conductivity, wet chemistry, ultrasound, etc., are becoming available. A corresponding information increase is found from slaughterhouse lines, aquaculture etc. Information that may be derived from these multivariate sources is only partially implemented. In the dairy industry, a focus is to improve farm management to increase competitiveness. Multivariate data from daily records may be better-modeled using multivariate analysis approaches including multi-trait mixed models. Use of such information will also be central to research in animal sciences to develop practical animal husbandry and aquaculture. Competence in the use of multivariate information will be useful for better estimation of breeding, feeding and health-related parameters.

In articles presented in this thesis, multivariate techniques and mixed multi-trait models have been applied to milk infrared spectral data. Phenotypes have been predicted from such spectra in Norwegian and Polish dairy cattle populations, to verify methods for exploiting milk FT-MIR spectral information for prediction of breeding values and phenotypes. Infrared spectra acquisition and its potential for phenotyping, methods of reducing dimensionality of the spectra to few components or to single traits, and methods for prediction of genetic parameters, breeding values and phenotypes for the components or traits, are described in the remaining part of this introduction.

## 1.2. Infrared spectroscopy: a potential tool for rapid phenotyping

There are several definition for spectroscopy. The most general definition of spectroscopy is the study of the interaction (i.e. absorption, emission, and reflection) between matter and electromagnetic radiation (Gengler et al., 2016;McParland and Berry, 2016). Historical development of spectroscopy is summarized elsewhere (Gengler et al., 2016). The infrared (IR) part of the electromagnetic radiation has three regions: near-IR (12800 – 4000 $cm^{-1}$), mid-IR (4000 – 200 $cm^{-1}$) and far-IR (200 – 10 $cm^{-1}$). Milk analysis mostly uses absorption IR spectroscopy

associated with the mid-IR region. Measurements are taken at up to a thousand different wavenumbers (e.g. 1,060 for Foss instruments) per milk sample. They are expressed as an inverted function of wavelengths, in centimeters$^{-1}$. The absorbance values along this range of wavenumbers form what is generally called a mid-IR spectrum. Observations in a given spectrum are then combined in a linear equation to predict the concentration of the milk component of interest (e.g. fat percentage).

Genetic and genomic evaluation of dairy animals depend on routine access to large quantities of phenotypic information on the animal itself or its relatives (Daetwyler et al., 2008). Gold standard methods are often not applicable for population-wide phenotyping due to high cost or other practical limitations, and are not rapid enough to obtain sufficient observations for genetic evaluations. Fourier transform mid-IR (FT-MIR) spectrometry is a potential tool to collect data at population level for phenotypic and genetic purposes. It is a rapid, nondestructive and cost-effective tool used worldwide in regular milk recording schemes and milk payment systems to quantify major milk components (i.e., fat, protein, casein, lactose and urea). The usefulness of FT-MIR to obtain new milk phenotypes such as more detailed milk composition, technological properties of milk, or cow physiological status, has been discussed in detail elsewhere (De Marchi et al., 2014;Bastin et al., 2016;McParland and Berry, 2016;Bonfatti et al., 2017b). The analysis of milk by FT-MIR spectrometry offers an opportunity to record a whole range of phenotypes to develop tools increasing profitability and sustainability of the dairy sector (Gengler et al., 2015). The predicted phenotypes can be used as indicator traits in dairy breeding programs for enhanced fertility and health (Bastin et al., 2016). For example, it created an opportunity for evaluating subclinical disease (e.g. ketosis) based on large numbers of phenotypic records available for indicator traits (Pryce et al., 2016).

### 1.3. Multivariate techniques for dimension reduction

Variable selection and dimension reduction is a major task for multivariate statistical analysis, and for multivariate regression. Stepwise regression is a well-known method for dimension reduction in regression analysis (Maitra and Yan, 2008). However, this method is not sound when several of the predictor variables are highly correlated, as independence is one of the primary assumptions in such a method. Error variances in estimates of the regression parameters increase when there is a high degree of correlation among the predictor variables (Maitra and Yan, 2008). This is known

as multi-collinearity in regression literature. Milk FT-MIR spectral variables, for example, have high dimension and exhibit strong correlations among each other (Soyeurt et al., 2010;Dagnachew et al., 2013a), and prediction equations potentially over-parameterize (Gengler et la., 2016). Therefore, multivariate techniques that are capable of capturing relevant information from the high dimensional spectral data, handle multi-collinearity, and derive prediction equation coefficient are required. In studies presented in this thesis, commonly used multivariate methods such as partial least squares (PLS) regression were used for dimension reduction, calibration of prediction models, and solving the multi-collinearity problem. We have also used principal component analysis (PCA) for spectral dimension reduction. Results from such study are not included in this thesis, but reported elsewhere (Belay et al., 2015), and referred to in articles presented in this thesis.

### 1.3.1. PCA

PCA is an unsupervised multivariate technique commonly used for dimension reduction and solving the multi-collinearity problem. PCA as a dimension reduction method is applied without considering the correlation between the response variable and the predictor variables. The purpose of PCA is to derive a few, *k*, latent traits/variables that are a linear combination of the original many, *m*, variables and that can be used to summarize the data without losing too much information contained in the *m* original variables (Martens and Naes, 1989). All the derived latent traits are orthogonal to each other. Mathematically, PCA decomposes a data matrix **X** with *n x m* dimension for *n* individual samples into an orthogonal *n x k* score matrix **T** (X-score) and an *m x k* loading matrix **P** (X-loading):

$\mathbf{X} = \mathbf{TP}' + \mathbf{E}$, where $\boldsymbol{E}$ is residual matrix of dimension *n x m*.

### 1.3.2. PLS regression

PLS regression is a more recent multivariate technique that generalizes and combines features from PCA and ordinary multiple regression. The goal of PLS regression is to predict response variables **Y** from predictors **X** and to describe their common structure (Martens and Naes, 1989;Abdi, 2003). Unlike the PCA, PLS searches for a set of factors that performs a simultaneous decomposition of **X** and **Y** with the constraint that these factors explain as much as possible of the covariance between **X** and **Y**, and use a decomposition of **X** to predict **Y**. When **X** is full rank and **Y** is a vector, the prediction of the response variable could be done using ordinary multiple

regression. PLS performs well even when the number of predictors is greater than the number of observations, or when strong collinearity exists between them: for example in spectroscopic data, where ordinary regression is no longer an option.

PLS explains most of the variation in both predictors and responses with fewer factors than the number of latent traits from PCA. The PLS model has the form:

$$\mathbf{X} = \mathbf{TP}' + \mathbf{E}$$

$$\mathbf{Y} = \mathbf{UQ}' + \mathbf{F}$$

Where **P** is a matrix of X-loadings; **U** is a matrix of Y-scores; **Q** is a matrix of Y-loadings that represents the correlation between the Y-variables and the X-score matrix **T**. **E** and **F** are error terms. The X-scores matrix **T** is then computed as **T**=**XW** and the PLS regression coefficients β of Y on X are computed as β=**WQ**'. **W** is a matrix of X-weights that reflects the covariance structure between the predictor and response variables.

### 1.4. Multivariate mixed model analysis

One important assumption in most linear model analyses, where only one or more fixed effects and one random effect (i.e. an error term) are considered, is independence of observations. For this assumption to hold, the data points should come from different individuals (not related) and each individual should only contribute one data point (no multiple responses). However, observations are not independent when individuals are clustered or grouped (e.g., clustered data) or when each individual is measured more than once in space (e.g., repeated measurements) or in time (e.g., longitudinal data). This is a common scenario in animal and plant sciences at least. The dependencies are resolved by introducing additional random effects with structure to the random residual in the so-called mixed models. These models are important for analysis of dependent data. When modeling genetic relationships between individuals and prediction of genetic merits or breeding values and covariance components, unbalanced designs often occur.

One of the potential targets of modern mixed modeling tools are multivariate analyses. Mixed model approaches have some advantages over classical multivariate analysis of variance techniques. In particular, they allow unbalanced data and relaxation of some of the usual assumptions of the linear models. In animal breeding, mixed models are commonly used for genetic analyses of traits in either single or multi-trait model settings. It may be challenging to

estimate (co)variance components of additive genetic and other random effects needed to run the mixed models, especially when more than 3-4 traits from a large number of animals with many records each are analyzed multivariately. Therefore, traits are often analyzed uni- or bi-variately or in subgroups. The precision of estimated genetic parameters and accuracy of predicted breeding values from the single trait (ST) models might be low, especially when the traits have low heritability and number of records are small. A way to increase the accuracy of estimated breeding values for traits in the breeding goal is to use multi-trait (MT) methods that enable fuller exploitation of the data, and to combine direct with indirect information from correlated traits (Pollak et al., 1984;Schaeffer, 1984;Thompson and Meyer, 1986;Van der Werf et al., 1992).

The use of MT analysis is not limited to improving accuracy of estimated breeding values (**EBV**), but also to give less biased estimates when animals have been selected based on values of another correlated trait (Henderson, 1975;Pollak et al., 1984). A model including information on the correlated trait is able to correct for this type of selection. The fact that the main selection in dairy cattle for many years has been on production, may lead to biased predicted EBVs for traits not included in the breeding goal when ST estimation is used, because genetic correlations to traits selected for are not accounted for in the EBV prediction. This may show up as bias in estimated genetic trends (Pollak et al., 1984). Despite the aforementioned importance of MT genetic evaluation, this method has often not been used in breeding schemes. There is for example no published work on Norwegian dairy cattle data using multivariate mixed models. Therefore, in one article presented in this thesis, we have compared the genetic evaluation of Norwegian dairy cattle under ST with MT animal models using test-day records.

### 1.5. Indirect and direct prediction approaches

Milk FT-MIR spectra are mainly utilized for phenotyping of individual traits. The predicted phenotypes are then, together with pedigree information and variance component estimates, used in best linear unbiased prediction (**BLUP**) to calculate individual EBV and other random components included into the model. This is the conventional method used today for genetic evaluation of animals and such approach has been referred to as indirect prediction (IP) (Dagnachew et al. 2013b). Alternatively, analyses for genetic evaluation of animals can directly be applied on the latent traits of milk FT-MIR spectral variables. BLUP predictions (EBV, herd

test-day, permanent environment and residual) for the traits of interest are predicted as correlated traits to the corresponding random components of spectra. Dagnachew et al. (2013b) referred to such an approach as direct prediction (DP).

The IP and DP approaches have been used to predict EBV for milk fat, protein and lactose contents in goats (Dagnachew et al., 2013b) and for traits related to fine milk composition and technological properties of milk in cows (Bonfatti et al., 2017a). Dagnachew et al. (2013b) showed that the DP approach reduced prediction error variance, resulting in 3-5% improved relative genetic gain using DP instead of the IP approach. They also reported high rank correlation coefficients (0.93 to 0.96) between EBV predicted using the IP and the DP. However, independent chemical analyses (reference values) for the milk contents were not available in that study. Possibly because of this, the coefficients of determination ($\mathbf{R^2}$) in calibrations were very high (> 0.96). Moreover, the accuracies of EBV were estimated based on coefficient matrices of the mixed model equations in that study. The DP and IP approaches have not been compared using independent reference data obtained by reference methods, except in the work of Bonfatti et al. (2017a). Bonfatti et al. (2017a) compared the two approaches for calibration equations using independent reference data measured by reference methods and for traits predicted with medium (0.35) to high (0.86) $R^2$ values. They reported rank correlation estimates ranging from 0.07 to 0.96. It has been indicated that the DP approach is more likely to be effective when traits of interest have high correlation with major sources of variation of the spectra (e.g. milk protein and fat contents) (Bonfatti et al., 2017). However, it is difficult to distinguish whether the IP or DP approach performed better for the cases of that study, because the IP and DP approaches were evaluated based on rank correlations.

Based on the studies of Dagnachew et al. (2013b) and Bonfatti et al. ( 2017a), it is difficult to make a conclusive remark on whether or when the DP approach is better than the IP approach for EBV prediction. The DP and IP approaches have not been compared systematically under different genetic and environmental correlation scenarios between traits of interest and spectral variables. In addition, in both previous studies (i.e. Dagnachew et al., 2013b and Bonfatti et al., 2017a) covariance components of the latent traits estimated by the DP approach were converted to variance components to be used in the IP approach using PLS regression coefficients estimated based on phenotype ($\beta_p$). Similarly, EBV of the latent traits were converted into EBV of traits of interest using the $\beta_p$. Utilization of a $\beta_p$ to convert EBV or covariance components of latent traits

into EBV or variance components of trait of interest does not seem appropriate. Therefore the effect of using different PLS regression coefficients estimated from phenotypes ($\beta_p$) and true breeding values ($\beta_g$) for converting multi-trait structures to single-trait structures on performances of the two approaches is unknown. Moreover, rank correlations between EBV obtained by the IP and the DP approach have been shown not to be related to accuracy of calibration equations (Bonfatti et al., 2017a). However, the relationships between accuracies of EBV obtained by the two approaches, and accuracy of calibration equations are not established. Furthermore, the potential of the DP approach for phenotypic prediction has not been evaluated.

### 1.6.Ketosis and its indicator traits

Monitoring of metabolic disorders in early lactation is important to evaluate transition cow management and as a guide to strategies to improve health and fertility on dairy farm (van der Drift et al., 2012a). Ketosis is a common metabolic disorder frequently observed in dairy cows during the early lactation period (Enjalbert et al., 2001;Zhang et al., 2012;Koeck et al., 2014). It is characterized by increased levels of ketone bodies (BHB, acetoacetate and acetone) in the blood, urine, and milk (Andersson, 1988). It is caused by severe negative energy balance and excessive body fat mobilization (De Roos et al., 2007). Ketosis can have a clinical and subclinical presentation in dairy cows. The clinical ketosis (**KET**) that has visible clinical signs occurs less frequently than the subclinical ketosis (**SCK**). The KET results in decreased feed intake, weight loss, and drop in milk yield (Foster, 1988;Radostits et al., 2007;Youssef et al., 2010). Reported incidences of KET vary from 0.24% in first parity to 17.2% in third parity with a median incidence of 3.3% (Pryce et al., 2016). The SCK is defined as an excessive level of circulating ketone bodies without clinical signs of ketosis (Andersson, 1988). It is associated with less milk yield (Duffield et al., 2009), reduced reproductive performance (Walsh et al., 2007;McArt et al., 2012), and higher risks for KET (Seifi et al., 2011) and displaced abomasum (LeBlanc et al., 2005;Duffield et al., 2009). Prevalences of SCK can vary between farms, ranging from 8.9 to 43% (McArt et al., 2012;Van der Drift et al., 2012a;Suthar et al., 2013). Usually the SCK is detected by testing the ketone concentrations in blood, urine, or milk.

It is difficult to assess the degree of ketosis problems in a herd based on the incidence of KET because many metabolic events including ketosis are subclinical by nature, and information on subclinical cases are mostly missing because it is difficult to detect (Pryce et al., 2016). Moreover,

diagnosis of KET is subjective, as definition of KET in herds and ability to detect clinical signs in early lactation cows may vary dramatically (Oetzel, 2007). Heritability estimates for ketosis have mostly been based on clinical records and are low, ranging from 0.01 to 0.16 (Pryce et al., 2016), partly due to the subjective nature of its diagnosis and to the low frequency of KET (Van der Drift et al., 2012b). Response to selection against KET is hampered by low reliabilities associated with the low heritability (Pryce et al., 2016). Use of information from correlated traits or from subclinical diagnosis could be an alternative to improve the accuracy of EBV and increase the selection response. Phenotypes derived from routinely collected data through milk recording such as fat-to-protein ratio and fatty acid profiles are promising ketosis indicators (Van Knegsel et al., 2010). Phenotypes more closely associated with ketosis, such as BHB and acetone in milk may also be valuable.

Concentration of BHB in blood has been used as a gold standard method for detection of SCK and several studies have used a threshold of 1.2 mmol/L (e.g. Van Knegsel et al., 2010; McArt et al., 2012; van der Drift et al., 2012a) or 1.40 mmol/L (Oetzel, 2004;Denis-Robichaud et al., 2014) to identify cows with SCK. However, the gold standard method does not allow routine testing of all animals at risk due to practical limitations such as difficulty in blood sampling (especially for farmers) and capacity for analyzing many blood samples at a time. Determination of ketone bodies in milk could make the sampling easier (Enjalbert et al., 2001;De Roos et al., 2007). As milk sampling is performed monthly in milk recording procedures, more routine measurements of milk BHB can be done by FT-MIR spectroscopy analysis in milk samples at test-days (De Roos et al., 2007;Van der Drift et al., 2012a). The BHB predicted from milk spectra have moderate heritability (0.07 to 0.40; Oikonomou et al., 2008; Jamrozik et al.,2016) and genetic correlations with KET (0.25 to 0.75; Koeck et al., 2014, 2016; Jamrozik et al., 2016); hence, indirect selection for ketosis using BHB as indicator trait should result in better genetic gain than direct selection for KET. KET itself has a very low heritability. For use and implementation of blood BHB predicted from milk spectra in dairy cattle breeding programs, knowledge of genetic parameters and genetic associations with clinical events and other traits in the breeding goal is essential. However, no report on genetic parameters and associations of predicted blood BHB with KET and milk production traits for cows in early lactation has been found.

## 2. AIM AND OUTLINE OF THIS THESIS

The overall aim of the research presented in this thesis was to verify methods for exploiting milk FT-MIR spectral information for prediction of breeding values and phenotypes. The specific objectives have been to:

- Compare ST and MT repeatability test-day animal models for genetic evaluations, and evaluate the practical usefulness of MT models compared to ST models in Norwegian dairy cattle.

- Verify whether multivariate mixed modeling of FT-MIR milk spectra in the form of factor scores (DP) gives better prediction of phenotypic blood BHB than the univariate approach (IP).

- See if an established connection between milk FT-MIR spectra and blood BHB in Polish dairy cattle could be used to identify Norwegian Red cows treated for ketosis, and estimate genetic parameters for the predicted blood BHB.

- Evaluate prediction accuracy of direct (DP) and indirect (IP) approaches, and their relationships with accuracy of calibration models using simulation.

Paper I assessed if there is a benefit from MT genetic evaluations compared to ST evaluation in test-day milk data.

Paper II evaluated the main objective of this thesis. The predictive ability of the IP and DP approaches for prediction of blood BHB from milk FT-MIR spectra were investigated. Prediction equations for blood BHB from milk FT-MIR spectra and reference blood BHB values in Polish dairy cattle were also developed.

Paper III applied the prediction equation developed for blood BHB using Polish data on FT-MIR spectra of Norwegian Red cows to predict blood BHB from milk spectra. Phenotypic relationships of the predicted blood BHB with veterinarian recorded ketosis (KET) and milk production traits were assessed. Also estimated was genetic parameters for the predicted blood BHB and clinical ketosis, and their genetic association with milk production traits.

Paper IV tried to wrap up importance of the IP and DP approaches using simulated data. Prediction accuracy of the IP and DP approaches for phenotype and EBV under different genetic and residual scenarios were assessed. Relationships between accuracy of the two approaches and accuracy of calibration models were also evaluated. Also evaluated was effect of using phenotype ($\beta_p$) or genetic ($\beta_p$) based calibration outputs for converting covariance components or EBV of latent traits into variance components or EBV of trait of interest on performance of the two approaches.

# 3. BRIEF SUMMARY OF PAPERS

## 3.1. Paper I

### Comparison of single-trait and multi-trait animal models for genetic evaluation of milk production traits predicted from milk infrared spectra in Norwegian dairy cattle

Estimation of (co)variance components of additive genetic and other random effects is challenging, especially when more than 3-4 traits from a large number of animals with many records each are analyzed multivariately. Consequently, traits are often analyzed univariately or in subgroups in breeding schemes. Information from correlated traits may improve the accuracy of predictions for a particular trait. Therefore, aim of this study was to evaluate the practical usefulness of MT models compared to ST models. The ST and MT repeatability test-day animal models were applied to 875,460 test-day records from Norwegian dairy cattle. Genetic parameters, accuracy of breeding values, rank correlation and genetic trends were estimated for milk production traits (milk, fat and protein yields, and fat and protein contents).

**Main results**

- Estimates of heritability ranged from 0.119 for fat yield to 0.377 for milk protein content but were very similar in both types of analyses.
- The MT analyses improved accuracies of EBVs for cows and for sires with less than 50 daughters, but were similar in analyses for sires with >50 daughters. Sires were less re-ranked compared to cows and all animals in pedigree when comparing MT and ST predictions.
- Favorable genetic trends were observed in all traits. The genetic trends were slightly higher under MT for protein and fat contents, and fat yield, but slightly lower for the remaining traits.

**Conclusion**

Comparison of the ST and MT animal models applied for the genetic evaluation of Norwegian dairy cattle revealed small differences between the estimates obtained by the two methods. Multi-trait analysis was more useful for cow genetic evaluations than for sire evaluations.

### 3.2. Paper II

**An attempt at predicting blood β-hydroxybutyrate from Fourier-transform mid-infrared spectra of milk using multivariate mixed models in Polish dairy cattle**

The aim of this study was to evaluate whether direct genetic analyses on milk FT-MIR spectra (DP) would give better prediction of blood BHB than univariate genetic analysis of BHB predicted from spectra before mixed model analysis (IP). In both cases, the prediction of phenotypes was the ultimate goal. The study also aimed to develop calibration equations to predict blood BHB from milk spectra. Links between (untransformed or log-transformed) blood BHB and (raw or pre-processed) milk FT-MIR spectra were developed (n=496) and validated (n=330). Calibration outputs were used to reduce dimension of milk FT-MIR spectral variables (n=158,028) into factor scores (DP) or into single-trait prediction of BHB (IP). Covariance components for the factor scores estimated and used for BLUP analyses in either covariance (DP) or variance structure (IP) were estimated. BLUP predictions of the random and estimations of the fixed effect parts of the model were used to predict BHB phenotypes for observations in the validation set in both approaches. Blood BHB predicted by both approaches were then regressed to the reference blood BHB values to evaluate performance of the IP and DP approaches.

**Main results**

- Predictive ability of calibration models were low ranging from 0.21 to 0.32 for untransformed blood and from 0.31 to 0.38 for log-transformed BHB in cross-validation analyses. The corresponding estimates in validation analyses were from 0.29 to 0.37 and 0.21 to 0.43, respectively, for untransformed and logarithmatic BHB.
- Predictive ability of the IP and DP approaches were also low, with slightly better prediction in IP (0.28-0.31) than in DP (0.26-30) approaches for phenotypic BHB prediction. Predictive ability of the two approaches were lower than prediction ability of calibration equations in the validation analysis.

**Conclusion**

Contrary to our expectation, slightly better predictions of BHB were found when univariate variance structure was used (IP) than when multivariate covariance structures were used (DP). Blood BHB log-transformation, spectral pre-processing and use of extreme blood BHB values improved prediction accuracy of the calibration models and the two approaches.

### 3.3. Paper III

**Genetic parameters of blood β-hydroxybutyrate predicted from milk infrared spectra and clinical ketosis, and their associations with milk production traits in Norwegian Red cows**

A few reports exist on genetic studies of plasma BHB measured by reference methods. However, there is no report on genetic parameters and associations of predicted blood BHB from milk spectra with clinical events and other traits in the breeding goal. The aim of this study was to estimate genetic parameters for milk spectra predicted blood BHB and for KET, and to examine their genetic association with milk production traits. Data on milk traits, KET and milk spectra were obtained from the Norwegian Dairy Herd Recording System. Data recorded up to 120 days after calving were considered. Blood BHB were predicted from Norwegian milk spectra using a model developed based on data from Polish dairy cows (Paper II). The milk components were also predictions from the milk spectra using Foss calibration. Veterinarian recorded KET data within 15 d before calving to 120 d after calving were used. Data were analyzed using ST or bivariate linear animal models.

**Main results**

- Contents of predicted blood BHB were in the range of literature values. Mean predicted blood BHB was higher at the beginning of lactation and decreased as DIM progressed.
- Heritability estimates for the predicted blood BHB at different DIM intervals were 0.230 to 0.365, while that for KET was 0.078 in ST, but 0.002 in bivariate analyses with either BHB or milk traits.
- Blood BHB at 11-30 DIM was moderately genetically correlated with KET (0.469) and milk traits (from -0.367 with protein content to 0.277 with milk yield), except with milk fat content (0.033). Estimates of genetic correlation of KET with milk production traits were -0.333 (with protein content) to 0.178 (with milk yield).

**Conclusion**

Prediction equations developed for the Polish dairy cows can be used for Norwegian Red cows to predict blood BHB to be used for management or breeding purpose. Predicted blood BHB at different DIM intervals or across lactation stages are heritable. Blood BHB can routinely be predicted from milk spectra analyzed from test-day milk samples, and thereby provides a practical alternative for selecting cows with lower susceptibility to ketosis, even though the correlations are moderate.

### 3.4. Paper IV

**Prediction accuracy of direct and indirect approaches, and their relationships with accuracy of calibration models**

Few studies have compared performance of the IP and DP approaches for EBV or phenotype prediction. It is difficult to make a conclusive remark on whether the DP approach is better than the IP approach based on studies done so far. The aim was to compare the IP and DP approaches for predictions of EBV and phenotypes under different genetic and residual correlation scenarios. We also evaluated relationships between performances of the two approaches and the accuracy of calibration equations. Moreover, effect of using different regression coefficients (e.g., phenotypic: $\beta_p$, genetic: $\beta_g$ etc.) on performance of the two approaches were evaluated. In this study, we simulated three traits under different genetic (low: 0.10 to high: 0.90) and residual (zero to high: ± 0.90) correlation scenarios between the traits and assumed that the first trait is a linear combination of the other two traits. The simulated data contained 2,100 parents (100 sires and 2000 cows) and 8,000 offspring (four offspring per cow). Of the 8,000 observations, 2,000 were randomly selected and used to develop links between the first and the other two traits using PLS regression analysis. The different PLS regression coefficients (such as $\beta_p$, $\beta_g$ etc.) were used in subsequent predictions following the IP and DP approaches. BLUP analyses were done on the remaining 6,000 observations using the 'true' (co)variance components that had been used for the simulation. Accuracy of prediction (of EBV and phenotype) was calculated as a correlation between predicted and true values from the simulations.

**Main results**

- Accuracies of EBV prediction were higher in the DP than in the IP approach. The reverse was true for accuracy of phenotypic prediction ($\beta_p$), but not when using $\beta_g$ and $\beta_r$ (regression coefficients for residual) where accuracy of phenotypic prediction in the DP was slightly higher than in the IP approach.
- Within the DP approach, accuracies of EBV when using $\beta_g$ were higher than when using $\beta_p$, especially at the low genetic correlation scenario. However, there were no differences in EBV prediction accuracy between the $\beta_p$ and $\beta_g$ in the IP approach. For phenotypic prediction, there was generally no difference in accuracy under $\beta_g$ and $\beta_p$ within either the DP or IP approach.

- Accuracy of the calibration models increased with increase in genetic and residual correlations between the traits. Performance of both approaches increased with increase in accuracy of the calibration models. Differences in phenotypic prediction accuracy between the two approaches became clearer as prediction ability of calibration models increased, but this was not the case for EBV prediction accuracy.

**Conclusion**

The DP approach is a good strategy for EBV prediction, but not for phenotypic prediction, where the classical PLS regression based equations or IP approach provided better results. Type of calibration outputs ($\beta_g$ or $\beta_p$) used for converting covariance components or EBV of latent traits into univariate structure had impact on accuracy of EBV, but not on accuracy of phenotypic prediction.

# 4. GENERAL DISCUSSION

The focus of this PhD study has been to verify methods for exploiting milk FT-MIR spectra for prediction of phenotypes and breeding values. Accuracy or validity of phenotypes (or genetic parameters) and breeding values predicted from milk spectra by different methods were discussed. Advantages and limitations of each method, as well as effect of multivariate techniques used in dimension reduction on accuracy EBV prediction in the DP approach were highlighted.

## 4.1. FT-MIR spectra for prediction of phenotypes and breeding values

### 4.1.1. Predicting phenotypes from milk FT-MIR spectra

Routine and accurate phenotype prediction is important in farm management and genetic and/or genomic evaluations of livestock. In studies presented in this thesis, phenotypes for traits of interest were predicted from FT-MIR spectra in either using 1) the classical PLS regression based prediction equations or 2) mixed model analyses following the IP and DP approaches. It is important that the predicted phenotypes are reasonable before using them for the desired purposes. In paper III, we used PLS regression based prediction equation to predict phenotypic blood BHB from FT-MIR spectra, and their validity could be assessed in several ways. From a biological point of view, the predicted blood BHB values made sense. For example, mean blood BHB concentrations were higher in early lactation and then decreased as DIM progressed, and its phenotypic distribution was similar to reference and published values. Heritability of predicted phenotypic blood BHB was in a range reported in literature (Van der Drift et al., 2012b;Koeck et al., 2014;Jamrozik et al., 2016;Koeck et al., 2016), but this alone is not sufficient. This is because estimates of heritability for traits predicted from FT-MIR might differ from estimates based on reference values for the same trait (Rutten et al., 2009;Poulsen et al., 2014;Bonfatti et al., 2017b). Validity of predicted phenotypes can also be assessed by studying its phenotypic associations with other traits, and we found meaningful phenotypic relationships between predicted blood BHB, KET and milk traits. For example, high-yielding cows had higher blood BHB concentration and were more prone to the risk of developing ketosis in early lactation compared with lower-yielding cows in Paper III. Similarly, cows with higher predicted blood BHB values had higher frequencies of KET (3.41%) compared with cows with lower blood BHB values (1.01%). Cows with higher predicted blood BHB values also had higher milk fat content throughout early lactation stage

compared with cows with lower blood BHB values. These results suggested that models developed for Polish dairy cows work and give reasonable results with Norwegian milk spectral data.

Application of mixed models for phenotypic prediction is not common in animal breeding. However, in paper II and IV, we used mixed model methods (i.e., IP and DP approaches) to predict phenotypes from milk spectra for blood BHB and other milk constituents. Accuracy of the predicted phenotypes were evaluated based on either coefficient of determination (paper II) or correlation coefficient (paper IV) between predicted and true (or measured) phenotypes. In using real data (paper II), slightly more accurate predictions of blood BHB phenotypes were observed when using the IP than the DP method. This is in contrast with previous reports (Dagnachew et al., 2013b), who found better accuracy of EBV prediction in using the DP than the IP approaches. It was also in contrast to our expectation that multivariate information would give better prediction than those that are univariate. Several possible reasons were postulated for the inferior phenotypic prediction by the DP approach. These include low genetic correlations between the latent traits, lack of enough information about contemporary cows in validation set, low correlation of blood BHB with milk contents (fat, protein etc.), and low accuracy of the calibration models developed. In paper IV, using simulated data, we tried to address some of the aforementioned limiting factors that we thought affected the performance of the DP approach. Results from the simulation study supported our previous finding from the real data (paper II) i.e. the IP approach performed better than the DP approaches for phenotypic prediction when using phenotypic regression coefficient ($\beta_p$). The exception was when using $\beta_g$ and $\beta_r$, where accuracy of phenotypic prediction in the DP was slightly higher than in the IP approach. Accuracy of phenotypic prediction in both approaches were affected by genetic and residual correlation structures between traits and accuracy of the calibration models used. Use of different regression coefficients had marginal effect on accuracy of phenotypic prediction. We also observed that spectral preprocessing and blood BHB log-transformation had an effect on phenotypic prediction accuracy in both IP and DP approaches.
Phenotypic prediction accuracies of the two approaches were also compared with classical PLS regression based prediction equations. Both with simulation and real data, the PLS regression based prediction equations performed better than the mixed model (IP and DP) methods, especially for phenotypes predicted without including the residual effects. For phenotypes predicted with residual effects, performance of the IP and PLS was equal, but better than the DP approach. This

indicated that inclusion of cows' circumstances at a given test-day into the IP or DP model did not improve prediction of phenotypes over the classical PLS method. Therefore, it does not seem to be a good strategy to use the IP or DP approach for phenotypic prediction, where the classical PLS regression based equations provided better results. Alternatively, information related to cow at a given test-day could be directly added to the spectra before PLS. For example, Vanlierde et al. (2015) included DIM directly into spectra using Legendre polynomial to predict methane, and prediction equations developed in such a way were shown to be more robust than equations that did not integrate the DIM information. Similarly, Shetty et al. (2017) used milk yield and live weight as predictors along with spectral variables to predict residual feed intake and dry matter intake. They showed improvement in accuracy of models that included spectral information along with milk yield and live weight as predictors for dry matter intake. Therefore, inclusion of cows' circumstances directly into spectra before PLS or using them as predictors along with spectral information during PLS can be an alternative to improve prediction accuracy for blood BHB from milk FT-MIR spectra.

### 4.1.2. Genetic parameters of traits predicted from FT-MIR spectra

In addition to using the FT-MIR predicted phenotypes in payment systems to reward or penalize producers or in monitoring metabolic status of cows (Gengler et al., 2016), they could be used as indicator traits in breeding programs for dairy cattle populations (Bastin et al., 2016; Gengler et al., 2016). The potential of FT-MIR predicted phenotypes in indirect selective breeding relies on heritability of the prediction and genetic correlation between the predicted and measured trait (Bonfatti et al., 2017b).

#### 4.1.2.1. Covariance components and heritabilities

In papers presented in this thesis, we have estimated (co)variance components and corresponding heritabilities for latent traits, blood BHB, and milk production traits predicted from milk spectra, using ST and MT models (paper I and III), or IP and DP approaches (paper II). Estimates of covariance components, heritabilities and other variance ratios of these traits were similar in ST and MT analyses. This is as expected because genetic and environmental variances of the traits themselves are not expected to be different, regardless of analyzing them univariately or multivariately. In paper II, variance components and heritability estimates for the DP predicted

blood BHB were slightly higher than the IP predicted BHB, indicating better information content in the DP approach. We also observed that spectral preprocessing had an effect on genetic parameter estimates. For example, in Paper II, most of the factor scores and blood BHB that were predicted from unprocessed spectra had higher estimates of heritability and proportion of variance due to permanent environment and herd test-date effects than those from preprocessed spectra. In paper III, genetic parameters of milk spectra-derived blood BHB and milk production traits increased with increase in DIM in both ST and MT analyses. This suggest that expression of additive genetic effects of these traits increase with the progress in lactation stage of the cow. A higher genetic variance rather than a decrease in environmental variance is the main cause for the increased heritabilities with the progress in DIM.

Recently, Bonfatti et al. (2017b) evaluated effect of predictive ability of calibration models on genetic parameter estimates (additive genetic and phenotypic variances, heritability, and genetic correlation between measured and infrared predicted traits). Those authors found a decrease in additive and phenotypic variances of predicted traits compared with measured traits, and the reduction in the variances were lower for traits predicted with higher $R^2$. However, the magnitude of heritability estimates for predicted traits was not related to $R^2$ of calibration models (Bonfatti et al., 2017b). Gengler et al. (2016) indicated that random noise of prediction (prediction error) would affect the heritability of the predicted traits.

Heritability of predicted traits also varies depending on the types of lactation records used (i.e., test-day or 305-d data). Milk, fat and protein yields are part of the Norwegian red (NRF) breeding goals and the genetic evaluations for these traits are based on the conventional 305-d lactation records. Estimates of heritabilities for the 305-d lactation milk, fat, and protein yields are 0.28, 0.21, and 0.24, respectively (http://www.genoglobal.com/Start/Norwegian-Red/about-norwegian-red/Norwegian-Red-Total-Merit-Index/), which is higher than corresponding estimates based on test-day records (paper I). Several studies (Shadparvar and Yazdanshenas, 2005;Miglior et al., 2007;Ptak et al., 2012) also reported higher heritabilities of whole-lactation traits compared with test-day heritabilities. Those authors implied that increased heritabilities of 305-d yield could be related to residual variance, which might be decreased when taking an overall value of all test-days used for making the 305-d. In addition, there is a difference in defining heritability between our study and NRF breeding company (e.g. Geno), where heritability was defined as ratio of

genetic variance to the sum of genetic and residual variance. If such a definition of heritability is adopted in our study, heritability estimates for 305-d yield and test-day yields become very similar at least for milk and protein yields in the Norwegian dairy cattle population.

#### 4.1.2.2. Genetic correlations

In addition to heritability and additive genetic variance of predicted traits, genetic correlation between the predicted traits (indicator traits – blood BHB) and traits of interest (e.g., ketosis) is a key factor affecting the potential usefulness of the predicted trait for indirect selective breeding programs. In paper III, we found moderate genetic correlations (0.469) between KET and its indicator (BHB at 11-30 DIM), as well as between BHB at 11-31 DIM and milk production traits (ranged from -0.367 with protein content to 0.277 with milk yield). Those correlations decreased as DIM progressed. Given its heritability, genetic correlation and routine availability, predicted blood BHB in early lactation could be used as an indicator trait in a routine genetic evaluation for resistance to ketosis, but selective breeding for lower BHB might have negative impact on yield traits. Genetic correlations between predicted blood BHB and other traits in breeding goal of NRF breeding program (e.g., health, fertility and conformation traits) are not known and need to be evaluated before considering the BHB in genetic selection. In a Canadian dairy population, lower EBV for milk BHB at early lactation stage was favorably correlated with several health and fertility measures, including somatic cell score, calving to first service, number of services, first service to conception, and days open (Koeck et al., 2014). Oikonomou et al. (2008) found moderate genetic correlations between blood BHB and several fertility traits, with estimates ranging from -0.65 (between blood BHB and conception rate in the first 305 d of first lactation) to 0.56 (between blood BHB and number of inseminations per conception). Longevity, overall score for conformation and for feet and legs, were favorably associated with milk BHB EBV (Koeck et al., 2014).

### 4.1.3. Predicting breeding values from milk FT-MIR spectra

As indicated earlier, EBV prediction from milk FT-MIR spectra (or from predicted phenotypes) for the traits of interest could be obtained using the conventional ST and MT methods or the contemporary IP and DP approaches.

#### 4.1.3.1. EBV prediction using the ST and MT mixed model analyses

For traits predicted from FT-MIR spectra, EBV would be estimated in either an ST or MT mixed model setting. In paper I, generally, we found little improvement in accuracy of EBV prediction for milk production traits in using the MT instead of ST model analysis. This could be due to the similarities in heritability and small differences in genetic and residual correlations between the investigated traits, as concluded in the older studies (Schaeffer, 1984;Falconer and Mackay, 1996). In a simulation study by Schaeffer (1984), he speculated that MT analyses for milk and fat yields would result in around 5% reduction of prediction error variance (PEV) (hence increase EBV accuracy) because heritabilities of the trait are similar and the difference between genetic and residual correlations is small. He also speculated that a larger percentage reduction of PEV would be had if milk yield, fat and protein percentages were analyzed simultaneously. In using real data, however, we observed little to no difference in accuracy when milk yield has been analyzed with yields or percentages fat and protein, especially for sire genetic evaluation. This would not be the case for MT analyses of fertility traits with production traits. For example, several studies have shown that MT analyses improved accuracy of EBV prediction for fertility traits when analyzed with milk production traits compared to analyzing only fertility traits together or separately (Kadarmideen et al., 2003;Biffani et al., 2005;Sun et al., 2010).

As discussed in paper I and elsewhere (Gengler and Coenraets, 1997;Guo et al., 2014), MT analyses would be useful when numbers of observations are small. With smaller numbers of observations (e.g., for sires with <50 daughters), EBV from MT evaluation tend to gain more in accuracy (e.g., 3.25% improvement for protein yield) compared to EBV accuracy from ST analysis. Moreover, MT analysis was found more useful for cow genetic evaluations than sire evaluations. Improvement in accuracy of cow EBV due to MT ranged from 2.5 % for milk yield to 9.83 % protein yield. We found similar results when using simulated data in paper IV, where the difference in EBV accuracy between the IP and DP approaches was more noticeable for cows than for sires' genetic evaluations. Relatively higher improvement in EBV accuracy due to MT analysis for cows than for sires in lactation yields of dairy cattle in Belgium were also reported (Gengler and Coenraets, 1997). The relative improvement in accuracy we found in paper I for milk yield for both cows and sires were similar to their results, but that of protein yield was higher in paper I than reported by Gengler and Coenraets (1997).

#### 4.1.3.2. EBV prediction using the IP and DP approaches

In addition to the indirect EBV prediction, using ST or MT models for traits predicted from spectra, EBV could be predicted from the spectra following the DP approach. In both the DP and IP approaches, in paper IV, we observed increase in accuracy of EBV prediction with increase in genetic and residual correlations between the traits, as well as with increase in calibration accuracy. In paper IV, we simulated traits using genetic correlations that were almost in opposite direction to residual correlations and found higher accuracy of EBV prediction in the DP than the IP approach; confirming results of previous study on goat milk contents (Dagnachew et al., 2013b). Nevertheless, the difference in performance between the IP and DP approaches was marginal when we simulated the traits using genetic correlations that were similar (in magnitude and sign) to residual correlations between the traits. This is in line with earlier reports (Schaefer, 1984; Thompson and Meyer, 1986) that indicated that the larger the difference between the residual and genetic correlations the better the accuracies from MT analysis. Generally, improvement in accuracy of EBV prediction in the DP over the IP approach depends on genetic and residual correlations between traits, accuracy of calibration models, type of PLS regression coefficients and amount of information (number of observations) used. Effects of each of these factors on EBV accuracies are described in the next paragraphs.

The DP approach outperformed the IP approach in the cases simulated, especially when using $\beta_g$ and at low genetic correlation scenario. Average rate of improvement in the DP over the IP approach when using $\beta_p$ ranged from 4.09% for high genetic correlation to 54.43% for low genetic correlation scenarios. The corresponding values when using the $\beta_g$ were from 4.22% for high genetic correlation to 248.94% for low genetic correlation scenarios. Here, two questions could possibly be raised: 1) why much better accuracy in DP than IP at low genetic correlation regardless of the type of βs used? 2) Why much better accuracy in DP than IP at low genetic correlation when using $\beta_g$ rather than $\beta_p$?

The first question might be related to the accuracy of the calibration model from which the βs were estimated. Accuracy of calibration itself depends on degree of genetic correlations between the traits. At low accuracy of calibration (i.e. when it is difficult to predict the intended trait from the spectra), the estimated βs that generated the predicted phenotype (for IP) or EBV (for DP) capture little information relevant for the intended traits. The EBV prediction from such poorly predicted

phenotypes (IP) would be inaccurate. In the DP approach, however, existing info in the βs regarding the intended trait is fully utilized, resulting in better accuracy of EBV prediction than in the IP approach. As the accuracy of the calibration model increased, however, the βs contained almost all information about the traits of interest and hence accurately predict phenotypes that later used for better EBV prediction in the IP approach. Breeding values prediction from such accurately predicted phenotypes are expected to be accurate. Hence, under high accuracy of calibration models, the DP approach utilizes little information that was not utilized by the IP approach, resulting in smaller differences in performance between the approaches.

Why much better accuracy in DP than IP at low genetic correlation when using $\beta_g$ rather than $\beta_p$? In the DP approach, EBV of the latent traits are predicted more accurately due to utilization of covariance between traits and are combined through βs into EBV of the intended trait. That means accuracy of EBV for the trait of interest depends on the type of β and amount of information contained in the β used to convert the EBV of latent traits into EBV of the traits of interest. When using $\beta_p$ that contained little information about genetic part of the intended trait, EBV accuracy of the intended traits might be lower than when using $\beta_g$ that are expected to contain more genetic information. That is why we observed larger variation in performance between the two approaches (as well as within the DP approach) at low genetic scenario when using $\beta_g$ than using $\beta_p$. The difference in performance of the IP and DP approaches due to the type of βs used would be reduced as accuracy of calibration models increased. Unlike for the phenotypic prediction accuracy, difference in EBV prediction accuracy between the two approaches did not vary widely as the accuracy of calibration model increased. Therefore, even though the prediction accuracy of EBV increased with increase in the accuracy of calibration model, it is not necessary to have a calibration model with high accuracy to see the difference in performance between the two approaches for EBV accuracy prediction. As indicated above, the difference in EBV accuracy between the IP and DP approaches was more noticeable when using small number of records (e.g., for cows) than large number of observations (for sire genetic evaluation).

### 4.2. Multi-trait mixed model vs direct prediction approach

The conventional MT mixed model is similar in principle to the DP approach, as both involve analyses of two or more traits simultaneously. Both perform better under small number of

observations. However, they differ conceptually. Unlike in the MT model, traits in the DP approach are not 'full' traits, but component traits that later combine into a 'full' trait through calibration outputs. Hence, performance of the DP approach depends on not only genetic and residual relationships between component traits but also on the relationship between the component traits (spectral variables) and the trait of interest. This has been confirmed using simulated data in paper IV. In the conventional MT mixed model analysis, however, only genetic and residual correlation structures between traits matter, especially for traits measured independent of spectra. MT model accounts for selection bias, but the DP may not, and performance of DP depends on type of regression coefficients and accuracy of calibration model.

In addition, EBV or phenotype prediction in the DP approach is more complex and tedious compared to the conventional MT model or the IP approach. In the DP approach, all steps from spectral dimension reduction to combining component estimates into estimates for traits of interest have to be done. Furthermore, the number of components retained and/or amount of original spectral variations captured by the components have an impact on the performance of the DP approach. The techniques used for spectral dimension reduction also influences performance of the DP approach. For example, when the DP approach is based on latent traits from PCA rather than from PLS regression, relevant information needed for prediction could be lost, as PCA does not take into account information of the response variable while decomposing spectral variables.

Despite improving EBV accuracy, increasing precision of genetic parameters and accounting for selection bias, the MT or DP analysis are complex and computationally demanding. In general, parameter estimation in MT mixed models is cubic in the number of traits. It was worse when traits were interdependent (e.g., MT analysis of fat and protein yields with milk yield, milk fat and protein contents), as well as when phenotypic values of (latent) traits are very small. Under such conditions, the analysis went through a number of iterations with small change in log likelihood and spent long time per iteration. In addition to computational challenges, covariance component estimations suffer from overfitting as the dimension of the trait-covariance increases (Bickel and Levina, 2008), mainly because the number of parameters of these models grow quadratically whereas data only linearly (Lippert et al., 2014).

### 4.3. Multivariate calibration and dimension reduction

As indicated earlier, PLS regression was used to extract relevant spectral information from high dimensional milk FT-MIR spectra into few factor scores and develop link between the spectra and traits of interest. Optimal number of PLS factors is usually determined based on minimum of prediction errors for cross-validation, external validation etc. Several strategies are applied to determine the optimal number from the prediction errors: global minimum – often result in overfitting; local minimum – most often used; and one standard error rule (Varmuza and Filzmoser, 2009). In paper II, five to ten PLS factors were retained based on local minimum value in root mean square error of cross-validation resulting in $R^2$ ranging from 0.21 to 0.32 for untransformed BHB and from 0.31 to 0.38 for log-transformed BHB. Those PLS factor scores explained from 96 to 99% variation in original predictors (spectral variables), but explained < 45% variation in the response variable (blood BHB). The low accuracies of prediction models could be due to non-linear or weak relationship between blood BHB and milk spectra, low concentrations of BHB in milk, and difference in metabolism of BHB in blood and milk due to sampling time and genetic differences between cows.

In using PCA for dimension reduction, components that explain about 99% of the original spectral variations are usually considered. However, the remaining 1% of the total spectral variation could contain relevant information needed for prediction (Dagnachew et al., 2013b; Belay et al., 2015; Bonfatti et al., 2017a). For example, Bonfatti et al. (2017a) showed that a considerable amount of information needed to predict phenotypes is lost when using 99% of original spectral variability, and that loss of such information could affect prediction of EBV from spectral information. One way of capturing more of the remaining 1% spectral variation is to increase the number of components retained. Doing that, however, might lead to overfitting and poses difficulties in estimation of covariance components for the retained large number of latent traits. The latter could be addressed by integrating PCA with canonical transformation as implemented in the study by Soyeurt et al. (2010), who managed to estimate covariance components for 46 latent traits.

Another way of capturing relevant information from spectra is to retain latent traits based on their associations with the traits of interest, instead of retaining them based on amount of their eigenvalues (Bonfatti et al., 2017a). Latent traits with the greatest eigenvalues do not necessarily contain the greatest part of spectral genetic variation (Dagnachew et al., 2013a, b; Belay et al.,

2015; Bonfatti et al., 2017a; Paper II). Those authors showed that latent traits having the largest eigenvalues are not necessarily those having the highest heritability estimates. Other latent traits that explained limited variation might play fundamental roles in the prediction of some traits (Bonfatti et al., 2017a). Moreover, PLS regression captures relevant variation of spectra associated with traits included in the PLS calibration models and hence are expected to give better prediction for those traits. This has been confirmed in paper II and Bonfatti et al. (2017a). For example, phenotypic prediction accuracy of DP was much lower than the IP approach when PCA was used (Belay et al., 2015) compared to when PLS was used for spectral dimension reduction (paper II) for the same dataset. The retained eight latent traits from PCA that explained 99% of the total spectral variation (Belay et al., 2015) could not contain as much relevant information about the blood BHB as the five PLS factors used in paper II did. Bonfatti et al. (2017a) showed that traits that were difficult to predict from spectra had much lower $R^2$ from principal component regression (PCR – that used the PCA principle) than the $R^2$ from PLS regression, but for traits that were easily predicted from spectra the PLS and PCR based calibrations had similar performance. However, PLS regression will not guarantee that information for other milk composition traits, which are not included in the calibration model, are retained in the factor scores.

### 4.4. Data quality and quantity

For papers presented in this thesis, already existing datasets were used, except the simulated data in Paper IV. Milk production traits (milk yield, and milk fat, protein and lactose contents), milk FT-MIR spectra, ketosis and pedigree information were obtained from Norwegian dairy control recording system (TINE). Measured blood BHB, milk FT-MIR spectra and pedigree information were obtained from Polish Federation of Cattle Breeders and Dairy Farmers (PFCBDF). These datasets are from other completed projects (e.g., blood BHB) and routine measurements as part of genetic evaluation programs.

Quality of the data obtained from milk FT-MIR spectra depends on accuracy of prediction models used, which is influenced at least by sample size from which reference values to be used in the calibration are measured. As the reference values are not routinely available, prediction equations are often developed based on small sample sizes and hence possibly with low accuracy, especially for traits with low concentrations. As a result, calibration models are validated internally and their predictive ability for new data sets becomes low. A way to increase sample size and hence improve

prediction accuracy is collaboration with different institutes within and across countries to create a common database that combines reference values and FT-MIR spectra. This may need standardization of both the spectra and the reference measurements as different instruments and technicians measure them. This could improve prediction ability of calibration models and enable the detection of relevant phenotypic traits and their reflection in the spectra (Friedrichs et al., 2015). In addition to the sample size, the way in which blood BHB was measured (i.e., values with few digits: 0.1, 0.2, … 6.3) have influenced predictive ability of the prediction models. Many samples had the same BHB values, resulting in a large number of few distinct values that reduced variation or range of values used. Instruments that can give more digits may help to increase the observed variability of blood BHB values and hence possibly the prediction accuracy.

## 5. GENERAL CONCLUSIONS

Based on findings of studies included in this thesis, the following concluding remarks are made:

- Multi-trait analysis were found to improve accuracy of EBV prediction; for example from 2.50 to 9.83 % due to MT analysis in Paper I and from 4.09% to 56.43% due to the DP approach in Paper IV. However, the benefits obtained from the MT models are conditional on several factors (heritability, genetic and residual correlation structures, number of observations, types of regression coefficient etc.).
- A more accurate prediction of blood BHB phenotypes were found with the IP than with the DP approach. Both the IP and DP approaches had lower predictive ability for phenotyping than the classical PLS regression based prediction equations, indicating unnecessity of doing mixed model to account for cows' information in phenotypic prediction from milk spectra. This has also been confirmed in the simulation study (Paper IV). Therefore, for phenotypic predication, the classical PLS regression based prediction equation seems the method of choice.
- Prediction equations developed for the Polish dairy cows can be used for Norwegian Red cows to predict blood BHB to be used for management or breeding purposes. Blood BHB at different DIM intervals or across lactation stages is heritable and has moderate genetic correlations with ketosis and milk production traits. Blood BHB can routinely be predicted from milk spectra analyzed from test-day milk samples, and thereby provides a practical alternative for selecting cows with lower susceptibility to ketosis, even though the genetic correlations are moderate.
- The DP approach is confirmed to be a method of choice for EBV prediction directly from heritable parts of spectra, but it's performance was inferior for phenotypic prediction, where the classical PLS regression based equations provided better prediction accuracy. Even though it is difficult practically to get $\beta_g$, use of the $\beta_g$ for converting EBV of latent traits into EBV of traits of interest improved accuracy of EBV in the simulations.

## 6. FURTHER RESEARCH

Based on the results presented, the following topics are recommended for further research:

- In paper I, the conventional repeatability animal model was applied to test-day data, and difference in estimates between the ST and MT repeatability models were small. Other test-day models that are expected to better suit the nature of such repeated measurement may be evaluated (e.g. using random regression).

- The link between milk spectra and phenotypic blood BHB was developed with a simple classical method (i.e., PLS) and the circumstances of cows at a given test day was included in the mixed model after PLS regression. Predictive ability of prediction equations developed in paper II was low. Further research on how to improve accuracy of blood BHB prediction from milk FT-MIR is needed. For example, inclusion of cow related information (yield traits, DIM, etc.) directly into spectra before PLS or using them as predictors along with spectral information. Use of Bayesian methods may be an alternative to improve prediction accuracy for blood BHB from milk FT-MIR spectra.

- Paper III evaluated only the genetic and phenotypic associations of predicted blood BHB with ketosis and milk production traits. Hence, further studies on genetic associations of BHB with health, fertility and other traits are needed before commencing selection for a lower BHB in NRF dairy cattle. Moreover, the benefit of using FT-MIR predicted indicator trait (e.g. BHB) in addition to the directly observed ketosis should be studied.

- Paper IV confirmed the importance of using the heritable part of spectra for better prediction of EBV for the traits of interest. However, importance of environmental components (HTD, PE etc.) of spectra in animal husbandry (for example as farm management tool) was not yet studied.

## 7. REFERENCES

Abdi, H. 2003. Partial least square regression (PLS regression). Encyclopedia for research methods for the social sciences:792-795.

Andersson, L. 1988. Subclinical ketosis in dairy cows. Veterinary clinics of north america: Food animal practice 4(2):233-251.

Bastin, C., L. Théron, A. Lainé, and N. Gengler. 2016. On the role of mid-infrared predicted phenotypes in fertility and health dairy breeding programs. J. Dairy Sci. 99(5):4080-4094.

Belay, T. K., K. Słoniewski, Z. Kowalski, and T. Adnoy. 2015. Predicting ketosis from milk mid infrared (MIR) spectra using multivariate mixed models in Proc. Third DairyCare Conference, Croatia, Zadar.

Biffani, S., M. Marusi, F. Biscarini, and F. Canavesi. 2005. Developing a genetic evaluation for fertility using angularity and milk yield as correlated traits. Interbull Bulletin 33:63-66.

Bonfatti, V., D. Vicario, L. Degano, A. Lugo, and P. Carnier. 2017a. Comparison between direct and indirect methods for exploiting Fourier transform spectral information in estimation of breeding values for fine composition and technological properties of milk. J. Dairy Sci. 100(3):2057-2067.

Bonfatti, V., D. Vicario, A. Lugo, and P. Carnier. 2017b. Genetic parameters of measures and population-wide infrared predictions of 92 traits describing the fine composition and technological properties of milk in Italian Simmental cattle. J. Dairy Sci. (in press).

Daetwyler, H. D., B. Villanueva, and J. A. Woolliams. 2008. Accuracy of predicting the genetic risk of disease using a genome-wide approach. PloS one 3(10):e3395.

Dagnachew, B. S., A. Kohler, and T. Ådnøy. 2013a. Genetic and environmental information in goat milk Fourier transform infrared spectra. J. Dairy Sci. 96(6):3973-3985.

Dagnachew, B. S., T. H. E. Meuwissen, and T. Ådnøy. 2013b. Genetic components of milk Fourier-transform infrared spectra used to predict breeding values for milk composition and quality traits in dairy goats. J. Dairy Sci. 96(9):5933-5942.

De Marchi, M., V. Toffanin, M. Cassandro, and M. Penasa. 2014. Invited review: Mid-infrared spectroscopy as phenotyping tool for milk traits. J. Dairy Sci. 97(3):1171-1186.

De Roos, A., H. Van Den Bijgaart, J. Hørlyk, and G. De Jong. 2007. Screening for subclinical ketosis in dairy cattle by Fourier transform infrared spectrometry. J. Dairy Sci. 90(4):1761-1766.

Denis-Robichaud, J., J. Dubuc, D. Lefebvre, and L. DesCôteaux. 2014. Accuracy of milk ketone bodies from flow-injection analysis for the diagnosis of hyperketonemia in dairy cows. J. Dairy Sci. 97(6):3364-3370.

Duffield, T., K. Lissemore, B. McBride, and K. Leslie. 2009. Impact of hyperketonemia in early lactation dairy cows on health and production. J. Dairy Sci. 92(2):571-580.

Enjalbert, F., M. Nicot, C. Bayourthe, and R. Moncoulon. 2001. Ketone bodies in milk and blood of dairy cows: Relationship between concentrations and utilization for detection of subclinical ketosis. J. Dairy Sci. 84(3):583-589.

Falconer, D. S. and T. F. Mackay. 1996. Introduction to quantitative genetics. Harlow: Addison Wesley Longman.

Foster, L. 1988. Clinical ketosis. Veterinary Clinics of North America: Food Animal Practice 4(2):253-267.

Friedrichs, P., C. Bastin, F. Dehareng, B. Wickham, and X. Massart. 2015. Final OptiMIR Scientific and Expert Meeting: From milk analysis to advisory tools (Palais des Congrès, Namur, Belgium, 16-17 April 2015): Optimir-a project aiming the development of novel mid-infrared based management tools for dairy herds. Biotechnol Agron, Soc Environ 19(2):97.

Gengler, N. and P. Coenraets. 1997. Use of multiple-trait animal models for genetic evaluation of milk, fat and protein lactation yields of dairy cattle in Belgium. Biotechnol Agron, Soc Environ 1(1):26-33.

Gengler, N., H. Soyeurt, F. Dehareng, C. Bastin, F. Colinet, H. Hammami, and P. Dardenne. 2015. Overview of possibilities and challenges of the use of infrared spectrometry in cattle breeding. Page 353 in Proc. Proc. EAAP 66th Annual Meeting. Wageningen Academic Publishers, Wageningen, the Netherlands.

Gengler, N., H. Soyeurt, F. Dehareng, C. Bastin, F. Colinet, H. Hammami, M.-L. Vanrobays, A. Lainé, S. Vanderick, and C. Grelet. 2016. Capitalizing on fine milk composition for breeding and management of dairy cows. J. Dairy Sci. 99(5):4071-4079.

Guo, G., F. Zhao, Y. Wang, Y. Zhang, L. Du, and G. Su. 2014. Comparison of single-trait and multiple-trait genomic prediction models. BMC genetics 15(1):1.

Henderson, C. R. 1975. Best linear unbiased estimation and prediction under a selection model. Biometrics:423-447.

Jamrozik, J., A. Koeck, G. Kistemaker, and F. Miglior. 2016. Multiple-trait estimates of genetic parameters for metabolic disease traits, fertility disorders, and their predictors in Canadian Holsteins. J. Dairy Sci. 99(3):1990-1998.

Kadarmideen, H. N., R. Thompson, M. P. Coffey, and M. A. Kossaibati. 2003. Genetic parameters and evaluations from single-and multiple-trait analysis of dairy cow fertility and milk production. Livest. Prod. Sci. 81(2):183-195.

Koeck, A., J. Jamrozik, G. Kistemaker, F. Schenkel, R. Moore, D. Lefebvre, D. Kelton, and F. Miglior. 2016. Genetic and phenotypic associations of milk ß-hydroxybutyrate with ketosis in Canadian Holsteins. Can. J. Anim. Sci. 96:302-305.

Koeck, A., J. Jamrozik, F. Schenkel, R. Moore, D. Lefebvre, D. Kelton, and F. Miglior. 2014. Genetic analysis of milk β-hydroxybutyrate and its association with fat-to-protein ratio, body

condition score, clinical ketosis, and displaced abomasum in early first lactation of Canadian Holsteins. J. Dairy Sci. 97(11):7286-7292.

LeBlanc, S., K. Leslie, and T. Duffield. 2005. Metabolic predictors of displaced abomasum in dairy cattle. J. Dairy Sci. 88(1):159-170.

Maitra, S. and J. Yan. 2008. Principle component analysis and partial least squares: Two dimension reduction techniques for regression. Applying Multivariate Statistical Models 79.

Martens, H. and T. Naes. 1989. Multivariate Calibration John Wiley & Sons. Chichester, UK.

McArt, J., D. Nydam, and G. Oetzel. 2012. Epidemiology of subclinical ketosis in early lactation dairy cattle. J. Dairy Sci. 95(9):5056-5066.

McParland, S. and D. Berry. 2016. The potential of Fourier transform infrared spectroscopy of milk samples to predict energy intake and efficiency in dairy cows. J. Dairy Sci. 99(5):4056-4070.

Miglior, F., A. Sewalem, J. Jamrozik, J. Bohmanova, D. Lefebvre, and R. Moore. 2007. Genetic analysis of milk urea nitrogen and lactose and their relationships with other production traits in Canadian Holstein cattle. J. Dairy Sci. 90(5):2468-2479.

Oetzel, G. R. 2004. Monitoring and testing dairy herds for metabolic disease. Veterinary clinics of North America: Food animal practice 20(3):651-674.

Oetzel, G. R. 2007. Herd-level ketosis–diagnosis and risk factors. in Proc. Proceedings of the 40th annual conference of bovine practitioners, Vancouver, Canada. Citeseer.

Oikonomou, G., G. Valergakis, G. Arsenos, N. Roubies, and G. Banos. 2008. Genetic profile of body energy and blood metabolic traits across lactation in primiparous Holstein cows. J. Dairy Sci. 91(7):2814-2822.

Pollak, E., J. Van der Werf, and R. Quaas. 1984. Selection bias and multiple trait evaluation. J. Dairy Sci. 67(7):1590-1595.

Poulsen, N. A., C. Eskildsen, T. Skov, L. B. Larsen, and A. J. Buitenhuis. 2014. Comparison of Genetic Parameters Estimation of Fatty Acids from Gas Chromatography and FT-IR in Holsteins. in Proc. 10th World Congress on Genetics Applied to Livestock Production (WCGALP).

Pryce, J., K. P. Gaddis, A. Koeck, C. Bastin, M. Abdelsayed, N. Gengler, F. Miglior, B. Heringstad, C. Egger-Danner, and K. Stock. 2016. Invited review: Opportunities for genetic improvement of metabolic diseases. J. Dairy Sci. 99(9):6855-6873.

Ptak, E., P. Brzozowski, and J. Bieniek. 2012. Genetic parameters for lactose percentage in the milk of Polish Holstein-Friesians. J. Anim.Feed Sci. 21(2):251-262.

Radostits, O., C. C. Gay, K. Hinchcliff, P. Constable, D. Jacobs, B. Ikede, R. McKenzie, D. Colwell, G. Osweiler, and R. Bildfell. 2007. Veteriary medicine: a textbook of the diseases of cattle, sheep, pigs, goats and horses.

Rutten, M., H. Bovenhuis, K. Hettinga, H. van Valenberg, and J. Van Arendonk. 2009. Predicting bovine milk fat composition using infrared spectroscopy based on milk samples collected in winter and summer. J. Dairy Sci. 92(12):6202-6209.

Schaeffer, L. 1984. Sire and cow evaluation under multiple trait models. J. Dairy Sci. 67(7):1567-1580.

Seifi, H. A., S. J. LeBlanc, K. E. Leslie, and T. F. Duffield. 2011. Metabolic predictors of post-partum disease and culling risk in dairy cattle. The Veterinary Journal 188(2):216-220.

Shadparvar, A. and M. Yazdanshenas. 2005. Genetic parameters of milk yield and milk fat percentage test day records of Iranian Holstein cows. Asian-Aust. J. Anim. Sci 18(9):1231-1236.

Shetty, N., P. Løvendahl, M. Lund, and A. Buitenhuis. 2017. Prediction and validation of residual feed intake and dry matter intake in Danish lactating dairy cows using mid-infrared spectroscopy of milk. J. Dairy Sci. 100(1):253-264.

Soyeurt, H., I. Misztal, and N. Gengler. 2010. Genetic variability of milk components based on mid-infrared spectral data. J. Dairy Sci. 93(4):1722-1728.

Sun, C., P. Madsen, M. Lund, Y. Zhang, U. Nielsen, and G. Su. 2010. Improvement in genetic evaluation of female fertility in dairy cattle using multiple-trait models including milk production traits. J.Anim. Sci. 88(3):871-878.

Suthar, V., J. Canelas-Raposo, A. Deniz, and W. Heuwieser. 2013. Prevalence of subclinical ketosis and relationships with postpartum diseases in European dairy cows. J. Dairy Sci. 96(5):2925-2938.

Thompson, R. and K. Meyer. 1986. A review of theoretical aspects in the estimation of breeding values for multi-trait selection. Livest. Prod. Sci. 15(4):299-313.

Van der Drift, S., R. Jorritsma, J. Schonewille, H. Knijn, and J. Stegeman. 2012a. Routine detection of hyperketonemia in dairy cows using Fourier transform infrared spectroscopy analysis of β-hydroxybutyrate and acetone in milk in combination with test-day information. J. Dairy Sci. 95(9):4886-4898.

Van der Drift, S., K. Van Hulzen, T. Teweldemedhn, R. Jorritsma, M. Nielen, and H. Heuven. 2012b. Genetic and nongenetic variation in plasma and milk β-hydroxybutyrate and milk acetone concentrations of early-lactation dairy cows. J. Dairy Sci. 95(11):6781-6787.

Van der Werf, J., J. Van Arendonk, and A. De Vries. 1992. Improving selection of pigs using correlated characters.

Van Knegsel, A., S. van der Drift, M. Horneman, A. de Roos, B. Kemp, and E. Graat. 2010. Short communication: Ketone body concentration in milk determined by Fourier transform infrared spectroscopy: Value for the detection of hyperketonemia in dairy cows. J. Dairy Sci. 93(7):3065-3069.

Vanlierde, A., M.-L. Vanrobays, F. Dehareng, E. Froidmont, H. Soyeurt, S. McParland, E. Lewis, M. H. Deighton, F. Grandl, and M. Kreuzer. 2015. Hot topic: Innovative lactation-stage-

dependent prediction of methane emissions from milk mid-infrared spectra. J. Dairy Sci. 98(8):5740-5747.

Varmuza, K. and P. Filzmoser. 2009. Introduction to multivariate statistical analysis in chemometrics. CRC press, New York.

Walsh, R., J. Walton, D. Kelton, S. LeBlanc, K. Leslie, and T. Duffield. 2007. The effect of subclinical ketosis in early lactation on reproductive performance of postpartum dairy cows. J. Dairy Sci. 90(6):2788-2796.

Youssef, M. A., S. A. El-Khodery, W. M. El-deeb, and W. E. A. El-Amaiem. 2010. Ketosis in buffalo (Bubalus bubalis): clinical findings and the associated oxidative stress level. Trop.Anim.Health Prod. 42(8):1771-1777.

Zhang, Z., G. Liu, H. Wang, X. Li, and Z. Wang. 2012. Detection of subclinical ketosis in dairy cows. Pak Vet J 32(2):156-160.

# Paper I

**Comparison of single-trait and multi-trait animal models for genetic evaluation of milk production traits predicted from milk infrared spectra in Norwegian dairy cattle**

Tesfaye Kebede Belay, Morten Svendsen and Tormod Ådnøy

*Under resubmission to Acta Agriculturae Scandinavica, Section A - Animal Science*

**Comparison of single-trait and multi-trait animal models for genetic evaluation of milk production traits predicted from milk infrared spectra in Norwegian dairy cattle**

Tesfaye Kebede Belay[1*], Morten Svendsen[1, 2] & Tormod Ådnøy[1]

[1] Department of Animal and Aquacultural Sciences, Norwegian University of Life Sciences, P. O. Box 5003, 1432 ÅS, Norway

[2] GENO Breeding and A.I. Association, P. O. Box 5025, N-1432 Ås, Norway

**Abstract**

The objective of this study was to compare single-trait (ST) and multi-trait (MT) animal models and to evaluate the practical usefulness of MT models compared to ST models. The methods were applied to 875,460 test-day records from Norwegian dairy cattle. Data were analyzed using repeatability test-day animal models. Traits studied include milk, fat and protein yields, and fat and protein contents. Estimates of heritability were very similar in both analyses. The MT analyses improved accuracies of estimated breeding values (EBV) for cows and sires with < 50 daughters, but they were similar for sires with >50 daughters. Sires were less re-ranked compared to cows' rank correlation. Estimated genetic trends were slightly higher under MT for protein and fat contents, and for fat yield, but slightly lower for the remaining traits. In conclusion, small differences were found between the estimates obtained by the ST and MT animal models for this population.

**Keywords:** EBV accuracy; genetic trend; heritability; rank correlation

* Corresponding authors: tesfaye.kebede.belay@nmbu.no

## Introduction

In animal breeding, traits can be analyzed either univariately or multivariately. It may be challenging to estimate (co)variance components of additive genetic and other random effects needed to run the models, especially when more than 3-4 traits from a large number of animals with many records each are analyzed multivariately. Therefore, traits are often analyzed univariately or in subgroups. Information from correlated traits may improve the accuracy of predictions for a particular trait if covariance are correct (Schaeffer, 1984).

The breeding program for Norwegian Red (**NRF**) depends on multi-trait selection (total merit index). Predicted breeding values for individual traits in the breeding goal are combined based on relative weights for the traits. The breeding goal comprises traits with low and high heritability. Genetic parameters and predicted breeding values for the traits are obtained from single-trait (**ST**) animal models. The precision of predicted breeding values from ST models may be low when the traits have low heritability and number of records are small (Guo *et al.*, 2014). A way to increase the accuracy of estimated breeding values for traits in the breeding goal is to use multivariate methods that enable full exploitation of the data and to combine direct with indirect information from correlated traits (Pollak *et al.*, 1984, Thompson and Meyer, 1986, Van der Werf *et al.*, 1992).

In addition to improving the accuracy of estimated breeding values (EBV), multi-trait (MT) analyses give less biased estimates when animals have been selected based on values of another correlated trait (Henderson, 1975, Pollak *et al.*, 1984). A model including information on the correlated trait is able to correct for this type of selection. The fact that the main selection in dairy cattle for many years has been on production may lead to biased predicted breeding values when ST estimation is used since genetic correlations are not included in the breeding value prediction. This will show up as bias in estimated genetic trends (Pollak *et al.*, 1984). Despite the expected

importance of MT genetic evaluation, this has often not been done in breeding schemes. There is for example no published work on Norwegian dairy cattle data using multi-trait mixed models. In this paper, we will compare the genetic evaluation of Norwegian dairy cattle under ST with MT animal models to see if there is a benefit from MT evaluation.

On the other hand, MT genetic evaluation is computationally demanding (even not feasible if all traits in the breeding goal are to be analyzed together) due to the increasing number of equations to be solved (Ducrocq, 1994, Konstantinov and Erasmus, 1993). A second drawback of MT evaluation is that the correlations used are only estimates, – often estimated with high uncertainty. Studies show that response to selection may depend highly on the precision of the variance components applied (Villanueva *et al.*, 1993). Under such conditions, univariate models may provide more precise estimates than MT models.

Genetic evaluations of dairy cattle in Norway have been based on 305-d lactation records of milk, fat, and protein yields. Within the dairy herd recording system, daily milk yields are recorded monthly while milk content and other traits are measured either monthly or bimonthly. These recordings are subsequently aggregated into a measure of lactation yields, and the individual test-day (**TD**) records are weighted by duration of testing periods to obtain the 305-d yields. Use of original TD records instead of the aggregated lactation yields could help to improve accuracy of EBV estimation (Ptak and Schaeffer, 1993, Kettunen *et al.*, 2000) and to provide more comprehensive management information to farmer (Kettunen *et al.*, 2000). Other advantages of TD records for genetic evaluation of dairy cattle over the lactation records are described in the literature (Ptak and Schaeffer, 1993, Bilal and Khan, 2009, Dzomba *et al.*, 2010).

Various TD models have been used to analyze TD records in genetic evaluation of dairy cattle and all the models have their advantages and disadvantages. Type of TD model to use might depend

on the objective of the study, number of TD records per lactation, and computing environments. In the present study, the simplest TD model, the repeatability TD model was applied assuming constant genetic variance throughout lactation and unity genetic correlations among TD records. TD records within and across lactation were considered repeated measurement of the same trait. This study compares ST and MT repeatability TD animal models for estimation of genetic parameters, prediction of breeding values and genetic trends of TD milk, fat, and protein yields, and milk fat and protein contents. Our intention is not to fit models that best suit the TD records, but primarily to compare the ST and MT genetic evaluations.

## Materials and methods

### *Data and data edits*

Norwegian dairy herd recording system TD data on milk, fat, and protein yields, and milk fat and protein contents collected in 2007 to 2013 were used in this study. The milk samples had been analyzed by a Fourier-transformed mid-infrared spectrometer (Milkoscan$^{TM}$ Combifoss 6500, Foss Electric, Hillerød, Denmark) to find milk fat and protein contents. The fat and protein yields were calculated by multiplying the respective milk contents with the observed TD milk yield.

Original size of the TD data was 5,301,687 records from 638,743 cows kept in 13,100 herds and daughters of 2088 sires. Multivariate genetic covariance estimation with such a number of records is very difficult (even impossible) with the currently available statistical packages and computer hardware. Therefore, certain criteria were implemented to reduce the data into manageable size. Cows with unknown sires or dams, herds with less than 1250 TD records and sires with less than 10 daughters were excluded from the dataset. Moreover, only cows with age at calving of 18 to 39, 30 to 51, 42 to 63 and 52 to 74 months in the first, second, third and fourth lactations, respectively, were considered. Number of records per herd test*date (**HTD**) were kept to at least

two. The final edited dataset contained 875,460 TD records from 91,186 cows that were progeny of 1282 sires and kept in 529 herds. A pedigree file containing animals with record and their ancestors was also available and the total number of animals in the pedigree file that had a link to the data file was 197,497.

***Statistical analysis***

In a preliminary step, the PROC GLM procedure of SAS was used to study the statistical significance of fixed factors. The factors were lactation stage (between 10 and 320 days in milk categorized into 10 days intervals with 32 levels), lactation number (parity 1 to 4), region (9 levels), year * month of test, and herd * year of test (HY). Cows were grouped by age at calving (in months) within lactation numbers. There were 15 possible age categories (classes) within each lactation. To get reasonable numbers of records for age classes at the peripheries in each age category, the first and the last few age classes were merged into the next and preceding age class, respectively. Furthermore, housing type (tie stall or loose housing) and milking system (robot milking or manual milking: bucket, pipe or milking parlor) were modelled. Beside the main factors, interactions among region*year*month of test (RYM), region*parity*lactation stage (RPS), parity*age at calving (PA) and housing*milking system (HMS) were fitted. All factors and interactions were found to be significant ($P<0.01$) and explained at least 60% of variation in all response variables. Interaction terms such as RYM (743 levels), RPS (1116 levels), HY (3633 levels), PA (60 levels) and HMS (4 levels) were retained in the final mixed model used for genetic analyses. Random factors included in the model were animal additive genetic, permanent animal environment (**PE**), HTD (28,256 levels) and residual effects.

(Co)variance components were estimated using ST and MT repeatability TD animal models with the restricted maximum likelihood (REML) method using the parameter expanded and average information (PX-AI) algorithm of the software WOMBAT (Meyer, 2007). Preliminary bivariate analyses were performed, and covariance component estimates from such analyses were pooled using the iterative summing of expanded part matrices approach (Mäntysaari, 1999) implemented in WOMBAT (Meyer, 2007). The pooled covariance matrix was used as prior to estimate covariance components in MT mixed model analysis.

The ST and MT repeatability TD animal models used in this study in matrix notation:

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Za} + \mathbf{Wp} + \mathbf{Hd} + \mathbf{e}$$

where, **y** is the vector of TD records for milk, fat and protein yields, and milk fat and protein contents; **b** is a vector of fixed effects of RYM, RPS, HY, PA, and HMS; **a** is a vector of random animals' additive genetic effects; **p** is a vector of random PE effects due to the cow; **d** is a vector of random HTD effects; **e** is a vector of random residual effects. **X, Z, W** and **H** are design matrices that relate records to the corresponding effects.

Assumed variance structure in the ST analysis was:$\mathrm{var(a)} = \sigma_{\mathrm{a}}^2\mathbf{A}$, $\mathrm{var(p)} = \sigma_{\mathrm{pe}}^2\mathbf{I}$, $\mathrm{var(d)} = \sigma_{\mathrm{d}}^2\mathbf{I}$ and $\mathrm{var(e)} = \sigma_{\mathrm{e}}^2\mathbf{I}$ , where $\sigma_{\mathrm{a}}^2$ is additive genetic variance, $\sigma_{\mathrm{pe}}^2$ is PE variance, $\sigma_{\mathrm{d}}^2$ is HTD variance, and $\sigma_e^2$ is residual variance. The **I** are identity matrices of appropriate sizes and **A** is additive relationship matrix. For the MT analyses, the following covariance structures were assumed:

$$\mathrm{var}\begin{bmatrix}\mathbf{a}\\ \mathbf{p}\\ \mathbf{d}\\ \mathbf{e}\end{bmatrix} = \begin{bmatrix}\mathbf{G}\otimes\mathbf{A} & 0 & 0 & 0\\ 0 & \mathbf{P}\otimes\mathbf{I} & 0 & 0\\ 0 & 0 & \mathbf{H}\otimes\mathbf{I} & 0\\ 0 & 0 & 0 & \mathbf{R}\otimes\mathbf{I}\end{bmatrix}$$

where, **G** is the genetic covariance matrix, **P** is the PE covariance matrix, **H** is the covariance matrix for HTD effects, **R** is the residual covariance matrix, **I** and **A** are as defined above, and $\otimes$ is the Kronecker product.

Heritability was calculated as $h^2 = \frac{\sigma_a^2}{\sigma_a^2+\sigma_{pe}^2+\sigma_d^2+\sigma_e^2}$. The proportion of variance due to PE cow effect was calculated as $c^2 = \frac{\sigma_{pe}^2}{\sigma_a^2+\sigma_{pe}^2+\sigma_d^2+\sigma_e^2}$. The proportion of variance due to HTD was calculated as $d^2 = \frac{\sigma_d^2}{\sigma_a^2+\sigma_{pe}^2+\sigma_d^2+\sigma_e^2}$. Standard errors (SE) for the variance components and variance ratios were estimated from the inverse of average information matrix whereas SE for the fixed and random effects were estimated with one additional iteration (after convergence) using EM-REML algorithm. With this algorithm, WOMBAT calculates the inverse of the coefficient matrix in the mixed model equations to obtain solutions for all fixed and random effects fitted, and provides corresponding SE, obtained from the diagonal elements of the inverse coefficient matrix. The approximate accuracy (r) of EBV for the *i*$^{th}$ individual was calculated from the SE of the EBV as: $r_i = \sqrt{1 - \frac{SE_i^2}{\sigma_a^2}}$.

To assess re-ranking of individuals, Spearman rank correlation analysis was conducted using animal rankings based on EBV obtained from the ST and MT models. The rank correlation analysis was done either for all animals in pedigree, for cows with records, or for sires having daughters with records. Genetic trends were determined by averaging the EBV of cows obtained from the ST and MT analysis by birth year of cows.

## Results

### *Descriptive statistics*

Table 1 summarizes descriptive statistics for the studied traits. Yield traits had a much higher coefficient of variation than the milk content traits. Mean and standard deviation of milk protein yield and content were lower than for the corresponding values for fat.

### *Estimates of genetic parameters*

Estimates of variance components and the corresponding variance ratios for the milk production traits from ST and MT models are presented in Table 2 and were similar in both types of analyses. The milk production traits had moderate heritability estimates ranging from 0.119 to 0.377 in ST models and 0.121 to 0.376 in MT models. Under both types of analyses, heritability estimates of milk fat and protein contents were higher than estimates for the corresponding yield traits. Approximate standard errors of the heritability estimates were similar in ST and MT analyses and ranged from 0.003 for milk fat content to 0.005 for milk protein content, protein yield and milk yield.

Estimates of variance ratios of HTD and PE to total variance effects were very similar in both types of analyses for the same traits. The proportion of phenotypic variance attributed to the HTD effects ranged from 6% for milk yield to 13% for milk fat content (Table 2). The effect of HTD on milk fat content (13%) was comparable with the animal additive genetic effect (16%) for the same trait, suggesting that, in addition to genetics, management (e.g. feeding) is an important factor influencing milk fat content. For yield traits, PE effects were larger than additive genetic effects, but for the milk content traits, additive genetic effects were larger than the PE effects.

Estimates of genetic, PE, HTD and residual correlations among the milk production traits from bivariate analyses are presented in Table 3. These correlations were also estimated using MT model where three traits (milk with fat and protein yields: MT-Y model, or milk with fat and protein contents: MT-P model) were analyzed simultaneously. The estimates from such analyses (results not shown) were very similar to the ones from the bivariate analyses. Genetic correlations among yield traits were high positive, ranging from 0.772 (between milk and fat yields) to 0.869 (between milk and protein yields). Genetic correlation between milk fat and protein contents was moderate positive (0.684) whereas the genetic correlations between yield and composition traits were low to moderate and ranged from -0.489 (between milk yield and its protein content) to 0.156 (between fat yield and milk fat content). Among the yield traits, phenotypic, genetic and residual effects were correlated in similar magnitude and in the same direction.

PE correlations among yield traits were positive and very high (0.937 to 0.955) whereas the PE correlation between milk fat and protein contents was moderate (0.315). Permanent environmental correlations between yield and composition traits were moderately low except for correlations of fat content with fat yield (0.068) and protein yield (-0.082). Unlike the other correlations, HTD correlation of fat content with fat yield was moderately high (0.710). The standard errors of the genetic correlation estimates ranged from 0.001 to 0.015, and a similar range of standard errors was found for the correlation estimates of PE effects. The standard errors for the correlation estimates of residual effects were from 0.000 to 0.008 and for HTD effects from 0.001 to 0.008.

***Distribution of estimated breeding values***

Estimated breeding values from ST and MT analyses were compared for each of the milk production traits studied. Standard deviations and range of EBVs for all animals in pedigree, and for sires having daughters with records are given in Table 4 for ST and MT analyses. For all

animals and sires, the standard deviation and the range (minimum and maximum) of EBV from both ST and MT models were similar, except the slightly lower estimates from ST than MT-Y analyses for milk yield (when milk was analyzed with fat and protein yields). In general, the ranges of EBV for yield traits (milk, fat and protein yields) were slightly higher for MT than ST analyses. As an example, a graphical distribution of milk EBV from the ST, MT-Y, and MT-P models for all animals in pedigree are given in Figure 1. It shows that EBV for milk yield from MT-Y model had a wider variation than those from ST and MT-P models. Plots of EBV for milk yield from ST and MT-P models overlapped indicating similarities among estimates. Similar distributions are given in the form of ranges for other milk production traits in Table 4.

***Accuracy of EBV***

Table 5 show approximate accuracy (r) of EBV for sires categorized into four groups based on numbers of daughters, and for cows with records. The average accuracies of sire EBVs were high, ranging from 0.852 to 0.996, and increased with number of daughters per sire. Estimates of accuracy of sire EBV from MT model were very similar to the ones from ST model, except for protein yield for sires with 10-49 daughters where accuracy increased from 0.861 in ST to 0.889 in MT (3.25% improvement). Milk EBV accuracy for sire from MT-Y was similar to the one from MT-P, but with a little improvement in MT-Y over MT-P (0.10 to 0.69% improvement). Accuracies of cows EBVs were much lower than those of the sires and ranged from 0.69 for fat yield to 0.86 for milk protein content. Accuracies of cows EBV were higher in MT than ST analysis. The highest improvement in accuracy of cows EBV due to MT was obtained for protein yield (9.83%), followed by milk yield (4.58% in MT-Y or 2.50% in MT-P) and fat content (3.06%). For the remaining traits, accuracies of cows' EBVs were very similar in both types of analyses.

Milk EBV accuracy for cow from MT-Y was higher than the one from MT-P (2.03% improvement).

***Rank and EBV correlations***

EBV were predicted for each of the 197,498 animals in pedigree for all traits with ST and MT (MT-Y or MT-P) analyses. Correlations between ST and MT EBVs and rank correlations based on animal ranks using the same EBVs are presented in Table 6. The EBV correlation and rank correlation were also calculated for the 91,186 cows with records and the 1,282 sires having daughters with records.

For all animals in pedigree and for cows with records, rank (or EBV) correlations were significantly ($P$<0.05) different from one at least for milk yield (range 0.792-0.968 for rank and 0.820-0.974 for EBV), protein yield (range 0.776-0.803 for rank and 0.81 for EBV) and milk fat content (range 0.955-0.961 for rank and 0.96 for EBV). For fat yield, there was some re-ranking for all animals in pedigree and for cows with records as the rank correlation was about 0.989. For milk protein content some re-ranking of animals and cows were observed by going from ST to MT analysis as correlations were high (0.993-0.995 for rank and 0.994-0.995 for EBV). For sires, rank (or EBV) correlations were not as severely affected as for all animals in pedigree and cows with records. Sire's correlations were very close to one (0.992-0.999), except for protein and milk yields where the correlations between ST and MT-Y were far from one (0.854-0.920). Therefore, except for the protein and milk (from MT-Y) yields, re-ranking of sires was limited. For all three subsets of animals, rank (or EBV) correlations were highest for protein content, but lowest for protein yield, and rank correlations were lower than the respective EBV correlations for each trait. Rank (or EBV) correlations of milk yield found with ST and MT-Y models were much lower than the corresponding correlations from ST and MT-P models.

***Genetic trends***

Average EBV per year for cows born in 2001 through 2011 are given in Figure 2 for milk yield; in Figure 3 for fat and protein yields; and in Figure 4 for milk fat and protein contents. Regardless of the type of analysis, average EBV of milk yield decreased for cows born from 2001 to 2004, where the lowest average estimate of EBV was obtained. After 2004, an increasing trend was observed in general (Figure 2). Estimates of genetic trend under ST model were slightly higher than the corresponding average estimates under MT-P model, with noticeable differences between the ST and MT-P EBV after 2005. For all years studied, however, estimates of genetic trend from MT-Y were lower ($P<0.0001$) than estimates from ST and MT-P models, but had the same genetic pattern as in ST and MT-P models (Figure 2).

Estimates of genetic trend for fat yield under MT was slightly higher ($P<0.0001$) than the values estimated under ST models, but the reverse was true for protein yield (Figure 3). In both types of analyses, like in milk yield, average EBV of cows for fat and protein yields decreased in the first four years, but had increasing trend for cows born in later years (Figure 3).

Figure 4 shows similar genetic trend for milk fat content or protein content under both models, but slightly higher in MT than ST models. Regardless of the type of analysis, average EBV of fat and protein contents were decreasing in the first few years, but generally showed increasing trend for cows born from year 2004 onwards (Figure 4).

**Discussion**

Single-trait and MT genetic evaluation of milk production traits were investigated in this study. Accuracy of EBV is an important parameter in livestock genetic improvement. It is used to calculate response to selection (Falconer and Mackay, 1996) and to express the credibility of individual EBV, and it is related to the risk that this EBV will change over time when more information becomes available (Bijma, 2012). The MT analyses should improve accuracy of EBV for each trait involved, at least theoretically, by reducing prediction error variances (**PEV**) as they add information from correlated traits (Schaeffer, 1984). From quantitative genetic theory, the accuracy of EBV for a trait with low heritability can be improved using MT models when correlated traits that have relatively high heritability are included. The benefit obtained from MT analysis (improvement in accuracy of EBV) depends on the correlation structure and the differences in heritability between the traits involved in the model (Falconer, 1996, Schaeffer, 1984). It also depends on the absolute difference between error and genetic correlations of two traits (Schaeffer, 1984). The greater the absolute difference in correlations, the greater is the reduction of PEV (and hence increase in accuracy) for both traits. When the error correlation is less (greater) than the genetic correlation, traits with lower (higher) heritability achieve a greater increase in accuracy (Schaeffer, 1984). In the present study, the estimates of heritability were moderate both for yield traits (0.121 to 0.201) and for composition traits (0.157 to 0.377). Moreover, there were low to moderate genetic correlations (-0.489 to 0.156) and error correlations (-0.319 to 0.597) between yield and composition traits. Genetic correlations either among yield traits or among composition traits were similar to the corresponding error correlations. These parameter estimates indicate that the accuracy of EBV for milk production traits would not increase or only slightly increase in a MT model including more milk production traits (Table 6).

Accuracies of EBV also depend on number of observations (amount of information) used for prediction of breeding values. Regardless of the type of analysis, EBV accuracy increased with increase in number of daughters per sire in all traits. With smaller numbers of observations (e.g., for sires with <50 daughters, and cows), traits with high heritability had better accuracy than those with relatively lower heritability. This difference in accuracy due to difference in heritability diminished as the number of daughters per sire increased. Multi-trait models would be useful when number of observations are small (Guo et al., 2014; Table 5). For example, for sires with less than fifty daughters, EBV from MT evaluation tend to gain more in accuracy (3.25% improvement) compared to EBV accuracy from ST analysis (e.g., for protein yield, Table 5). However, with large numbers of daughters per sire, EBV accuracies in both ST and MT models were very similar. Cows (with direct information from own production and with few daughters) had accuracies even lower than sires with less than fifty daughters.

Cow genetic evaluations gain more from multi-trait analysis than sire evaluations. For cows, improvement of EBV accuracy was larger as additional information provided by other traits was added. The highest improvement in accuracy of cows EBV was obtained for protein yield (9.83%), followed by milk yield (4.58 in MT-Y or 2.50% in MT-P) and milk fat content (3.06%). For sires, very little improvement in accuracy due to MT analysis compared to ST was obtained, except for protein yield where accuracy was improved by 3.25%. Gengler and Coenraets (1997) also observed higher relative improvement due to MT analysis for cows than for sires in lactation yields of dairy cattle in Belgium. The relative improvement in accuracy found in the current study in milk yield for both cows and sires were similar to their results, but that of protein yield was higher here than reported by Gengler and Coenraets (1997).

Another advantage of MT models may be reduction or elimination of selection bias for the trait of interest, if correlated traits on which the selection was based are included in the analysis (Pollak et al., 1984; Schaeffer, 1984). Among the traits investigated in this study, the yield traits (milk, fat and protein yields), in addition to fertility and disease traits, are part of the NRF breeding goals – i.e. milk fat and protein contents are not included. Selection is currently done on milk, fat and protein yields, so ST genetic evaluation of either milk fat or milk protein content might suffer from selection bias and the resulting EBV could lead to less effective selection of future parents. Multi-trait genetic evaluation of traits outside the breeding goal (e.g. milk fat and protein contents) with correlated traits (e.g. yield traits) would correct for possible selection bias (if any). However, there was no or little difference in parameter estimates for the composition traits between ST and bivariate (or MT) analyses in this study. This is probably because composition traits have been indirectly considered in selection through protein or fat yields, as they are a result of milk yield and respective percentages.

In animal breeding programs, the main purpose of breeding value prediction is ranking animals in order to select superior individuals of the current generation that will be used as parents of next generations. The rank (or EBV) correlations between ST and MT sire evaluations were very close to unity, except for protein and milk yields in MT-Y model. The rank (or EBV) correlations between ST and MT calculations for cows (or all animal) evaluations differed from unity for milk, fat and protein yields, and fat content. Therefore, genetic evaluation of these milk production traits using an ST model or an MT model could lead to different ranking of candidates. Re-ranking of cows or all animals for milk protein content would be limited as correlations were close to one. For fat yield, rank correlations found in this study (0.9890 for cows or 0.9917 for sires) was similar to values reported by Gengler and Coenraets (1997) who found rank correlation of 0.981 for cows

or 0.991 for sires between ST and MT (milk, protein and fat yields) evaluations. Correlation estimates reported for protein and milk yields by Gengler and Coenraets (1997) were higher than values obtained in this study while rank correlations between ST and MT-P model for milk yield for sires ranking were similar to the report of Gengler and Coenraets (1997).

There were two sets of rank (or EBV) correlations for milk yield: between ST and MT-Y, and between ST and MT-P evaluations. For all the three subsets of animals, the correlations between ST and MT-Y were far from unity and much lower than the correlations between ST and MT-P models. In addition to rank (or EBV) correlations, genetic trends were lower in MT-Y than in ST model for protein or in both ST and MT-P for milk yield. This could be related to the heritabilities and genetic correlation structures of the traits involved in the MT analysis. In an MT-P model, heritabilities of milk fat and protein contents were similar to or higher than the heritability of milk yield (Table 2) that were moderately correlated with these composition traits (Table 3). Because of this similarity in heritabilities and moderate genetic correlations, information gained from correlated traits by milk yield would be very limited. Hence, genetic evaluation of milk yield in MT-P would be very similar to the evaluation of milk yield from ST model. In the MT-Y evaluation, fat and protein yields had lower heritabilities than that of the milk yield, which had strong positive genetic correlations with fat (0.772) and protein (0.869) yields. Fat yield had the lowest heritability among yield traits and was strongly correlated with protein yield (0.836). In this case, milk yield and possibly protein yield would give information to traits with low heritability in the model (e.g. fat yield). As a result, genetic evaluation of milk and protein yields in MT-Y models would be lower than in the ST evaluation whereas genetic evaluation of fat yield would be similar in both analyses, or higher in MT than in ST evaluation.

We have also analyzed the traits bivariately to investigate if the difference in rank (or EBV) correlations and genetic trend observed between ST and MT-Y or MT-P would be repeated between ST and bivariate evaluations. In such analyses, average rank (or EBV) correlations were very close to unity for milk (0.9978), fat (0.9975) and protein (0.9988) yields, indicating limited re-ranking of sires. This is in contrast to the re-ranking of sires observed for milk and protein yields when all three yield-traits were analyzed together (MT-Y). For all traits and the three subsets of animals, estimates from bivariate analysis (rank (or EBV) correlations, genetic trends, annual genetic change and accuracy of EBVs) were slightly higher than estimates from ST or MT analysis. Therefore, given the computational cost and modelling complexity associated with multivariate analysis, ST or bivariate analysis of the milk production traits especially for the yield traits could be a good strategy for genetic evaluation of Norwegian dairy cattle.

**Conclusions**

Comparison of the ST and MT animal models applied for the genetic evaluation of NRF revealed small differences between the estimates (variance components, heritability, EBV, accuracy of EBV, rank (or EBV) correlations, and genetic trend obtained by the methods. Variances due to HTD effects was found to be an important source of variation for milk fat content under both types of analyses while effects of the PE were larger than the additive genetic effects for yield traits. Genetic and phenotypic correlations among the yield traits were high positive whereas these values were low to moderate positive or negative between yield traits and composition traits. Accuracies of EBV was high for sires (0.852 to 0.996), increased with daughter number, and was similar in both models except some noticeable improvement in MT (3.06%) for protein yield for sires in the 10-49 daughter group. Accuracies of EBV for cows was higher with MT than ST: a 2.5% to 9.8% improvement in accuracy due to using MT models was found. Re-rankings of sires based on EBV

were very limited while for cows or all animals in pedigree some re-ranking was found. The three subsets of animals were highly re-ranked for protein and milk yield as the rank correlations between ST and MT-Y were far from unity (0.78 to 0.91). Annual genetic trends (Figures 2-4) were slightly higher under MT analyses for protein content, fat yield and fat content, but slightly lower for the remaining traits. Considering computing demands, it could be a good strategy for genetic evaluation of milk production traits in Norwegian Red to use a single-trait or bivariate models.

**Acknowledgement**

The authors gratefully acknowledge TINE SA, the Norwegian dairy association that manages the central database, for providing data for this work.

**References**


Bijma, P. (2012) Accuracies of estimated breeding values from ordinary genetic evaluations do not reflect the correlation between true and estimated breeding values in selected populations. *J. Anim.Breed.Genet.,* 129**,** 345-358.

Bilal, G., Khan, M. (2009) Use of test-day milk yield for genetic evaluation in dairy cattle: a review. *Pak Vet J.,* 29**,** 35-41.

Ducrocq, V.(1994) Multiple trait prediction: principles and problems. In: Proceedings of the 5th World Congress on Genetics Applied to Livestock Production, 1994. 455-462.

Dzomba, E., Nephawe, K., Maiwashe, A., Cloete, S., Chimonyo, M., Banga, C., Muller, C., Dzama, K. (2010) Random regression test-day model for the analysis of dairy cattle production data in South Africa: Creating the framework. *S.Afr J. Anim. Sci.,* 40**,** 273-284.

Falconer, D. S., Mackay, T. F. (1996) Introduction to quantitative genetics. *Harlow: Addison Wesley Longman*.

Gengler, N., Coenraets, P. (1997) Use of multiple-trait animal models for genetic evaluation of milk, fat and protein lactation yields of dairy cattle in Belgium. *Biotechnol.Agron. Soc.Environ.,* 1**,** 26-33.

Guo, G., Zhao, F., Wang, Y., Zhang, Y., Du, L., Su, G. (2014) Comparison of single-trait and multiple-trait genomic prediction models. *BMC genetics,* 15**,** 1.

Henderson, C. R. (1975) Best linear unbiased estimation and prediction under a selection model. *Biometrics***,** 423-447.

Kettunen, A., Mäntysaari, E. A., Pösö, J. (2000) Estimation of genetic parameters for daily milk yield of primiparous Ayrshire cows by random regression test-day models. *Livest. Prod.Sci.,*66**,** 251-261.

Konstantinov, K., Erasmus, G. (1993) Using transformation algorithms to estimate (co) variance components by REML in models with equal design matrices. *S. Afr J.Anim Sci., 23(5/6)*.

Meyer, K. (2007) WOMBAT—A tool for mixed model analyses in quantitative genetics by restricted maximum likelihood (REML). *J.Zhejiang Univ.Sci.B.,* 8**,** 815-821.

Mäntysaari, E. (1999) Derivation of Multiple Trait Reduced Random Regression (RR) Model for the First Lactation Test Day Records of Milk, Protein and Fat. *50th Annual Meeting. Europ. Ass. Anim. Prod. Mimeo.*

Pollak, E., Van Der Werf, J., Quaas, R. (1984) Selection bias and multiple trait evaluation. *J.Dairy Sci.,* 67**,** 1590-1595.

Ptak, E., Schaeffer, L. (1993) Use of test day yields for genetic evaluation of dairy sires and cows. *Livest. Prod. Sci.,* 34**,** 23-34.

Thompson, R., Meyer, K. (1986) A review of theoretical aspects in the estimation of breeding values for multi-trait selection. *Livest. Prod.Sci.,* 15**,** 299-313.

Van Der Werf, J., Van Arendonk, J., De Vries, A. (1992) Improving selection of pigs using correlated characters. Abstract 43rd Ann. Meeting EAAP, Madrid, Vol. 1 (1992) 14.

Villanueva, B., Wray, N., Thompson, R. (1993) Prediction of asymptotic rates of response from selection on multiple traits using univariate and multivariate best linear unbiased predictors. *Anim.Sci.,* 57**,** 1-13.

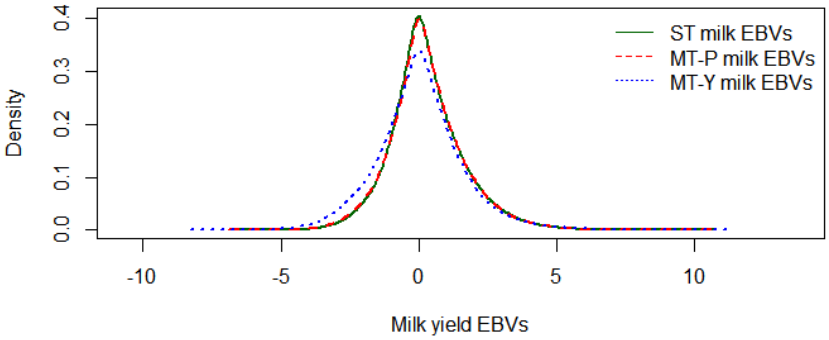
## FIGURES

**Figure 1** Comparison of distributions of estimated breeding values (EBV) of 197,497 animals from single-trait (ST milk EBVs) and multi-trait (MT-P milk EBVs and MT-Y milk EBVs) analyses for milk yield.

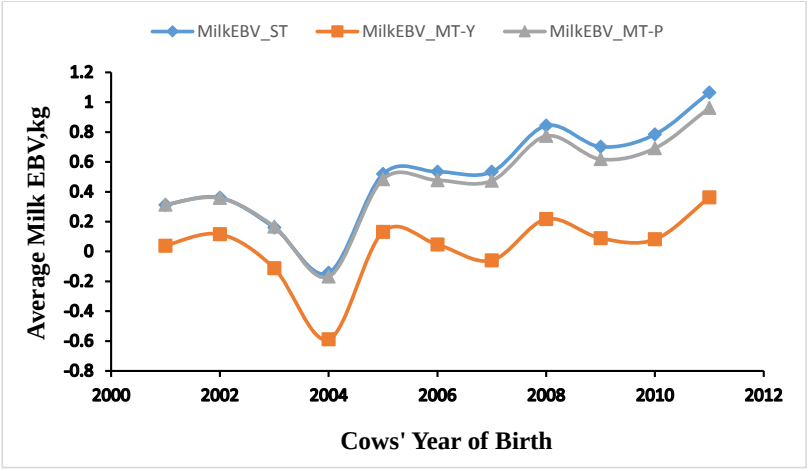
**Figure 2** Predicted breeding values for test-day milk yield averaged by birth year of cows under single-trait (MilkEBV_ST), and multi-trait (MT) models: either when milk yield was analyzed with fat and protein yields (MilkEBV_MT-Y) or with fat and protein contents (MilkEBV_MT-P).

**Figure 3** Estimated breeding values for test-day fat and protein yields averaged by birth year of cows under single-trait (ST) and multi-trait (MT) models.

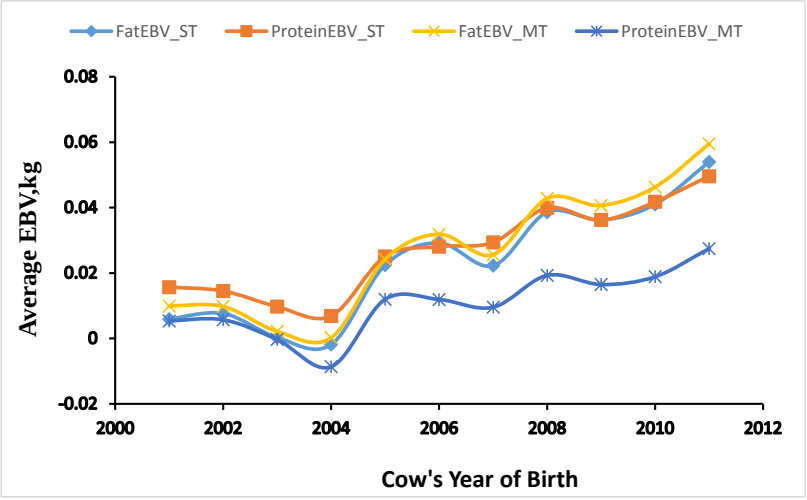
**Figure 4** Estimated breeding values for test-day milk fat and protein contents averaged by birth year of cows under the single-trait (ST) and multi-trait (MT) animal models.



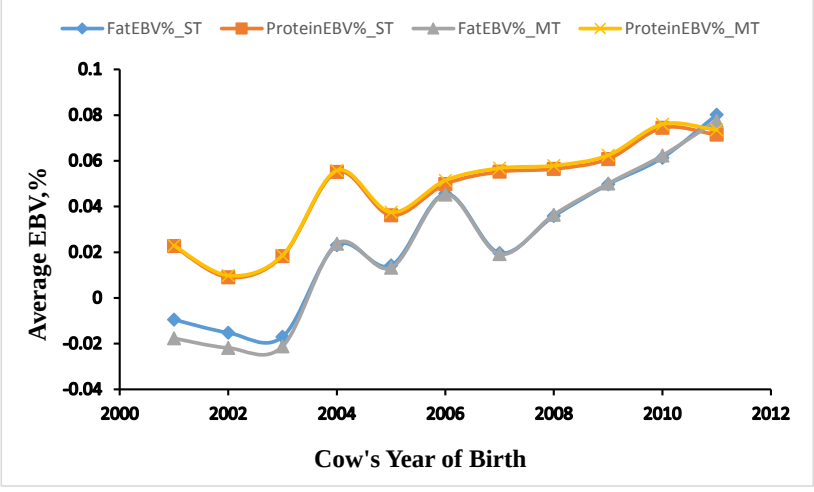

**Belay et al. Figure 1**




**Belay et al. Figure 2**




**Belay et al. Figure 3**




**Belay et al. Figure 4**

## TABLES

**Table 1** Mean, standard deviation (SD), coefficient of variation (CV, %), minimum and maximum values for milk, fat and protein yields, and fat and protein contents in milk

| Traits | Mean | SD | CV | Minimum | Maximum |
|---|---|---|---|---|---|
| Milk yield, kg | 25.86 | 8.13 | 31.44 | 5.00 | 50.00 |
| Fat yield, kg | 1.06 | 0.34 | 32.08 | 0.10 | 3.44 |
| Protein, kg | 0.88 | 0.25 | 28.41 | 0.10 | 2.39 |
| Fat, % | 4.16 | 0.75 | 18.03 | 1.75 | 7.00 |
| Protein, % | 3.45 | 0.34 | 9.86 | 1.04 | 6.98 |

**Table 2** Estimates of additive genetic($\sigma^2_a$), permanent animal environment ($\sigma^2_{pe}$), herd test-day ($\sigma^2_{htd}$) and residual ($\sigma^2_e$) variances; and variance ratios* for genetic($h^2$), permanent environment ($c^2$), herd*test-day ($d^2$) and residual ($e^2$) variances for milk, fat, and protein yields; and fat and protein contents from single-trait and multi-trait$ analyses

| Models &traits | $\sigma^2_a$ | $\sigma^2_{pe}$ | $\sigma^2_{htd}$ | $\sigma^2_e$ | $h^2$ | $c^2$ | $d^2$ |
|---|---|---|---|---|---|---|---|
| Single-trait | | | | | | | |
| Milk, kg | 6.1335 | 8.55680 | 1.67800 | 14.1850 | 0.201 | 0.280 | 0.055 |
| Fat, kg | 0.00829 | 0.01243 | 0.00616 | 0.04262 | 0.119 | 0.179 | 0.089 |
| Protein, kg | 0.00541 | 0.00867 | 0.00237 | 0.01543 | 0.170 | 0.272 | 0.074 |
| Fat, % | 0.07395 | 0.02403 | 0.05854 | 0.30928 | 0.159 | 0.052 | 0.126 |
| Protein, % | 0.02655 | 0.00598 | 0.00492 | 0.03301 | 0.377 | 0.085 | 0.070 |
| Multi-trait[1] | | | | | | | |
| Milk, kg | 6.1236 | 8.55267 | 1.67914 | 14.1866 | 0.200 | 0.280 | 0.055 |
| Fat, kg | 0.00852 | 0.01338 | 0.00614 | 0.04245 | 0.121 | 0.190 | 0.087 |
| Protein, kg | 0.00566 | 0.00943 | 0.00307 | 0.01542 | 0.169 | 0.281 | 0.092 |
| Multi-trait[2] | | | | | | | |
| Milk, kg | 6.1426 | 8.5479 | 1.6783 | 14.1865 | 0.201 | 0.280 | 0.055 |
| Fat, % | 0.07302 | 0.02374 | 0.05855 | 0.30975 | 0.157 | 0.051 | 0.126 |
| Protein, % | 0.02645 | 0.00605 | 0.00491 | 0.03301 | 0.376 | 0.086 | 0.070 |

*$h^2$with standard errors (SE) between 0.003 and 0.005; $c^2$with SE between 0.002 and 0.004; $d^2$ with SE = 0.001; $for multi-trait model, milk yield was either run [1]with fat and protein yields (MT-Y) or [2]with fat and protein contents (MT-P).

**Table 3** Estimate of genetic (above diagonal), permanent environmental (above diagonal in parenthesis), herd*test day (below diagonal) and residual (below diagonal in parenthesis) correlations among the milk production traits from bivariate analyses

| Traits | Milk, kg | Fat, kg | Protein, kg | Fat% | Protein% |
|---|---|---|---|---|---|
| Milk, kg | | 0.772(0.953) | 0.869(0.955) | -0.463(-0.211) | -0.489(-0.405) |
| Fat, kg | 0.589(0.659) | | 0.836(0.937) | 0.156(0.068) | -0.063(-0.293) |
| Protein, kg | 0.801(0.933) | 0.610(0.647) | | -0.160(-0.082) | -0.098(-0.218) |
| Fat% | -0.093(-0.137) | 0.712(0.597) | 0.0002(-0.096) | | 0.684(0.315) |
| Protein% | 0.046(-0.319) | 0.176(-0.151) | 0.364(-0.018) | 0.211 (0.187) | |

**Table 4** Standard deviations (SD), minimum (Min) and maximum (Max) of estimated breeding values from single-trait and multi-trait* evaluation of milk production traits based on all 197,497 animals in pedigree or on the 1,282 sires that had daughters with records.

| Traits | All animals in pedigree | | | | | | Sires having daughters with records | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Single-trait EBV | | | Multi-trait EBV* | | | Single-trait EBV | | | Multi-trait EBV* | | |
| | SD | Min | Max | SD | Min | Max | SD | Min | Max | SD | Min | Max |
| Milk[1], kg | 1.36 | -6.12 | +10.27 | 1.58 | -7.67 | +10.77 | 2.14 | -6.12 | +9.52 | 2.33 | -7.02 | +9.05 |
| Fat, kg | 0.05 | -0.25 | +0.31 | 0.05 | -0.24 | +0.33 | 0.08 | -0.17 | +0.28 | 0.08 | -0.18 | +0.30 |
| Protein, kg | 0.04 | -0.16 | +0.28 | 0.04 | -0.17 | +0.27 | 0.06 | -0.16 | +0.27 | 0.06 | -0.16 | +0.24 |
| Milk[2], kg | - | - | - | 1.38 | -6.37 | +10.25 | - | - | - | 2.13 | -6.25 | +9.66 |
| Fat,% | 0.16 | -0.80 | +0.10 | 0.16 | -0.85 | +0.99 | 0.22 | -0.74 | +0.74 | 0.22 | -0.74 | +0.74 |
| Protein, % | 0.11 | -0.53 | +0.93 | 0.11 | -0.52 | +0.19 | 0.13 | -0.39 | +0.51 | 0.13 | -0.39 | +0.51 |

*For multi-trait EBV, milk yield was either run [1]with fat and protein yields (MT-Y) or [2]with fat and protein contents (MT-P).

**Table 5** Accuracy* of predicted breeding values of cows and sires with different size of daughter groups for the milk production traits under single- and multi-trait models; and number of cows and sires (N) within different daughter groups

| Model | N | Yield traits (kg) | | | | Percentage traits (%) | |
|---|---|---|---|---|---|---|---|
| | | Milk[1] | Fat | Protein | Milk[2] | Fat | Protein |
| Single-trait | | | | | | | |
| Sire: | | | | | | | |
| 10-49 | 1075 | 0.872 | 0.852 | 0.861 | - | 0.902 | 0.930 |
| 50-99 | 89 | 0.940 | 0.930 | 0.934 | - | 0.958 | 0.971 |
| 100-399 | 69 | 0.976 | 0.971 | 0.973 | - | 0.983 | 0.988 |
| ≥400 | 49 | 0.992 | 0.991 | 0.991 | - | 0.995 | 0.996 |
| Cow | 91186 | 0.720 | 0.692 | 0.704 | - | 0.785 | 0.860 |
| Multi-trait | | | | | | | |
| Sire: | | | | | | | |
| 10-49 | 1075 | 0.881 | 0.854 | 0.889 | 0.875 | 0.906 | 0.930 |
| 50-99 | 89 | 0.945 | 0.930 | 0.947 | 0.941 | 0.959 | 0.971 |
| 100-399 | 69 | 0.977 | 0.971 | 0.978 | 0.976 | 0.983 | 0.988 |
| ≥400 | 49 | 0.993 | 0.991 | 0.993 | 0.992 | 0.995 | 0.996 |
| Cow | 91186 | 0.753 | 0.693 | 0.773 | 0.738 | 0.809 | 0.863 |

* Accuracy calculated as $r_i = \sqrt{1 - \frac{SE_i^2}{\sigma_a^2}}$; [1]When milk yield was analyzed with fat and protein yields in the MT model (MT-Y); [2]When milk yield was analyzed with fat and protein contents in the MT model (MT-P).

**Table 6** Correlations between single- and multi-trait estimated breeding values, and rank correlations based on animal rank using the same breeding values, for all 197,497 animals in pedigree, or 91,186 cows, or 1,282 sires

| Traits | All animals | | Cows | | Sires | |
|---|---|---|---|---|---|---|
| | EBV | Rank | EBV | Rank | EBV | Rank |
| Milk[1], kg | 0.8204 | 0.7918 | 0.8322 | 0.8246 | 0.9198 | 0.9068 |
| Fat, kg | 0.9906 | 0.9889 | 0.9897 | 0.9890 | 0.9925 | 0.9917 |
| Protein, kg | 0.8066 | 0.7757 | 0.8127 | 0.8031 | 0.8740 | 0.8538 |
| Milk[2], kg | 0.9744 | 0.9675 | 0.9712 | 0.9674 | 0.9969 | 0.9960 |
| Fat,% | 0.9653 | 0.9607 | 0.9595 | 0.9549 | 0.9942 | 0.9935 |
| Protein,% | 0.9952 | 0.9946 | 0.9940 | 0.9931 | 0.9993 | 0.9991 |

[1]When milk yield was analyzed with fat and protein yields in the MT model (MT-Y); [2]When milk yield was analyzed with milk fat and protein contents in the MT model (MT-P).

# Paper II

**An attempt at predicting blood β-hydroxybutyrate from Fourier-transform mid-infrared spectra of milk using multivariate mixed models in Polish dairy cattle**

T.K. Belay, B.S. Dagnachew, Z.M. Kowalski and T. Ådnøy

*Journal of Dairy Science (In Press)*

J. Dairy Sci. 100:1–15
https://doi.org/10.3168/jds.2016-12252


# An attempt at predicting blood β-hydroxybutyrate from Fourier-transform mid-infrared spectra of milk using multivariate mixed models in Polish dairy cattle

**T. K. Belay,*[1] B. S. Dagnachew,* Z. M. Kowalski,† and T. Ådnøy***
*Department of Animal and Aquacultural Sciences, Norwegian University of Life Sciences, PO Box 5003, 1432 Ås, Norway
†Department of Animal Nutrition and Dietetics, University of Agriculture in Krakow, Krakow 30-059, Al. Mickiewicza 24/28, Poland

## ABSTRACT

Fourier transform mid-infrared (FT-MIR) spectra of milk are commonly used for phenotyping of traits of interest through links developed between the traits and milk FT-MIR spectra. Predicted traits are then used in genetic analysis for ultimate phenotypic prediction using a single-trait mixed model that account for cows' circumstances at a given test day. Here, this approach is referred to as indirect prediction (IP). Alternatively, FT-MIR spectral variable can be kept multivariate in the form of factor scores in REML and BLUP analyses. These BLUP predictions, including phenotype (predicted factor scores), were converted to single-trait through calibration outputs; this method is referred to as direct prediction (DP). The main aim of this study was to verify whether mixed modeling of milk spectra in the form of factors scores (DP) gives better prediction of blood β-hydroxybutyrate (BHB) than the univariate approach (IP). Models to predict blood BHB from milk spectra were also developed. Two data sets that contained milk FT-MIR spectra and other information on Polish dairy cattle were used in this study. Data set 1 (n = 826) also contained BHB measured in blood samples, whereas data set 2 (n = 158,028) did not contain measured blood values. Part of data set 1 was used to calibrate a prediction model (n = 496) and the remaining part of data set 1 (n = 330) was used to validate the calibration models, as well as to evaluate the DP and IP approaches. Dimensions of FT-MIR spectra in data set 2 were reduced either into 5 or 10 factor scores (DP) or into a single trait (IP) with calibration outputs. The REML estimates for these factor scores were found using WOMBAT. The BLUP values and predicted BHB for observations in the validation set were computed using the REML estimates. Blood BHB predicted from milk FT-MIR spectra by both approaches were regressed on reference blood BHB that had not been used in the model development. Coefficients of determination in cross-validation for untransformed blood BHB were from 0.21 to 0.32, whereas that for the log-transformed BHB were from 0.31 to 0.38. The corresponding estimates in validation were from 0.29 to 0.37 and 0.21 to 0.43, respectively, for untransformed and logarithmic BHB. Contrary to expectation, slightly better predictions of BHB were found when univariate variance structure was used (IP) than when multivariate covariance structures were used (DP). Conclusive remarks on the importance of keeping spectral data in multivariate form for prediction of phenotypes may be found in data sets where the trait of interest has strong relationships with spectral variables.
**Key words:** direct prediction, indirect prediction, β-hydroxybutyrate, milk spectra, dairy cattle

## INTRODUCTION

Subclinical ketosis (**SCK**) is an economically important metabolic disorder in early-lactation dairy cows. It is associated with reduced milk production (Duffield et al., 2009), reduced reproductive performance (Walsh et al., 2007), and increased risk of displaced abomasum (LeBlanc et al., 2005; Duffield et al., 2009) and clinical ketosis (Seifi et al., 2011). The disorder is closely related to the negative energy balance occurring in early lactation. Prevalence of SCK can vary between farms; reported prevalence rates range from 8.9 to 43% in the first 2 mo of lactation (McArt et al., 2012; van der Drift et al., 2012; Suthar et al., 2013).

Clinical and SCK are characterized by increased concentrations of ketone bodies (BHB, acetoacetate, and acetone) in milk and blood (Andersson, 1988). Blood BHB concentration has been used as a gold standard for detection of SCK and several studies have used a threshold of 1.2 (Duffield et al., 1997; van der Drift et al., 2012) to 1.4 mmol/L (Geishauser et al., 2000; Oetzel, 2004; Denis-Robichaud et al., 2014) to discriminate between cows with and without SCK. However,

Received November 3, 2016.
Accepted April 5, 2017.
[1]Corresponding author: tesfaye.kebede.belay@nmbu.no

the gold standard method does not allow for routine testing of all animals at risk at regular intervals due to some practical limitations, such as difficulty in blood sampling, especially for farmers or testing many blood samples at a time. Determination of ketone bodies in milk could make the sampling easier (Enjalbert et al., 2001; de Roos et al., 2007). Milk sampling is performed monthly in the milk recording procedures. More routinely, measurements of milk BHB can be done by Fourier transform mid-infrared (**FT-MIR**) spectroscopy analysis in milk samples at test days (de Roos et al., 2007; van der Drift et al., 2012). Blood or milk BHB predicted from milk spectra could be used for detection of SCK in farm management for dairy cows.

The FT-MIR spectra acquisition of the milk sample is multivariate (e.g., 1,060 variables per sample). Hundreds of these spectral variables are used for phenotyping of traits of interest (e.g., BHB) through links developed between the traits and milk spectra. The predicted phenotypes are then together with pedigree information and variance component estimates used in BLUP to calculate individual breeding values (EBV) and other random components included into the model; this is the conventional method used today for genetic evaluation of animals. Dagnachew et al. (2013b) referred to such an approach as indirect prediction (**IP**) and also proposed an alternative approach called direct prediction (**DP**), where genetic analyses are directly applied on the milk FT-MIR spectral variables or on factor scores (latent traits). The BLUP predictions (EBV, herd test day, permanent environment, and residual) for the traits of interest are predicted as correlated traits to the corresponding random components of spectra. Milk FT-MIR spectral variables exhibit strong correlations among each other (Soyeurt et al., 2010; Dagnachew et al., 2013a), and a direct genetic analyses on such correlated spectral variables may result in better accuracy of genetic evaluations (Dagnachew et al., 2013b). In our study, the 2 approaches, IP and DP, were used to predict phenotypes using BLUP predictions of the random and fixed effects part of the models.

The 2 approaches (IP and DP) have been used to predict EBV for milk fat, protein, and lactose contents using Norwegian dairy goat data (Dagnachew et al., 2013b). Those authors showed that accuracy of EBV were improved by 3 to 5% when DP was used compared with the IP approach; they also reported high rank correlation coefficients (0.93 to 0.96) between EBV predicted using IP and DP. However, independent chemical analyses (reference values) for the milk content were not available in that study (i.e., the study relied on phenotypes predicted based on the same spectra for both model calibration and evaluation). Recently, Bonfatti et al. (2017) compared the IP and DP approaches to estimate EBV for several traits related to fine composition and technological properties of milk and reported rank correlations ranging from 0.07 to 0.96, but $<0.5$ for most traits. In the present study, we adopted the IP and DP approaches to predict phenotype (not EBV) for BHB having an independent reference value for this trait. Our hypothesis was that keeping spectra multivariate in the form of factor scores or latent traits throughout REML and BLUP analyses instead of converting the spectra into single-trait before the genetic analyses should keep more information, and possibly also give a better prediction of the derived single-trait BHB after multiple-trait mixed modeling accounting for the cows' circumstances. We hypothesized that with multivariate information, one variable may carry information about another variable and thus improve the predictions.

The main objective of our study was to verify whether multivariate mixed modeling of milk FT-MIR spectra that are in the form factor scores (DP) gives better prediction of blood BHB than the univariate (IP) approach, where traits are first predicted from the spectra and then the predicted traits used in genetic analysis for ultimate phenotypic prediction. To do so, the current study developed prediction models for blood BHB from milk spectra and blood BHB measured by reference method.

## MATERIALS AND METHODS

### *Data*

Two data sets (referred to as data set 1 and data set 2) were used in our study, made available by the Polish Federation of Cattle Breeders and Dairy Farmers, which provides the monthly milk recording of cows in Poland. Both data sets contained FT-MIR spectra of individual milk samples, pedigree information, milk yield, and other cow and farm information. The milk samples had been analyzed by the MilkoScan FT6000 instrument (Foss Analytical A/S, Hillerød, Denmark). Major milk components, such as protein, fat, lactose, fat composition (both group and individual fatty acids), and ketone bodies (acetone and BHB), had been predicted using the Foss calibration and were available in the data sets.

***Data Set 1.*** After merging the measured blood BHB and phenotypes predicted from the spectra with their corresponding spectral data, data set 1 consisted of data on 832 Polish Holstein Friesian cows (1,914 observations; i.e., at least 2 records per cow) that had been milked 2 or 3 times per day. The spectra and other phenotypes that were predicted from the spectra were recorded for each milking, whereas blood BHB

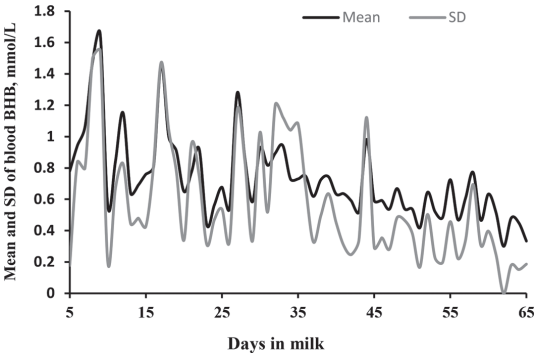



**Figure 1**. Mean and SD of blood BHB (mmol/L) by DIM from 826 lactating Polish dairy cows.

was measured only once using the glucometer Optium Xido (Abbott, Winey, UK) on test day at 1000 to 1400 h. For better correlation between blood BHB and milk spectra, milk and blood samples from the same milking were used. The data were collected between September 2013 and June 2014, and the cows were from 2 to 127 DIM. Cows with <5 (n = 1) or >65 DIM (n = 4) or with duplicated records (n = 1) were excluded from analysis, resulting in 826 cows kept in 55 herds. Mean blood BHB concentrations at each DIM were calculated for the cows (826) and is depicted in Figure 1.

Cows in data set 1 were randomly divided into a calibration and a validation set. The calibration set (n = 496 cows from 31 herds) was used to develop a link between milk spectra and blood BHB, whereas the validation set (n = 330 cows from 24 herds) was used to validate the prediction model and for evaluation of the IP and DP approaches. Descriptive statistics of the calibration and validation sets of data set 1 are presented in Table 1. Phenotypic associations of the measured blood BHB with Foss-predicted milk BHB and predicted blood BHB from spectra of validation set by models developed in our study were estimated.

***Data Set 2.*** The large data set (data set 2) originally contained 1,173,141 observations recorded from September to December 2014. Unlike data set 1, data set 2 did not contain BHB measured from blood. All cows with <5 (n = 67 observations) or >65 DIM (n = 934,600 observations) were excluded from analysis, resulting in 238,474 observations. Furthermore, cows with no pedigree information or with an unknown age at test day were removed from the data set. For better separation of animal effects from herd effects, herd test days (**HTD**) with less than 2 records were also excluded from the data set. Finally, 158,028 observations from 107,988 cows (daughters of 8,339 sires and 100,423 dams) kept in 9,102 herds remained for estimation of (co)variance components of the spectra. A pedigree file containing animals with records and their ancestors was available. The total number of animals in the pedigree file that had a link to the data file were 469,751.

### *Selection of Spectral Variables*

The major proportion of milk is water, hence the water spectrum influences the milk spectra. Both the O-H bending region (approximately between 1,620 and 1,700 $cm^{-1}$) and the O-H stretching region (above 3,025 $cm^{-1}$) of water are more or less opaque for the infrared light used, resulting in noise-like regions (Afseth et al., 2010; Dagnachew et al., 2013a). Therefore, the 2 regions comprising 536 spectral data points were excluded and only the remaining 524 spectral data points (926–1,617 $cm^{-1}$ and 1,701–3,025 $cm^{-1}$) were considered for further analysis. These 524 spectral variables are referred to as region I.

Spectral region between 1,803 and 2,800 $cm^{-1}$ (262 wave numbers) has been reported to have no specific bands or useful chemical information (Andersen et al., 2002; Iñón et al., 2004; Dagnachew et al., 2013a). Region I without these spectral variables (between 1,803–2,800 $cm^{-1}$) is referred to as region II (i.e., region II is a subset of region I).

### *Preprocessing of Spectra*

Calibrations of prediction models, for relationship of milk spectra and blood BHB, and genetic analyses were performed on both unprocessed and preprocessed spectra. The selected spectral variables were preprocessed by 2 methods. First, second derivatives of the spectra by the Savitzky-Golay (**SG**) numerical algorithm with

**Table 1**. Descriptive statistics for reference blood BHB (mmol/L) in data set 1 and its subsets: calibration, subset of calibration (extreme values ≤0.5 or ≥1.4 mmol/L), and validation sets

| Data | No. of cows | Mean | SD | Minimum | Maximum |
|---|---|---|---|---|---|
| Data set 1 | 826 | 0.760 | 0.743 | 0.1 | 6.3 |
| Calibration set | 496 | 0.734 | 0.725 | 0.1 | 6.3 |
| Calibration subset | 296 | 0.716 | 0.928 | 0.1 | 6.3 |
| Validation set | 330 | 0.800 | 0.768 | 0.1 | 5.5 |

9 window-size and second-order polynomials were calculated. Second, the spectra preprocessed by the SG were further preprocessed using extended multiplicative signal correction (**EMSC**). Preprocessing was performed on both region I and region II.

### Multivariate Calibration of Prediction Models

The calibration data set (n = 496) was used to develop a link between blood BHB and milk spectra using the pls package (Mevik and Wehrens, 2007) implemented in R (R Core Team, 2016). Partial least squares (**PLS**) regression analyses were done on all 496 blood BHB values in the calibration set, and on a subset with 296 observations with extreme blood BHB values (low: <0.6 mmol/L, high: ≥1.4 mmol/L). In the analyses, blood BHB was used as a response variable (y) whereas unprocessed or preprocessed spectra (region I or II) were used as predictor variables (X).

The calibration models were cross-validated using 10 random segments, and the optimum number of PLS factors were determined based on the first local minimum value in root mean squared error of prediction of the cross-validation ($\mathbf{RMSE_{cv}}$). The calibration models were then applied to the validation data set. The PLS regression parameters, such as regression coefficients $\left(\hat{\beta}_{\mathrm{PLS}}\right)$, matrices of weights (**W**) that reflect covariance structures between y and X, matrices of factor scores (**T**), matrices of y-loadings (**Q**), and matrices of X-loadings (**P**), were used in the subsequent predictions and calculations. Predictions were performed following the DP or IP approaches. Figure 2 shows a schematic representation of these 2 prediction approaches.

### DP

The DP approach has several steps: spectral variables dimension reduction into few factors, estimation of covariance components for the factor scores from data set 2, prediction of random components for factor scores from the validation set using the estimated covariance components, and conversion of predicted factor scores into predicted blood BHB. The steps are described in detail in subsequent sections.

***Spectral Variables Dimension Reduction.*** Direct genetic analysis to estimate (co)variance components for the random effects of the mixed model of all the selected spectral variables (i.e., 524 or 262 spectral data points) simultaneously was not possible with currently available analytical packages used in quantitative genetics. They are limited to fewer traits in multitrait analysis (Meyer, 2007; Madsen and Jensen, 2008; Gilmour et al., 2009). Moreover, many of the spectral variables are highly collinear and the redundancy needs to be removed or absorbed. Dimension reduction is usually done by principal component analysis (**PCA**), PLS regression, or factor analysis. In the current study, we reduced the dimension of the spectral variables into factor scores through a weight matrix (**W**) obtained from the PLS regression with respect to blood BHB, as described above. The PLS factors mainly contain information related to the response variable(s) in the regression, and hence are expected to give better prediction than PCA components that contain general info in spectra. Previously, PCA was used on the same data set and results from such an analysis has been reported (Belay et al., 2015).

***Estimation of Covariance Components for Factor Scores.*** A matrix of factor scores (**T**) were calculated for observations of spectra in data set 2 using the weight matrix (**W**) that had been obtained by PLS regression on the calibration part of data set 1:

$$\mathbf{T} = \mathbf{XW}, \qquad [1.1]$$

where **T** is an $n \times c$ matrix of factor scores, with $n$ being number of observations (n = 158,028) and $c$ the number of factors. Numbers of factors were chosen by cross-validation in PLS regression for the calibration part of data set 1. **X** is an $n \times k$ spectral data matrix for data set 2, with $k$ being number of spectral variables. **W** is a $k \times c$ weight matrix that reflects the covariance structure between milk spectra (**X**) and blood BHB (**y**); **W** is determined as a function of **X** and **y** by PLS regression in the calibration part of data set 1.

The factor scores characterize the relationship of the milk information to the blood BHB and give the relationship among the spectral variables. Factor scores were then treated as traits in multiple-trait mixed model analyses. A repeatability test day animal model was used to estimate variance components for the factor scores, **T**, from data set 2 (only spectra, no blood BHB). The model in matrix notation was

$$\mathbf{t} = \mathbf{Xb} + \mathbf{Za} + \mathbf{Wp} + \mathbf{Hd} + \mathbf{e}, \qquad [1.2]$$

where **t** is a vector of factor scores (the **t** of 1 milk sample above the other); **b** is a vector of fixed effects of breed (2 levels), lactation number (1 to 4), herd size (3 levels) × lactation stage (6 levels), and months of test (4 levels); **a** is a vector of random animal genetic effects; **p** is a vector of random permanent environmental effects; **d** is a vector of random HTD effects; and **e** is vector of random residual effects. **X**, **Z**, **W**, and **H** are design matrices that relate records to the corresponding (fixed and random) effects. The 2 breed levels are Polish Holstein-Friesian (black-white), which accounted

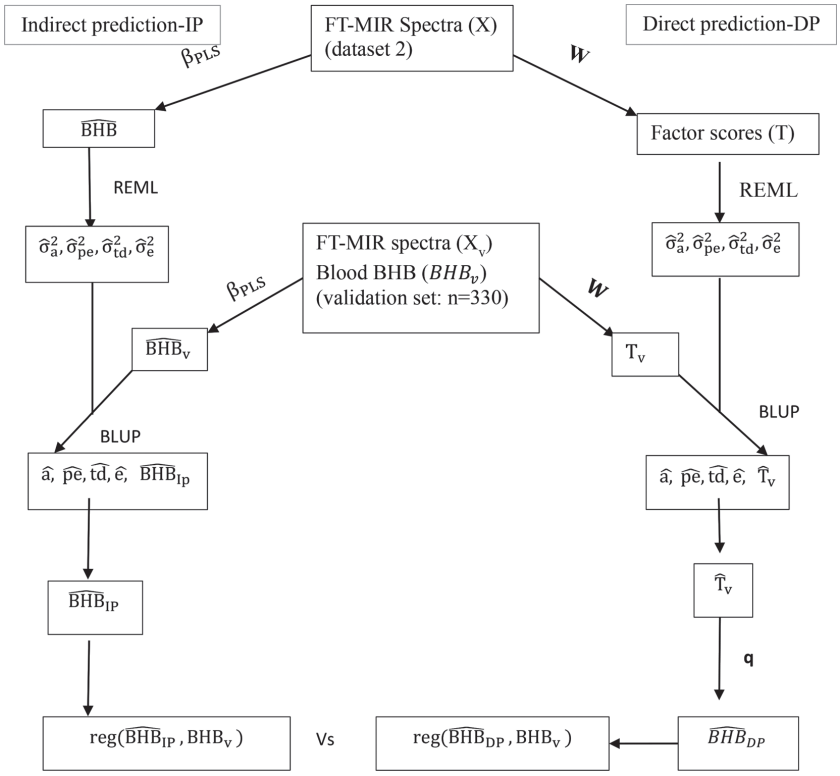



**Figure 2**. Schematic representation of the indirect (IP) and direct (DP) prediction approaches. In the IP, the phenotype for trait of interest (e.g., BHB) was predicted from milk spectra using regression coefficient ($\beta_{PLS}$), and this predicted trait was analyzed with a genetic model for ultimate phenotypic prediction. In the DP, multiple predicted factor scores that were obtained through a weight matrix (W) were analyzed with a genetic multivariate model before predicted model components were combined through the y-loading (q) to eventually predict the phenotype. Reg = regression; $T_v$ = factor scores for observations in the validation set; $\sigma^2_a, \sigma^2_{pe}, \sigma^2_{td}$, and $\sigma^2_e$ = estimates of variance components for genetic, permanent animal environment, herd test-day and residual, respectively; a, pe, td, and e = additive genetic, permanent animal environmental, herd test-day, and residual effects, respectively; $\widehat{BHB}$ = predicted BHB from spectra of data set 2 using the PLS regression-based prediction equation; $\widehat{BHB}_v$ = predicted BHB from spectra of validation set using the PLS regression-based prediction equation; $\widehat{BHB}_{IP}$ = predicted BHB using the IP approach; $\widehat{BHB}_{DP}$ = predicted BHB using the DP approach; $BHB_v$ = measured/reference blood in the validation set.

for 86.4% of the records, and the other 15 breeds (such as Polish Holstein-Friesian red-white, Simental, Polish red-white, and so on), which accounted for 13.6% of the records all together. Herd size was defined based on number of cows with records per herd in the original data set (before edition) and categorized as small (<35 cows), medium (35–99 cows), and large (≥100 cows). Each group contained similar numbers of cows. Days in milk (lactation stage) were categorized into 6 levels, each with 10 test days. Each of the 4 test months were modeled. The assumed (co)variance structure was

$$\text{var}\begin{bmatrix}\mathbf{a}\\ \mathbf{p}\\ \mathbf{d}\\ \mathbf{e}\end{bmatrix} = \begin{bmatrix}\mathbf{G}\otimes\mathbf{A} & 0 & 0 & 0\\ 0 & \mathbf{P}\otimes\mathbf{I} & 0 & 0\\ 0 & 0 & \mathbf{H}\otimes\mathbf{I} & 0\\ 0 & 0 & 0 & \mathbf{R}\otimes\mathbf{I}\end{bmatrix},$$

where $\mathbf{G}$ is genetic covariance (5 × 5) matrix, $\mathbf{P}$ is the covariance (5 × 5) matrix for within-cow permanent environmental effects, $\mathbf{H}$ is the covariance (5 × 5) matrix for HTD effects, and $\mathbf{R}$ is the residual covariance (5 × 5) matrix, $\mathbf{I}$ and $\mathbf{A}$ are identity and additive rela-

tionship matrices, respectively, and ⊗ is the Kronecker product.

The (co)variance component estimates were obtained by the REML method using the multivariate average information-REML algorithm of WOMBAT (Meyer, 2007). These estimated variance components are population parameters that should characterize any data coming from the population. Preliminary bivariate analyses were performed and (co)variance component estimates from the bivariate analyses of the factor scores were pooled using the iterative summing of expanded part matrices approach (Mäntysaari, 1999) implemented in WOMBAT (Meyer, 2007). The pooled covariance matrices were priors in the multivariate REML.

***BLUP Analyses for Factor Scores from Spectra of Validation Set.*** Once (co)variance components were estimated for factor scores from the large spectral data set (data set 2), BLUP values could be calculated for the random components of any new data set from the population using the estimated (co)variance components and the structural circumstances of the new data set (genetic relationship, permanent environment, and HTD design). Factor scores ($\mathbf{T}_v$) for observations in the validation set (n = 330) were calculated using the weight matrix $\mathbf{W}$ from the model calibration and milk spectra of the validation set as follows. Neither blood BHB nor milk spectra in the validation set were used in the model development:

$$\mathbf{T}_v = \mathbf{X}_v\mathbf{W}, \qquad [1.3]$$

where $\mathbf{T}_v$ is an $n_v \times c$ matrix of factor scores, $\mathbf{W}$ is as defined in Eq. [1.1], and $\mathbf{X}_v$ is an $n_v \times k$ spectral data matrix for the validation set, with $k$ being number of spectral variables. The subscript v is used to indicate validation set.

A model similar to Eq. [1.2], with some modification in the fixed part of the model, was used to run BLUP for the $\mathbf{T}_v$ using the covariance components estimated with Eq. [1.2] and the $\mathbf{I}$ and $\mathbf{A}$ relevant for the validation set. In this model [1.4], fixed effects of lactation number (4 levels), lactation stage (6 levels), and year (2 levels) × season (2 levels: April to September and October to March) were fitted:

$$\mathbf{t}_v = \mathbf{Xb} + \mathbf{Za} + \mathbf{Wp} + \mathbf{Hd} + \mathbf{e}, \qquad [1.4]$$

where $\mathbf{t}_v$ is a vector of factor scores for observations in the validation set (with the $\mathbf{t}_v$ of 1 milk sample above the other), and other model elements were as defined in the Eq. [1.2].

***Conversion of the Predicted Factor Scores into Predicted Blood BHB.*** In addition to predictions of the random effects $(\hat{\mathbf{a}}, \hat{\mathbf{p}}, \text{and } \hat{\mathbf{d}})$, predicted factor scores $(\hat{\mathbf{T}}_v)$ were given directly by WOMBAT from the BLUP run for the factor scores of the validation set. These predicted factor scores in multivariate form were converted into predicted blood BHB $(\widehat{\text{BHB}}_{\text{DP}})$ through the $\mathbf{Y}$-loading matrix ($\mathbf{Q}$) used in transposed vector form as

$$\widehat{\text{BHB}}_{\text{DP}} = \hat{\mathbf{T}}_v\mathbf{q}',$$

where $\mathbf{q}$ is a vector, not a matrix, as only a single response variable was in the PLS regression analysis. It had dimension $1 \times c$, where $c$ is the number of factors retained and it relates factors to response variables.

### *IP*

In this approach, BHB values were predicted from milk spectra using the PLS regression coefficient $(\hat{\beta}_{\text{PLS}})$ estimated above, and the predicted phenotypes used in further analyses. This is the conventional approach used for genetic evaluation and other purposes in animal sciences or in other fields. The BHB was predicted as

$$\widehat{\text{BHB}} = \mathbf{X}\hat{\beta}_{\text{PLS}}, \qquad [2.1]$$

where $\widehat{\text{BHB}}$ is predicted BHB phenotype from spectra $\mathbf{X}$ of data set 2 through PLS regression coefficient $(\hat{\beta}_{\text{PLS}})$ found in the calibration part of data set 1.

Covariance components and the corresponding variance ratios were estimated by REML for the predicted BHB by fitting single-trait animal model considering the same effects as in Eq. [1.2]:

$$\widehat{\text{BHB}} = \mathbf{Xb} + \mathbf{Za} + \mathbf{Wp} + \mathbf{Hd} + \mathbf{e}. \qquad [2.2]$$

The model elements are as defined in Eq. [1.2], but with univariate variance structure. We assumed $\text{var}(\mathbf{a}) = \mathbf{A}\sigma_a^2$, $\text{var}(\mathbf{p}) = \mathbf{I}\sigma_{pe}^2$, $\text{var}(\mathbf{d}) = \mathbf{I}\sigma_d^2$, and $\text{var}(\mathbf{e}) = \mathbf{I}\sigma_e^2$, where $\sigma_a^2$ is additive genetic variance, $\sigma_{pe}^2$ is permanent environmental variance, $\sigma_d^2$ is HTD variance, and $\sigma_e^2$ is residual variance. The BHB were then predicted for observations in the validation set $(\widehat{\text{BHB}}_v)$ using the $\hat{\beta}_{\text{PLS}}$ that was used in Eq. [2.1], but using spectra from the validation set ($\mathbf{X}_v$; i.e., $\widehat{\text{BHB}}_v = \mathbf{X}_v\hat{\beta}_{\text{PLS}}$).

Assuming similar effects as in Eq. [1.4], but with a single-trait animal model, BLUP solutions for fixed and random effects were found for $\widehat{\text{BHB}}_{\text{v}}$ from validation set:

$$\widehat{\text{BHB}}_{\text{v}} = \mathbf{Xb} + \mathbf{Za} + \mathbf{Wp} + \mathbf{Hd} + \mathbf{e}, \qquad [2.3]$$

where model elements were as defined in Eq. [1.2] and Eq. [2.2]. A similar variance structure as in Eq. [2.2] was assumed.

For this BLUP, run on $\widehat{\text{BHB}}_{\text{v}}$, the variance components used were estimated either from (1) single-trait REML (i.e., the one estimated in Eq. [2.2]), or (2) multiple-trait REML, as estimated in Eq. [1.2] after converting from multivariate covariance to univariate variance structures. The multivariate covariance structure from Eq. [1.2] for additive, permanent environmental, HTD, and residual covariance were converted into the corresponding univariate variance structure as

$$\hat{\sigma}^2_a = \mathbf{qGq}',$$

$$\hat{\sigma}^2_{pe} = \mathbf{qPq}',$$

$$\hat{\sigma}^2_d = \mathbf{qHq}', \text{ and}$$

$$\hat{\sigma}^2_e = \mathbf{qRq}. \qquad [2.4]$$

Predicted blood BHB $\left(\widehat{\text{BHB}}_{\text{IP}}\right)$ were directly obtained from WOMBAT together with predicted random effects and solutions for random residuals. Thus, we got 2 vectors of predicted BHB for observations in validation set, 1 from DP $\left(\widehat{\text{BHB}}_{\text{DP}}\right)$ and the other from IP $\left(\widehat{\text{BHB}}_{\text{IP}}\right)$. In addition to these predicted BHB, we measured blood BHB (reference values), which had not been used in calibration, from observations in the validation set.

### *Evaluation of the IP and DP Approaches*

The 2 sets of predicted blood BHB $\left(\widehat{\text{BHB}}_{\text{DP}} \text{ and } \widehat{\text{BHB}}_{\text{IP}}\right)$ are 2 different predictions of blood BHB. Performance of the 2 approaches for prediction of BHB was evaluated based on adjusted coefficient of determination ($\mathbf{R^2}$) estimated by regressing the IP or DP predicted blood BHB against measured blood BHB (reference values). Prediction with the IP or DP approach was also compared with prediction of BHB by PLS (using the PLS regression found in calibration on the milk spectra of the validation set).

## RESULTS

### *Description of Reference Blood BHB*

Table 1 shows descriptive statistics for reference blood BHB data. Content of BHB in the 826 blood samples analyzed ranged from 0.1 to 6.3 mmol/L, with an average of 0.760 mmol/L and a standard deviation of 0.743 mmol/L. More than 80% of the samples had <1.0 mmol/L. The most frequent concentrations observed were 0.3, 0.4, and 0.6 mmol/L. Out of the 826 cows sampled, 114 of them had a concentration ≥1.2 mmol/L of blood BHB. Mean blood BHB concentration for all cows (826) at each DIM was calculated and is given in Figure 1. The mean blood BHB concentration was generally high around the beginning of lactation and decreased as DIM progressed. Mean and standard deviation of BHB in calibration set were lower than the corresponding values in the validation set (Table 1). We also found a difference in ranges of BHB values between calibration and validation set.

### *Cross-Validation and Validation Results*

The link between untransformed or log-transformed blood BHB and milk spectra was developed using PLS regression analysis on the calibration set (all, n = 496) and on its subset (n = 296 with extreme values). The results from such analyses are presented in Tables 2 and 3. Based on the 3 sets of spectra (unprocessed and preprocessed by SG and EMSC), the 2 spectral regions (regions I or II), and the 2 categories of blood BHB values (all or extreme), 9 different prediction models were developed. The idea was to find models with the better fit. Five to 10 PLS factors were retained based on the first local minimum value in $\text{RMSE}_{\text{cv}}$. Table 2 shows cross-validation and validation statistics for untransformed blood BHB predicted using the 9 prediction models developed. In the cross-validations, averages of predicted values were the same as corresponding mean reference values, but with smaller standard deviations and ranges than the reference values. These results (i.e., the low variation and the reduced range of predicted values) indicate lack of precision of the models on high values. For untransformed blood BHB, the $R^2$ in cross-validation ($\mathbf{R^2_{cv}}$) ranged from 0.217 to 0.316, with $\text{RMSE}_{\text{cv}}$ ranging from 0.630 to 0.787 (Table 3). The $\text{RMSE}_{\text{cv}}$ were relatively high, which might be due to the lack of precision of the models on high values of the data sets (Table 2) that had a high proportion of low values. This is evident from models developed based on extreme BHB values, where a majority of them were low with few high values. The logarithmic transformation of blood BHB values increased the $R^2_{\text{cv}}$,

**Table 2**. Cross-validation and validation descriptive statistics[1] for untransformed blood BHB (mmol/L) predicted from milk spectra using different calibration models

| | Cross-validation | | | | Validation | | | |
|---|---|---|---|---|---|---|---|---|
| Model[2] | Mean | SD | Minimum | Maximum | Mean | SD | Minimum | Maximum |
| Unprocessed spectra | | | | | | | | |
| All BHB values with region I | 0.734 | 0.389 | 0.014 | 2.721 | 0.756 | 0.443 | −0.059 | 2.612 |
| Extreme BHB with region II | 0.716 | 0.565 | −0.156 | 3.304 | 0.756 | 0.598 | −0.389 | 3.082 |
| All BHB values with region II | 0.734 | 0.390 | −0.017 | 2.506 | 0.745 | 0.444 | −0.085 | 2.629 |
| 2nd derivative spectra (SG) | | | | | | | | |
| All BHB values with region I | 0.734 | 0.397 | −0.119 | 2.551 | 0.746 | 0.425 | −0.056 | 2.535 |
| Extreme BHB with region II | 0.716 | 0.608 | −0.436 | 2.905 | 0.774 | 0.623 | −0.426 | 3.327 |
| All BHB values with region II | 0.734 | 0.403 | −0.118 | 2.257 | 0.733 | 0.429 | −0.075 | 2.509 |
| EMSC preprocessed spectra | | | | | | | | |
| All BHB values with region I | 0.734 | 0.389 | −0.118 | 2.114 | 0.732 | 0.403 | −0.175 | 2.119 |
| Extreme BHB with region II | 0.716 | 0.586 | −0.845 | 2.900 | 0.730 | 0.593 | −0.730 | 2.722 |
| All BHB values with region II | 0.734 | 0.392 | −0.260 | 2.159 | 0.724 | 0.405 | −0.206 | 2.084 |

[1]Mean, SD, minimum, and maximum of predicted blood BHB values (mmol/L) are presented.

[2]Spectra were preprocessed by Savitzky-Golay (SG) algorithm and extended multiplicative signal correction (EMSC).

ranging from 0.313 to 0.381. The $RMSE_{cv}$ for the log-transformed blood BHB values were between 0.222 and 0.278.

In the validation, predicted BHB contents (Table 2) were smaller than the corresponding reference values (Table 1). For untransformed blood BHB, $R^2$ in validation ($\mathbf{R^2_v}$) ranged from 0.308 to 0.374, with root mean square error of validation ($\mathbf{RMSE_v}$) ranging from 0.607 to 0.638 (Table 3). Similar to the cross-validation, log-transformation of BHB values increased $R^2$ in the validation, except for the extreme blood BHB values. For both untransformed and log-transformed blood BHB, $R^2_v$ were generally higher than the corresponding estimates in cross-validation, except models developed on log-transformed extreme blood BHB values (Table 3). The $RMSE_v$ for untransformed BHB were also lower than the corresponding $RMSE_{cv}$, whereas the reverse was true for the log-transformed BHB. This indicates that prediction ability of models based on log-transformed BHB could be compromised compared with the untransformed BHB.

In both cross-validation and validation, preprocessing of spectra either by SG or EMSC generally increased $R^2$ or reduced prediction errors, except for some models with untransformed BHB in the validation analyses. In validation, better results (high $R^2_v$ or low $RMSE_v$) were obtained with unprocessed spectra for untransformed BHB and with EMSC for log-transformed BHB. In

**Table 3**. Summary of partial least squares (PLS) regression prediction models for untransformed and $log_{10}$-transformed blood BHB in cross-validation and validation under unprocessed, second derivative (SG), and EMSC preprocessed[1] spectra

| | | Cross-validation[3] | | | | Validation[4] | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | Untransformed BHB | | Transformed BHB | | Untransformed BHB | | Transformed BHB | |
| Model | No. of factors[2] | $RMSE_{cv}$ | $R^2_{cv}$ | $RMSE_{cv}$ | $R^2_{cv}$ | $RMSE_v$ | $R^2_v$ | $RMSE_v$ | $R^2_v$ |
| Unprocessed spectra | | | | | | | | | |
| All BHB values with region I | 6 (8) | 0.6396 | 0.2109 | 0.2318 | 0.3198 | 0.6065 | 0.3738 | 0.2469 | 0.3964 |
| Extreme BHB with region II | 5 | 0.7865 | 0.2760 | 0.2776 | 0.3130 | 0.6327 | 0.3186 | 0.2792 | 0.2277 |
| All BHB values with region II | 5 (7) | 0.6397 | 0.2172 | 0.2326 | 0.3169 | 0.6153 | 0.3554 | 0.2468 | 0.3966 |
| 2nd derivative spectra (SG) | | | | | | | | | |
| All BHB values with region I | 5 | 0.6383 | 0.2201 | 0.2223 | 0.3730 | 0.6153 | 0.3554 | 0.2462 | 0.3999 |
| Extreme BHB with region II | 5 | 0.7748 | 0.2875 | 0.2628 | 0.3814 | 0.6199 | 0.3457 | 0.2765 | 0.2421 |
| All BHB values with region II | 5 (10) | 0.6384 | 0.2186 | 0.2210 | 0.3807 | 0.6150 | 0.3562 | 0.2446 | 0.4075 |
| EMSC preprocessed spectra | | | | | | | | | |
| All BHB values with region I | 5 (10) | 0.6302 | 0.2368 | 0.2227 | 0.3701 | 0.6312 | 0.3217 | 0.2414 | 0.4228 |
| Extreme BHB with region II | 5 | 0.7622 | 0.3159 | 0.2662 | 0.3718 | 0.6378 | 0.3075 | 0.2707 | 0.2741 |
| All BHB values with region II | 5 (10) | 0.6301 | 0.2351 | 0.2228 | 0.3690 | 0.6313 | 0.3216 | 0.2397 | 0.4309 |

[1]Spectra were preprocessed by Savitzky-Golay (SG) algorithm and extended multiplicative signal correction (EMSC).

[2]Number of PLS factors, and numbers of factors in parentheses were for models based on log-transformed blood BHB.

[3]$RMSE_{cv}$ = root mean squared error of the cross-validation; $R^2_{cv}$ = coefficient of determination of the cross-validation.

[4]$RMSE_v$ = root mean squared error of the validation; $R^2_v$ = coefficient of determination of the validation.

**Table 4**. Estimates of variance ratios for genetic, permanent environment (PE), herd test days (HTD), and residual random effects for the factor scores calculated from raw spectra in region II and all (or extreme[1]) blood BHB values

| Factor score | Variance ratio[2] | | | |
|---|---|---|---|---|
| | Genetic | PE | HTD | Residual |
| 1 | 0.093 (0.093) | 0.143 (0.143) | 0.169 (0.169) | 0.595 (0.595) |
| 2 | 0.221 (0.227) | 0.212 (0.213) | 0.082 (0.080) | 0.485 (0.480) |
| 3 | 0.176 (0.180) | 0.119 (0.122) | 0.158 (0.166) | 0.547 (0.531) |
| 4 | 0.163 (0.156) | 0.165 (0.162) | 0.137 (0.130) | 0.534 (0.552) |
| 5 | 0.053 (0.056) | 0.070 (0.075) | 0.504 (0.480) | 0.374 (0.388) |

[1]Numbers in parentheses are variance ratio for factor scores calculated from raw spectra in region II and extreme blood BHB values.

[2]Ratio is relative to total phenotypic variance for each factor score. Standard error of variance ratios due to genetic, PE, HTD, and residual were 0.004–0.012, 0.004–0.01, 0.003–0.004, and 0.004, respectively.

the cross-validation, better results were obtained with EMSC for untransformed BHB and with EMSC or SG for log-transformed BHB. Despite the large number of spectral variables contained in region I, it had no effect on the $R^2$ of validation or cross-validation, except for validation of unprocessed spectra with untransformed BHB values (Table 3). Comparing the models with respect to the 2 sets of BHB values, prediction models with extreme BHB values (both untransformed and log-transformed) had generally higher $R^2_{cv}$ and $RMSE_{cv}$, but had lower $R^2_v$ than models with all BHB values (Table 3).

### *Genetic Parameters for the Factor Scores and BHB*

Out of the 9 prediction models that were developed based on untransformed blood BHB, 4 of them were selected to be used in the genetic analysis for ultimate phenotypic prediction: 2 models from raw and 2 from SG preprocessed spectra of region II with all or extreme BHB values. Models based on spectra of region I were not selected, as they did not give better accuracy despite the large number of spectral variables in region I. Models developed based on log-transformed BHB were not used for genetic analyses, as IP and DP approaches can be evaluated independent of the BHB scale. Moreover, models based on log-transformed BHB had slightly higher prediction error in the validation than in the cross-validation, whereas the reverse is true for models from untransformed BHB (Table 3). Estimates of variance ratios for each factor score, calculated from raw and preprocessed spectra using the 4 selected calibration models, are presented in Tables 4 and 5. Genetic variance ratios (heritabilities) for the 5 factor scores calculated from unprocessed spectra ranged from 0.053 to 0.227 (Table 4) and from 0.081 to 0.158 (Table 5) for SG preprocessed spectra. The corresponding variance ratios of the permanent environmental effects ranged from 0.070 to 0.213 and from 0.074 to 0.153. Variance ratios of the HTD ranged from 0.080 to 0.504 and from 0.130 to 0.374 for the factors from raw and preprocessed spectra, respectively. The corresponding variance ratios for the residual effects varied from 0.374 to 0.595 and from 0.437 to 0.595.

Corresponding variance components for blood BHB were calculated from the estimated covariance structures of factor scores using Eq. [2.4]. Table 6 presents estimated variance ratios and variance components for

**Table 5**. Estimates of variance ratios for genetic, permanent environment (PE), herd test day (HTD), and residual random effects for the factor scores calculated from Savitzky-Golay (SG) preprocessed spectra in region II and all (or extreme[1]) blood BHB values

| Factor score | Variance ratio[2] | | | |
|---|---|---|---|---|
| | Genetic | PE | HTD | Residual |
| 1 | 0.097 (0.095) | 0.140 (0.140) | 0.169 (0.169) | 0.595 (0.595) |
| 2 | 0.081 (0.084) | 0.118 (0.111) | 0.257 (0.259) | 0.544 (0.545) |
| 3 | 0.158 (0.144) | 0.114 (0.106) | 0.209 (0.231) | 0.519 (0.519) |
| 4 | 0.102 (0.096) | 0.086 (0.074) | 0.376 (0.349) | 0.437 (0.481) |
| 5 | 0.113 (0.118) | 0.118 (0.153) | 0.173 (0.130) | 0.595 (0.599) |

[1]Numbers in parentheses are variance ratio for factor scores calculated from Savitzky-Golay (SG) preprocessed spectra in region II and extreme blood BHB values.

[2]Ratio is relative to total phenotypic variance for each factor. Standard error of variance ratios due to genetic, PE, HTD, and residual is 0.006–0.008, 0.000–0.007, 0.003–0.004, and 0.004, respectively.

**Table 6**. Multifactor (direct prediction) REML based estimates of variance ratios and variance components for genetic, permanent environment (PE), herd test day (HTD), and residual random effects for BHB found from raw or Savitzky-Golay (SG) preprocessed milk spectra from data set 2 (n = 158,028) using the 4 selected calibration models[1]

| | Variance ratio (variance component) | | | |
|---|---|---|---|---|
| Model | Genetic | PE | HTD | Residual |
| Unprocessed spectra | | | | |
| All BHB values with region II | 0.111 (0.018) | 0.142 (0.023) | 0.279 (0.045) | 0.468 (0.076) |
| Extreme BHB with region II | 0.109 (0.036) | 0.144 (0.047) | 0.275 (0.090) | 0.473 (0.156) |
| 2nd derivative spectra (SG) | | | | |
| All BHB values with region II | 0.083 (0.017) | 0.158 (0.032) | 0.342 (0.070) | 0.416 (0.085) |
| Extreme BHB with region II | 0.088 (0.036) | 0.145 (0.060) | 0.337 (0.139) | 0.430 (0.177) |

[1]Estimated multivariate covariances have been converted into one-trait variance structure relevant for BHB prediction. Ratios are relative to total phenotypic variance for BHB from each model. Values in parentheses are estimates for variance components.

genetic, PE, HTD, and residual of BHB. For BHB from unprocessed spectra, average estimates of variance ratios (variance ratios for BHB from all and extreme BHB values were averaged within spectral sets) for genetic, PE, HTD, and residual were 0.110, 0.143, 0.277, and 0.471, respectively. The corresponding values for BHB from SG preprocessed spectra were 0.086, 0.152, 0.340, and 0.423. Variance components for BHB were also estimated by univariate REML using Eq. [2.2], where spectral variables had first been converted into a single trait (BHB) through the PLS regression coefficient. Variance components and variance ratios for such BHB were slightly lower than the genetic parameters presented in Table 6, except the estimates for HTD, and estimated residual effects that were the same (Tables 6 and 7).

Most of the factor scores and BHB that were predicted from unprocessed spectra had higher estimates of heritability and proportion of variance due to PE and HTD effects than those from SG preprocessed spectra (Tables 4–7). The larger estimates for factors and BHB from unprocessed spectra may be due to unprocessed spectra possibly containing unwanted heritable variation, which could be removed by preprocessing. Spectral preprocessing removes not only unwanted variations (such as variation in intensity of light sources, scattering, contaminants, optical path length, and so on) in spectra, but also some real molecular structures or milk constituents, which might be heritable.

### *Prediction Ability of the IP and DP Approaches*

Performance of the IP and DP approaches were evaluated based on $R^2$ estimated by regressing the IP- or DP-predicted blood BHB on the reference blood BHB values of the validation data set that had not been used in model calibrations. Table 8 presents the estimated $R^2$ for the IP and DP approaches. The $R^2$ for the DP method were intermediate and ranged from 0.263 to 0.298, whereas the corresponding estimates for IP method, when variance components from multiple REML were used, ranged from 0.281 to 0.301 and from 0.278 to 0.306 when variance components from single-trait REML were used (Table 8). The predictability of the IP approach was slightly higher compared with the predictability of the DP approach; this means that

**Table 7**. Univariate (indirect prediction) REML based estimates of variance ratios and variance components for genetic, permanent environment (PE), herd test day (HTD), and residual random effects for BHB found from raw or Savitzky-Golay (SG) preprocessed milk spectra from data set 2 (n = 158,028) using the 4 selected calibration models[1]

| | Variance ratio (variance component) | | | |
|---|---|---|---|---|
| Model | Genetic | PE | HTD | Residual |
| Unprocessed spectra | | | | |
| All BHB values with region II | 0.103 (0.017) | 0.141 (0.023) | 0.288 (0.047) | 0.468 (0.076) |
| Extreme BHB with region II | 0.101 (0.033) | 0.142 (0.047) | 0.283 (0.094) | 0.473 (0.156) |
| 2nd derivative spectra (SG) | | | | |
| All BHB values with region II | 0.075 (0.015) | 0.155 (0.032) | 0.353 (0.072) | 0.416 (0.085) |
| Extreme BHB with region II | 0.081 (0.033) | 0.142 (0.059) | 0.347 (0.143) | 0.430 (0.178) |

[1]Ratio is relative to total phenotypic variance for BHB from each model. Values in parentheses are estimates for variance components.

**Table 8**. Coefficient of determination between reference blood BHB values and blood BHB predicted by the direct and indirect prediction approaches from milk spectra[1]

| Calibration model | Indirect prediction (IP) | | Direct prediction (DP) |
|---|---|---|---|
| | Variances from single REML | Variances from multiple REML | |
| Unprocessed spectra | | | |
| All BHB values with region II | 0.2865 | 0.2898 | 0.2692 |
| Extreme BHB with region II | 0.2775 | 0.2805 | 0.2631 |
| 2nd derivative spectra (SG) | | | |
| All BHB values with region II | 0.2943 | 0.2972 | 0.2804 |
| Extreme BHB with region II | 0.3061 | 0.3091 | 0.2978 |

[1]In the IP approach, where spectral variables first converted into single-trait and then genetic analysis was applied on the trait for ultimate phenotypic prediction, variances estimated from single-trait REML or multiple REML (after converting into variance structure) were used. In the DP approach, spectral variables reduced to factor scores that were analyzed using multitrait genetic analysis and eventually combined into the phenotypic trait.

a more accurate prediction of BHB was found when univariate variance structure was used than when multivariate covariance structures were used. Predictability of the 2 approaches were compared with the predictability of models given in Table 3 (PLS regression based predictions equations) for untransformed BHB in the validation analyses. The PLS regression-based prediction equations are the commonly used methods for phenotyping of trait of interest from milk FT-MIR spectra. Predictability of the IP and DP approaches were lower than predictability of equations developed based on the classical PLS regression in validation (Table 3).

As in the calibration models, preprocessing of spectra slightly improved accuracy of BHB prediction in both the IP and DP approaches. The improvement in accuracy due to preprocessing was slightly better in DP than in IP approaches. This indicated that the DP approach could perform better with spectra that contain less noisy information; noisy information in multivariate form could result in inferior performance.

## DISCUSSION

### *Multivariate Calibration Models*

The distribution of the data in the calibration set was slightly different than in the validation set, mainly due to lower mean and standard deviation of the reference values (Table 1). This could explain some difference in cross-validation and validation statistics. That the $R^2_v$ was generally higher than the $R^2_{cv}$ might be due to higher mean and standard deviation of the reference values in validation. It has been shown that $R^2$ is highly dependent on distribution of the data and especially on the range of data (Grelet et al., 2016). Because of the way in which blood BHB was measured (i.e., values with few digits: 0.1, 0.2, …, 6.3), many samples had the same BHB values; this resulted in a large number of few distinct values. Such duplication in BHB values (not in the corresponding spectra) could also influence the $R^2_{cv}$ by reducing variation or range of values in the random segments used for the cross-validation. In the validation set, data were not divided into random segments, so existing variation in the blood BHB was available. That could possibly result in the higher observed $R^2_v$ than $R^2_{cv}$.

The $R^2_{cv}$ of the prediction models developed in our study were low, but in the range of estimates reported for untransformed milk BHB (0.10 to 0.64) or for log-transformed milk BHB (0.09 to 0.63; de Roos et al., 2007). Grelet et al. (2016) found $R^2_{cv}$ of 0.71 and $R^2_v$ of 0.63 for milk BHB, larger than estimates found in the current study. With blood BHB used as a reference value in calibration, Broutin (2015, 2016) also found higher $R^2_{cv}$ (0.7360 or 0.5999) than that observed in our study. The predictive ability of calibration models developed in the present study may not be sufficient to determine exact values of blood BHB, but may allow for a rough screening to distinguish cows with high or low values. It has been concluded that FT-MIR–predicted ketone bodies may be promising as screening tool for ketosis at herd level, but not accurate enough for management decisions at an individual animal level (de Roos et al., 2007; van der Drift et al., 2012; Grelet et al., 2016).

The correlations between reference BHB and predicted BHB obtained by the models developed in our study (averaged to 0.584) were higher than the correlation between reference blood BHB and Foss-predicted milk BHB (0.567). This indicates that these models may be more appropriate to indicate ketosis, as they predict blood BHB instead of milk BHB. It also shows the interest of predicting blood values directly from FT-MIR spectra rather than using milk BHB from FT-MIR spectra. The $R^2$ between reference and predicted blood BHB (Table 3) also indicate that milk spectra would contain substantial amount of information about

BHB. Reported phenotypic correlations between reference blood BHB and reference milk BHB vary widely, ranging from 0.66 to 0.89 (Enjalbert et al., 2001; Denis-Robichaud et al., 2014; Friedrichs et al., 2015); correlation coefficients found in the current study were in the lower range of the values reported in those studies. However, only Broutin (2016) reported on the correlation between reference blood BHB and predicted blood BHB from milk spectra, finding a correlation of 0.7370.

Several factors could contribute to the degree of accuracy of prediction models observed in our study. Relationship between blood BHB and milk spectra might not be linear, which could in part explain the observed low $R^2$. The $R^2$ of prediction models and concentration of analyte (e.g., fat composition and so on) are known to have direct relationships (Soyeurt et al., 2006; Rutten et al., 2009). Infrared absorbance is directly proportional to concentration of analyte or substance (Beer's law), indicating that analytes with very low concentrations (e.g., BHB) are difficult to detect by the FT-MIR spectroscopy. The concentration of BHB in milk is very low (21.7 mg/L given its molar mass of 104.11 g/mol), which is below the detection limit (100 mg/L) of the FT-MIR spectroscopy (Dardenne et al., 2015). Therefore, it is important to note that calibration of BHB in milk can only be done by indirect links with global milk composition, not by the specific spectral responses of BHB in milk (Grelet et al., 2016).

Moreover, the 2 information sources that were used in our study, milk spectra and blood BHB, were from different media (milk and blood). Genetic differences between cows in udder ketone body metabolism may exist and could influence excretion of ketone bodies from blood to milk (van der Drift et al., 2012). van der Drift et al. (2012) also found that the random effect of herd explained considerable variation in the probability of hyperketonemia for cows. Those authors explained the herd differences in the association between blood and milk ketone body concentrations by time of milk sampling, feeding, and blood sampling that were not identical on the different farms. Time of sampling (before or after feeding, morning or evening milking) could result in variation of BHB in blood and milk, as there might be difference in metabolism of BHB production in milk and blood.

### *Evaluation of the IP and DP Approaches*

The slightly better prediction of blood BHB from milk FT-MIR spectra by the IP than the DP approach was not in line with our expectation. It is also in contrast to the work of Dagnachew et al. (2013b), who reported better prediction in accuracy of EBV for milk contents in goat by DP than IP approaches. Bonfatti et al. (2017) also reported results that are mostly in contrast with the work of Dagnachew et al. (2013b), who found high rank correlations (>0.94) between IP and DP predicted EBV of all traits investigated. Bonfatti et al. (2017) reported <0.5 rank correlations between EBV predicted by IP and DP for most traits included in their study. Reasons why the DP approach resulted in better prediction for EBV (e.g., Dagnachew et al., 2013b), but not for phenotypic are not clear, but could be due to difference in methods of comparison (correlation vs prediction error variance) and type of parameters compared (phenotype vs EBV). Genetic parameter estimates (e.g., heritability) for BHB using covariance components (DP) after converting into univariate variance structure were higher (Table 6) than corresponding estimates using variance components (IP; Table 7), indicating better information content in the DP approach. However, neither phenotypic prediction from the multivariate mixed model using spectral variables that were reduced into few components by PLS (Table 8) nor principal component analysis (**PCA**; Belay et al., 2015) were promising. It is therefore important to verify if such results from the current study or previous studies (Dagnachew et al., 2013b; Belay et al., 2015) will be reproducible and to look for reasons behind the reported results for example using simulated data.

In the DP approach, dimension of spectral variables can be reduced into a few factor scores by PLS regression, as in the current paper, or into latent traits by PCA. Covariance components for the latent traits from PCA are population parameters that characterize any information coming from the population, as they represent any information available in the milk spectra, whereas covariance components for factor scores from PLS regression mainly contain information related to the particular trait used in the calibration. The PLS factors are thus expected to give better prediction of the trait than the one with latent traits from PCA. However, with PLS, information about other milk composition traits not included in the calibration may not be retained in the factor scores, as also indicated by Dagnachew et al. (2013b) and Bonfatti et al. (2017). The expected better prediction of traits with PLS model was confirmed. For example, prediction accuracy of DP was much lower than the IP approach when PCA was used (Belay et al., 2015) compared with when PLS was used for spectral dimension reduction (Table 8). One possible reason for this could be that the retained 8 latent traits from PCA, which explained 99% of the total spectral variation (Belay et al., 2015), did not contain as much relevant information about the blood BHB as those 5 PLS factors used in our study did. Dagnachew et al. (2013b) also used 8 latent traits to extract genetic component of the FT-MIR spectra and

indicated the possible existence of relevant information in the remaining 1% of the total spectral variation. Interestingly, Bonfatti et al. (2017) also showed that a considerable amount of information needed to predict phenotypes is lost when using 99% of original spectral variability, and loss of such information could affect prediction of EBV from spectral information. Those authors further showed that information left in 0.01% of original spectral variability is fundamental for prediction of some of the traits included.

Several possible reasons exist for the slightly lower prediction accuracy in the DP compared with the IP approach found in our study. It could be due to the low genetic correlations observed among the latent traits (factor scores). Expected improvement in accuracy of prediction from multivariate analysis would be due to its ability to account for the covariance among the traits. When the covariance or correlation among the traits are very low, multivariate analysis might not perform better than the univariate one. Any errors in the covariance estimation or the modeling of the observations may also have reduced the accuracy. The genetic and environmental parameters used in BLUP analysis are estimates and possibly contain errors. Use of estimates with errors in multivariate analysis would affect accuracy of prediction (Schaeffer, 1984; Thompson and Meyer, 1986). Response to selection, which is directly proportional to accuracy of predicted breeding values, highly depends on the precision of the estimates and the applied variance components (Villanueva et al., 1993). Under such conditions, univariate models can provide more precise estimates than multitrait models. Accuracies will also depend on the relevance of the models used, and whether or not anything can be gained by using the mixed model.

Lack of enough information about contemporary cows in the validation data set with blood BHB samples could be another major contributing factor to the poor performance observed in multivariate analysis. If no structure of the random effects of the model exists in the data to be predicted, there may be no benefit from using a multivariate mixed model. The number of cows in the validation set were small and each cow had only 1 measured blood BHB; hence, the cows in validation set could possibly be not well connected to each other genetically. In addition, almost all of the HTD classes contained only 1 cow, which might impose difficulty in separating herd effect from genetic effect of cows and contribute to low accuracy in the multivariate analysis. With FT-MIR spectra, for more contemporary cows in the validation, more info that is multivariate could have been available. An attempt was made to merge the validation set with data set 2 to increase amount of information in the validation set (such as increasing number cows in the HTD classes); however, as these 2 data sets were collected in different years, they had only 1 cow in common. Hence, we failed to use information in both data sets for better accuracy in multiple model.

In our study, correlations of blood BHB with milk fat (0.35), protein (−0.06), and lactose (−0.20) contents were low to medium, which might have contributed to the lower accuracy in DP. These milk contents were prediction from spectra by Foss calibration. For example, Bonfatti et al. (2017) found a positive relationship between rank correlation and correlations of traits of interest with major sources of spectral variation (such as milk protein, fat, and lactose contents), and that variability of the traits of interest is better explained when they are highly correlated with the major sources of spectral variation. In addition, for traits more correlated with the major sources of spectral variation, the DP approach is more likely to be effective (Bonfatti et al., 2017). Those authors indicated that the better accuracy of EBV from the DP than from the IP that Dagnachew et al. (2013b) reported would be due to large contribution of milk protein, fat, and lactose contents to spectral variation.

In addition, $R^2_{cv}$ of the calibration models could affect accuracy of the DP approach. For example, in Dagnachew et al. (2013b), $R^2_{cv}$ was very high (>0.94) and the multivariate model performed better than the univariate one, possibly as a consequence of this. In the current study, the predictive ability of the calibration models was much lower (explained less than 45% of the variation in the blood BHB, Table 3), and the multivariate models performed slightly inferior to the univariate ones (Table 8). From these 2 studies, we can see that accuracy of predicting breeding values or phenotype seems to depend on predictive ability of the calibration model. This could lead us to the conclusion that, for DP to work better, there should be a strong relationship between the trait of interest and spectral variables. In other words, we should first be sure that the univariate method (IP) is working well with the data on hand before embarking on the DP method. Under such conditions, where the univariate method was found to be working, the multivariate method might perform better, but this should be an assumption that needs to be made. Bonfatti et al. (2017) reported that rank correlations between EBV obtained by the IP and the DP approach are not related to the accuracy the calibration equations; however, the relationships between accuracies of EBV obtained by the 2 approaches and accuracy of calibration equations are not well established.

Both the IP and DP approaches had lower predictability for the phenotype than the predictability of equations developed based on the classical PLS regression.

This indicated that inclusion of cows' circumstances at a given test day into the IP or DP model did not improve prediction of blood BHB from milk FT-MIR spectra. Therefore, for phenotypic prediction, the classical PLS regression-based prediction equation seems to be the method of choice. In our study, information related to the cow was added into the models after PLS regression or in the mixed model analysis of predicted trait or factor scores. On the other hand, it has been shown that it is possible to directly add the information to the spectra before PLS. For example, Vanlierde et al. (2015) included DIM directly into spectra using Legendre polynomial to predict methane, and prediction equations developed in such a way were shown to be more robust than equations that did not integrate the DIM information. Similarly, Shetty et al. (2017) used milk yield and live weight as predictors along with spectral variables to predict residual feed intake and DMI. They showed improvement in accuracy of model that included spectral information along with milk yield and live weight as predictors for DMI. Therefore, inclusion of cows' circumstances directly into spectra before PLS or using them as predictors along with spectral information during PLS can be an alternative to improve prediction accuracy for blood BHB from milk FT-MIR spectra.

## CONCLUSIONS

In this study, predictive ability of the DP and IP approaches were evaluated using measured blood BHB and milk spectra-predicted blood BHB. A calibration and an independent validation data set was used. Accuracy of prediction with the 2 approaches were similar. Slightly better prediction of BHB was found when univariate variance structure was used (IP) than when multitrait covariance structures were used (DP) in mixed models. Prediction accuracies of the developed calibration models were also low, which could partly be due to a weak relationship between milk spectra and blood BHB. Blood BHB log-transformation, spectral preprocessing, and use of extreme blood BHB values have improved prediction accuracy of the calibration models and the 2 approaches. Conclusive remarks on the importance of keeping spectral data in a multivariate form for prediction of phenotype and model components (EBV, HTD, and so on) may be found in data sets where the trait of interest has strong relationships with spectral variables.

## ACKNOWLEDGMENTS

The authors gratefully acknowledge Polish Federation of Cattle Breeders and Dairy Farmers (Warsaw, Poland) for providing data for this work. The authors acknowledge especially K. Słoniewski (Polish Federation of Cattle Breeders and Dairy Farmers) for transferring the data and providing explanation on the data when needed. The authors also thank Artur and Wojciech Jagusiak (University of Agriculture in Krakow, Poland) for helping us to understand the data and construct the pedigree file, respectively. The authors also acknowledge Achim Kohler and Valeria Tafintseva (Norwegian University of Life Sciences) for their help in model calibration.

## REFERENCES

Afseth, N. K., H. Martens, Å. Randby, L. Gidskehaug, B. Narum, K. Jørgensen, S. Lien, and A. Kohler. 2010. Predicting the fatty acid composition of milk: A comparison of two Fourier transform infrared sampling techniques. Appl. Spectrosc. 64:700–707.

Andersen, S. K., P. W. Hansen, and H. V. Andersen. 2002. Vibrational spectroscopy in the analysis of dairy products and wine. Accessed Jul. 14, 2016. www.foss.dk/~/media/B4E6F6119A7548F4B380C-3E6C958DBD1.ashx.

Andersson, L. 1988. Subclinical ketosis in dairy cows. Vet. Clin. North Am. Food Anim. Pract. 4:233–251.

Belay, T. K., K. Słoniewski, Z. Kowalski, and T. Adnoy. 2015. Predicting ketosis from milk mid infrared (MIR) spectra using multivariate mixed models. In Proc. Third DairyCare Conference, Croatia, Zadar. Accessed Jun. 20, 2016. www.dairycareaction.org/uploads/2/4/2/6/24266896/belay_tuesday.pdf.

Bonfatti, V., D. Vicario, L. Degano, A. Lugo, and P. Carnier. 2017. Comparison between direct and indirect methods for exploiting Fourier transform spectral information in estimation of breeding values for fine composition and technological properties of milk. J. Dairy Sci. 100:2057–2067.

Broutin, P. 2015. Blood BHB determination by mid infrared spectroscopy for the monitoring of the cows metabolic activity and detection of subclinical ketosis—A new approach. ICAR Technical Workshop. Accessed Sep. 20, 2016. www.icar.org/wp-content/uploads/2015/09/4-Pierr.pdf.

Broutin, P. 2016. Determination of the concentration of a component in one fluid of an animal by spectroscopic analysis of another fluid. United States, Bentley Instrument (Lille, FR) 20160238520. Accessed Oct. 12, 2016. http://www.freepatentsonline.com/y2016/0238520.html.

Dagnachew, B. S., A. Kohler, and T. Ådnøy. 2013a. Genetic and environmental information in goat milk Fourier transform infrared spectra. J. Dairy Sci. 96:3973–3985.

Dagnachew, B. S., T. H. E. Meuwissen, and T. Ådnøy. 2013b. Genetic components of milk Fourier-transform infrared spectra used to predict breeding values for milk composition and quality traits in dairy goats. J. Dairy Sci. 96:5933–5942.

Dardenne, P., J. Fernandez, O. Baeten, and B. Lecler. 2015. Untargeted contaminant detection by IR. Wallon Agricultural Research Centre. Accessed Dec. 15, 2016. http://www.namur2015.be/docs/pdf/Session%205%20-%202%20-%20Dardenne.pdf.

de Roos, A., H. Van Den Bijgaart, J. Hørlyk, and G. De Jong. 2007. Screening for subclinical ketosis in dairy cattle by Fourier transform infrared spectrometry. J. Dairy Sci. 90:1761–1766.

Denis-Robichaud, J., J. Dubuc, D. Lefebvre, and L. DesCôteaux. 2014. Accuracy of milk ketone bodies from flow-injection analysis for the diagnosis of hyperketonemia in dairy cows. J. Dairy Sci. 97:3364–3370.

Duffield, T. F., K. Lissemore, B. McBride, and K. Leslie. 2009. Impact of hyperketonemia in early lactation dairy cows on health and production. J. Dairy Sci. 92:571–580.

Duffield, T. F., D. F. Kelton, K. E. Leslie, K. D. Lissemore, and J. H. Lumsden. 1997. Use of test day milk fat and milk protein to

detect subclinical ketosis in dairy cattle in Ontario. Can. Vet. J. 38:713–718.
Enjalbert, F., M. Nicot, C. Bayourthe, and R. Moncoulon. 2001. Ketone bodies in milk and blood of dairy cows: Relationship between concentrations and utilization for detection of subclinical ketosis. J. Dairy Sci. 84:583–589.
Friedrichs, P., C. Bastin, F. Dehareng, B. Wickham, and X. Massart. 2015. Final OptiMIR Scientific and Expert Meeting: From milk analysis to advisory tools. Palais des Congrès, Namur, Belgium, 16-17 April 2015. Biotechnol. Agron. Soc. Environ. 19:97–124.
Geishauser, T., K. Leslie, J. Tenhag, and A. Bashiri. 2000. Evaluation of eight cow-side ketone tests in milk for detection of subclinical ketosis in dairy cows. J. Dairy Sci. 83:296–299.
Gilmour, A. R., B. Gogel, B. Cullis, R. Thompson, and D. Butler. 2009. ASReml user guide release 3.0. VSN International Ltd., Hemel Hempstead, UK.
Grelet, C., C. Bastin, M. Gelé, J.-B. Davière, M. Johan, A. Werner, R. Reding, J. F. Pierna, F. Colinet, and P. Dardenne. 2016. Development of Fourier transform mid-infrared calibrations to predict acetone, β-hydroxybutyrate, and citrate contents in bovine milk through a European dairy network. J. Dairy Sci. 99:4816–4825.
Iñón, F. A., S. Garrigues, and M. de la Guardia. 2004. Nutritional parameters of commercially available milk samples by FTIR and chemometric techniques. Anal. Chim. Acta 513:401–412.
LeBlanc, S. J., K. Leslie, and T. Duffield. 2005. Metabolic predictors of displaced abomasum in dairy cattle. J. Dairy Sci. 88:159–170.
Madsen, P., and J. Jensen. 2008. DMU: A User's Guide. A package for analysing multivariate mixed models. Version 6, release 4.7. DJF, Foulum, Denmark.
Mäntysaari, E. 1999. Derivation of Multiple Trait Reduced Random Regression (RR) Model for the First Lactation Test Day Records of Milk, Protein and Fat. Page 8 in 50th Annual Meeting. Europ. Ass. Anim. Prod. Mimeo. Zurich, Switzerland, Aug. 23–26, 1999.
McArt, J. A., D. Nydam, and G. Oetzel. 2012. Epidemiology of subclinical ketosis in early lactation dairy cattle. J. Dairy Sci. 95:5056–5066.
Mevik, B.-H., and R. Wehrens. 2007. The pls package: Principal component and partial least squares regression in R. J. Stat. Softw. 18:1–24.
Meyer, K. 2007. WOMBAT—A tool for mixed model analyses in quantitative genetics by restricted maximum likelihood (REML). J. Zhejiang Univ. Sci. B 8:815–821.
Oetzel, G. R. 2004. Monitoring and testing dairy herds for metabolic disease. Vet. Clin. North Am. Food Anim. Pract. 20:651–674.
R Core Team. 2016. R: A language and environment for statistical computing R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.
Rutten, M. J., H. Bovenhuis, K. Hettinga, H. van Valenberg, and J. Van Arendonk. 2009. Predicting bovine milk fat composition using infrared spectroscopy based on milk samples collected in winter and summer. J. Dairy Sci. 92:6202–6209.
Schaeffer, L. 1984. Sire and cow evaluation under multiple trait models. J. Dairy Sci. 67:1567–1580.
Seifi, H. A., S. J. LeBlanc, K. E. Leslie, and T. F. Duffield. 2011. Metabolic predictors of post-partum disease and culling risk in dairy cattle. Vet. J. 188:216–220.
Shetty, N., P. Løvendahl, M. Lund, and A. Buitenhuis. 2017. Prediction and validation of residual feed intake and dry matter intake in Danish lactating dairy cows using mid-infrared spectroscopy of milk. J. Dairy Sci. 100:253–264.
Soyeurt, H., P. Dardenne, F. Dehareng, G. Lognay, D. Veselko, M. Marlier, C. Bertozzi, P. Mayeres, and N. Gengler. 2006. Estimating fatty acid content in cow milk using mid-infrared spectrometry. J. Dairy Sci. 89:3690–3695.
Soyeurt, H., I. Misztal, and N. Gengler. 2010. Genetic variability of milk components based on milk infrared spectral data. J. Dairy Sci. 93:1722–1728.
Suthar, V. S., J. Canelas-Raposo, A. Deniz, and W. Heuwieser. 2013. Prevalence of subclinical ketosis and relationships with postpartum diseases in European dairy cows. J. Dairy Sci. 96:2925–2938.
Thompson, R., and K. Meyer. 1986. A review of theoretical aspects in the estimation of breeding values for multi-trait selection. Livest. Prod. Sci. 15:299–313.
van der Drift, S., R. Jorritsma, J. Schonewille, H. Knijn, and J. Stegeman. 2012. Routine detection of hyperketonemia in dairy cows using Fourier transform infrared spectroscopy analysis of β-hydroxybutyrate and acetone in milk in combination with test-day information. J. Dairy Sci. 95:4886–4898.
Vanlierde, A., M.-L. Vanrobays, F. Dehareng, E. Froidmont, H. Soyeurt, S. McParland, E. Lewis, M. H. Deighton, F. Grandl, and M. Kreuzer. 2015. Hot topic: Innovative lactation-stage-dependent prediction of methane emissions from milk mid-infrared spectra. J. Dairy Sci. 98:5740–5747.
Villanueva, B., N. Wray, and R. Thompson. 1993. Prediction of asymptotic rates of response from selection on multiple traits using univariate and multivariate best linear unbiased predictors. Anim. Sci. 57:1–13.
Walsh, R. B., J. Walton, D. Kelton, S. LeBlanc, K. Leslie, and T. Duffield. 2007. The effect of subclinical ketosis in early lactation on reproductive performance of postpartum dairy cows. J. Dairy Sci. 90:2788–2796.

# Paper III

**Genetic parameters of blood β-hydroxybutrate and clinical ketosis, and their associations with milk production traits in Norwegian Red cows**

T.K. Belay, M. Svendsen, Z.M. Kowalski and T. Ådnøy

*Journal of Dairy Science (In Press)*

J. Dairy Sci. 100:1–14
https://doi.org/10.3168/jds.2016-12458


# Genetic parameters of blood β-hydroxybutyrate predicted from milk infrared spectra and clinical ketosis, and their associations with milk production traits in Norwegian Red cows

**T. K. Belay,*[1] M. Svendsen,† Z. M. Kowalski,‡ and T. Ådnøy***
*Department of Animal and Aquacultural Sciences, Norwegian University of Life Sciences, PO Box 5003, 1432 ÅS, Norway
†GENO Breeding and A.I. Association, PO Box 5025, N-1432 Ås, Norway
‡Department of Animal Nutrition and Dietetics, University of Agriculture in Krakow, Krakow 30-059, Al. Mickiewicza 24/28, Poland

## ABSTRACT

The aim of this study was to estimate genetic parameters for blood β-hydroxybutyrate (BHB) predicted from milk spectra and for clinical ketosis (KET), and to examine genetic association of blood BHB with KET and milk production traits (milk, fat, protein, and lactose yields, and milk fat, protein, and lactose contents). Data on milk traits, KET, and milk spectra were obtained from the Norwegian Dairy Herd Recording System with legal permission from TINE SA (Ås, Norway), the Norwegian Dairy Association that manages the central database. Data recorded up to 120 d after calving were considered. Blood BHB was predicted from milk spectra using a calibration model developed based on milk spectra and blood BHB measured in Polish dairy cows. The predicted blood BHB was grouped based on days in milk into 4 groups and each group was considered as a trait. The milk components for test-day milk samples were obtained by Fourier transform mid-infrared spectrometer with previously developed calibration equations from Foss (Hillerød, Denmark). Veterinarian-recorded KET data within 15 d before calving to 120 d after calving were used. Data were analyzed using univariate or bivariate linear animal models. Heritability estimates for predicted blood BHB at different stages of lactation were moderate, ranging from 0.250 to 0.365. Heritability estimate for KET from univariate analysis was 0.078, and the corresponding average estimate from bivariate analysis with BHB or milk production traits was 0.002. Genetic correlations between BHB traits were higher for adjacent lactation intervals and decreased as intervals were further apart. Predicted blood BHB at first test day was moderately genetically correlated with KET (0.469) and milk traits (ranged from −0.367 with protein content to 0.277 with milk yield), except for milk fat content from across lactation stages that had near zero genetic correlation with BHB (0.033). These genetic correlations indicate that a lower BHB is genetically associated with higher milk protein and lactose contents, but with lower yields of milk, fat, protein, and lactose, and with lower frequency of KET. Estimates of genetic correlation of KET with milk production traits were from −0.333 (with protein content) to 0.178 (with milk yield). Blood BHB can routinely be predicted from milk spectra analyzed from test-day milk samples, and thereby provides a practical alternative for selecting cows with lower susceptibility to ketosis, even though the correlations are moderate.
**Key words:** β-hydroxybutyrate, clinical ketosis, milk trait, genetic parameter

## INTRODUCTION

Clinical ketosis (**KET**) is one of the most frequent metabolic diseases affecting dairy cattle. In a recent literature review, Pryce et al. (2016) found a median ketosis frequency of 3.3% with a range from 0.24% in first lactation up to 17.2% in third lactation. Those authors reported a median incident rate of 10.25% for Norwegian Red (**NRF**) cows based on 2 previous studies. The frequency of KET in NRF cows has decreased markedly since the mid-1980s, from 10.6% in 1987 to 4.3% in 1998 in first-lactation cows (Heringstad et al., 2005) and from 23.88% in 1985 to 4.56% in 2005 (Østerås et al., 2007).

In Norway, health data including KET have been recorded on an individual cow basis since 1978 based on veterinarian treatments (Heringstad et al., 2000). Many metabolic events, including ketosis, are subclinical by nature, and information on subclinical cases is mostly missing in records because it is difficult to detect (Pryce et al., 2016). Subclinical cases are assumed to receive less veterinary intervention, thus leading to underestimation of the incidence in systems that depend on veterinary data (Schwarzenbacher et al., 2010). Failure

Received December 14, 2016.
Accepted April 5, 2017.
[1]Corresponding author: tesfaye.kebede.belay@nmbu.no

to detect subclinical events can be expensive to dairy producers, as it negatively affects overall performance of cows (Duffield et al., 2009); therefore, systems to detect ketosis at a subclinical stage in addition to the clinical one would be useful.

Previous genetic studies of ketosis were mostly based on clinical records. Heritability estimates for KET reported in those studies ranged from 0.01 to 0.16, as summarized in the literature review by Pryce et al. (2016). Heritability estimates for KET may be influenced by the subjective nature of its diagnosis and the low frequency of KET compared with subclinical ketosis in cows (van der Drift et al., 2012b). Moreover, response to selection in KET is hampered by low reliabilities often associated with the low heritability (Pryce et al., 2016). Pryce et al. (2016) suggested that information from correlated traits or from subclinical diagnosis could be used to improve the accuracy of predicted breeding values and increase the selection response. Phenotypes derived from routinely collected data through milk recording, such as fat-to-protein ratio and fatty acid profiles, are promising ketosis indicators (van Knegsel et al., 2010). Phenotypes more closely associated with ketosis, such as BHB and acetone in milk (Pryce et al., 2016), may also be valuable. Concentration of BHB in blood has been used as a gold standard indicator of ketosis, and thresholds of 1.2 (van der Drift et al., 2012a) or 1.4 mmol/L (Denis-Robichaud et al., 2014) have been used to identify cows with subclinical ketosis. However, the gold standard method does not allow routine testing of all animals at risk due to practical limitations, such as difficulty in blood sampling (especially for farmers) and capacity for analyzing many blood samples at a time.

Routine measurements of ketone bodies in milk can be done by Fourier transform mid-infrared (**FT-MIR**) spectrometer analysis of milk samples at test-days (de Roos et al., 2007; van der Drift et al., 2012a; Grelet et al., 2016). Those previous studies agreed that FT-MIR predicted milk ketone bodies adequately and that FT-MIR might be useful for ketosis-screening purposes. In addition to their routine availability, indicator traits (e.g., milk or blood BHB) have moderate heritability (0.07 to 0.40; Oikonomou et al., 2008; van der Drift et al., 2012b; Jamrozik et al., 2016). Milk BHB has also moderate genetic correlations with KET, with estimates ranging from 0.25 to 0.75 (Koeck et al., 2014, 2016; Jamrozik et al., 2016); hence, indirect selection for ketosis using BHB as an indicator trait should result in better genetic gain than direct selection for KET. For example, Koeck et al. (2016) estimated about 65% more selection response from indirect selection for the indicator trait (BHB) than direct selection for ketosis.

Ketone bodies have not routinely been measured by FT-MIR spectrometer analysis of milk samples at test days in Norway. Given the subclinical nature of ketosis, the potential of FT-MIR for routine prediction of ketosis indicators, and the moderate heritability and genetic correlations of the indicator traits with KET, it is important to assess the potential of FT-MIR predicted phenotypes (e.g., BHB) for their use in dairy farm management and breeding programs. For use and implementation of blood BHB predicted from milk spectra in dairy cattle breeding programs, knowledge of genetic parameters and their genetic associations with clinical events and other traits in the breeding goal is essential. Few reports exist on genetic studies of plasma BHB measured by reference methods (Oikonomou et al., 2008; van der Drift et al., 2012b); however, there is no report on genetic parameters and associations of blood BHB predicted from milk spectra with KET and milk production traits for cows in early lactation. Therefore, the objective of our study were (1) to describe the phenotype of predicted blood BHB and to examine its phenotypic associations with KET and milk production traits; (2) to estimate genetic parameters for the predicted blood BHB traits and KET; and (3) to determine genetic associations between the predicted blood BHB and KET and milk production traits (milk, fat, protein, and lactose yields, and milk fat, protein, and lactose contents) in NRF cows.

## MATERIALS AND METHODS

### *Data and Data Edits*

We used 3 data sets, referred to as data set 1, data set 2, and data set 3.

***Data Set 1.*** This data set consisted of test-day predicted blood BHB and milk production traits. Over 5 million FT-MIR spectra from milk test-day samples recorded from February 2007 to June 2014 were obtained from the Norwegian dairy herd recording system. The milk spectra were from test-day milk samples analyzed by FT-MIR spectrometer (Milkoscan Combifoss 6500, Foss Electric, Hillerød, Denmark). Blood BHB was predicted from milk spectra of the NRF cows with a previously developed calibration model for blood BHB from milk spectra and reference blood BHB of Polish dairy cows (Belay et al., 2017) by permission from our Polish collaborators (Polish Federation of Cattle Breeders and Dairy Farmers, Warsaw, Poland). The data set consisted of 826 Polish dairy cows (1 observation per cow), from 5 to 65 DIM, with measured blood BHB and milk spectra. It was randomly divided into a calibration (n = 496) and a validation (n = 330) set. The calibration model was developed by partial least square (**PLS**) regression, internally cross-validated using 10 random segments, and externally validated using the

**Table 1**. Number of records and summary statistics of data set 1 for blood BHB predicted from milk spectra at 11 to 30 ($BHB_1$), 31 to 60 ($BHB_2$), 61 to 90 ($BHB_3$), and 91 to 120 DIM ($BHB_4$) and across early lactation stage (11–120 DIM; $BHB_{all}$)

| Trait | Period, DIM | No. of records | Mean, mmol/L | SD, mmol/L | Minimum, mmol/L | Maximum, mmol/L |
|---|---|---|---|---|---|---|
| $BHB_1$ | 11–30 | 241,543 | 1.242 | 0.337 | −1.919 | 6.133 |
| $BHB_2$ | 31–60 | 350,560 | 1.209 | 0.293 | −4.047 | 6.317 |
| $BHB_3$ | 61–90 | 327,462 | 1.170 | 0.266 | −1.163 | 6.050 |
| $BHB_4$ | 91–120 | 307,573 | 1.157 | 0.256 | −3.264 | 4.211 |
| $BHB_{all}$ | 11–120 | 1,227,138 | 1.192 | 0.289 | −4.047 | 6.317 |

validation set. Optimum number of PLS factors (in this case 6 factors) were determined based on first local minimum value in root mean squared error of prediction. The PLS regression coefficients were applied on the Norwegian milk spectra to predict blood BHB. The predicted blood BHB was merged with milk production traits and other farm and cow information. Milk production traits, such as milk fat, protein, and lactose contents, were also predictions from the same spectra with machine integrated operational calibrations for the respective traits (Foss Analytical A/S, Hillerød, Denmark).

The predicted blood BHB and milk production traits were kept for cows in 11 to 120 DIM. Cows with unknown sires or dams, herds with less than 200 test day records, and sires with less than 25 daughters were excluded from further analysis. Only cows with age at calving of 18 to 40, 30 to 51, 42 to 63, and 52 to 74 mo in the first, second, third, and fourth lactations, respectively, were considered. To get a reasonable number of records for age classes at peripheries in each age category, the first and last few age classes were merged into the next and preceding age class, respectively. Number of records per herd test date (**HTD**) were kept to at least 2. Twelve observations with extreme values (potential outliers) were also removed from the data set. The final edited data set contained 1,227,138 test day records from 324,920 cows that were progeny of 1,427 sires and kept in 3,539 herds. Statistical analyses were carried out for the BHB across lactation number in early lactation at periods of 11 to 30 (**$BHB_1$**), 31 to 60 (**$BHB_2$**), 61 to 90 (**$BHB_3$**), and 91 to 120 DIM (**$BHB_4$**), and across all periods of early lactation, 11 to 120 DIM (**$BHB_{all}$**). Descriptive statistics of BHB are presented in Table 1. Summary of descriptive statistics for the milk production traits across the lactation stages considered are shown in Table 2. The BHB was found to be reasonably normally distributed (Figure 1) and not log-transformed for genetic analysis. A pedigree file containing animals with records and their ancestors was also available with a total number of 671,849 individuals.

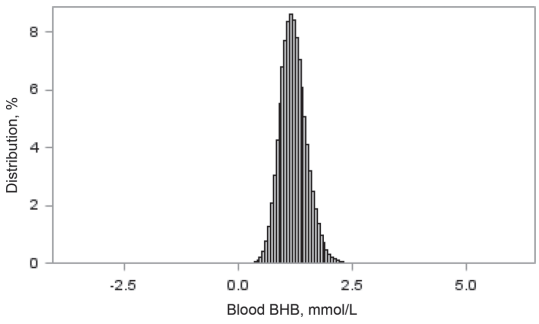



**Figure 1**. Distribution (%) of blood BHB concentrations (mmol/L) predicted from milk spectra of Norwegian Red cows in data set 1.

***Data Set 2.*** This data set contained information on KET. It consisted of 1,742,421 observations on cows in lactation 1 to 4 that gave birth from January 2007 through December 2015. Absence or presence of KET was scored as 0 or 1, respectively, based on whether or not the cow received veterinary treatment between 15 d before calving to 120 d after calving. For genetic analysis, animals with no sire or dam information and cows with less than 2 records were excluded. The final edited data set consisted of 1,054,381 records from 357,474 cows that were progeny of 2,455 sires and 282,657 dams and kept in 12,533 herds. Statistical analyses were carried out for the KET across lactation numbers. A summary of descriptive statistics of the analyzed data set is presented in Table 2. A pedigree file for animals with records and their ancestors was also available and contained 886,401 individuals.

***Data Set 3.*** The unedited data set of predicted blood BHB and milk production traits was merged with the unedited ketosis data set, keeping only cows with information on BHB. The merged data (data set 3) contained 1,520,446 records from 430,389 cows kept in 11,697 herds. Bivariate genetic analysis of such a large data set was not feasible as the number of elements in mixed model equations were beyond the limit of the program used for REML estimates. Cows with no sire or dam information, small herds (with <225 records

**Table 2**. Number of records and summary statistics for milk production traits based on data set 1 and for clinical ketosis as binary trait (0 = healthy, 1 = treated) across lactations based on data set 2

| Trait | No. of records | Mean | SD | Minimum | Maximum |
|---|---|---|---|---|---|
| Milk yield, kg | 1,227,138 | 29.215 | 6.992 | 5 | 50 |
| Fat yield, kg | 1,227,138 | 1.178 | 0.351 | 0.094 | 3.440 |
| Protein yield, kg | 1,227,138 | 0.945 | 0.219 | 0.127 | 2.864 |
| Lactose yield, kg | 1,227,138 | 1.385 | 0.333 | 0.090 | 2.625 |
| Fat, % | 1,227,138 | 4.050 | 0.789 | 1.750 | 7.000 |
| Protein, % | 1,227,138 | 3.250 | 0.254 | 1.350 | 6.980 |
| Lactose, % | 1,227,138 | 4.743 | 0.180 | 1.100 | 5.600 |
| Ketosis, % | 1,054,381 | 2.040 | 1.414 | 0 | 1 |

over all lactations), and sires with less than 30 daughters were excluded. The final edited merged data set contained 717,915 records from 179,691 cows that were a progeny of 1,169 sires and 135,985 dams and kept in 2,828 herds. The pedigree file of animals with records and their ancestors contained 469,672 individuals.

### Models

Data were analyzed with mixed linear animal models using the REML method with parameter expanded and average information algorithm (PX-AI) of the WOMBAT software (Meyer, 2007) or the DMU package (Madsen and Jensen, 2008). Univariate analyses of BHB traits, bivariate analyses among BHB traits, and BHB traits with milk production traits were done using WOMBAT (Meyer, 2007). Bivariate analysis of KET with either BHB or milk production traits were done using the DMU package (Madsen and Jensen, 2008), as DMU allows to fit separate models for each trait in bivariate analyses.

***Genetic Parameters for Blood BHB.*** The test day-predicted blood BHB records were considered as repeated measurement of the same trait across and within lactations. We assumed constant genetic variance across lactation stages and unity genetic correlations between test day records. Hence, predicted blood BHB across stages of early lactation (11–120 DIM) was analyzed with a single-trait repeatability test day animal model. The test day records of BHB were also grouped into adjacent DIM intervals, and each group was treated as a distinct trait and analyzed with single- or multiple-trait repeatability test day animal models. Single- and 2-trait repeatability test day animal models were applied to BHB at 11 to 30, 31 to 60, 61 to 90, and 91 to 120 DIM using the data set 1. In matrix notation, the models were as follows

$$\mathbf{y} = \mathbf{Xb} + \mathbf{Za} + \mathbf{Wc} + \mathbf{Hd} + \mathbf{e}, \qquad [1]$$

where $\mathbf{y}$ is the vector of predicted blood BHB at different DIM intervals or across DIM; $\mathbf{b}$ is a vector of fixed effects of region × year × month of test day, region × parity × DIM, herd × year of test day, parity × age at calving, and housing × milking system; $\mathbf{a}$ is a vector of random animal additive genetic effects; $\mathbf{c}$ is a vector of random permanent environmental ($\mathbf{PE}$) effects due to the cow; $\mathbf{d}$ is a vector of random HTD effects; $\mathbf{e}$ is vector of random residual effects; and $\mathbf{X}$, $\mathbf{Z}$, $\mathbf{W}$, and $\mathbf{H}$ are design matrices that relate records to the corresponding effects. Region had 9 levels; DIM was defined in 3 (or 11) classes of 10 d each. Interaction of housing type (tiestall or loose housing) and milking system (robot milking or manual milking: bucket, pipe, and milking parlor) was also modeled. All these fixed effects significantly ($P < 0.001$) affected the traits.

The assumed (co)variance structure in the single-trait analysis was $\text{var}(a) = \sigma_a^2\mathbf{A}$, $\text{var}(p) = \sigma_{pe}^2\mathbf{I}$, $\text{var}(d) = \sigma_d^2\mathbf{I}$, and $\text{var}(e) = \sigma_e^2\mathbf{I}$, where $\sigma_a^2$ was additive genetic variance, $\sigma_{pe}^2$ was PE variance, $\sigma_d^2$ was HTD variance, and $\sigma_e^2$ was residual variance. The $\mathbf{I}$ were identity matrices of appropriate sizes and $\mathbf{A}$ was the additive relationship matrix.

For 2-trait analyses, the following covariance structures were assumed for the random effect vectors included in the models:

$$\text{var}\begin{bmatrix}\mathbf{a}\\ \mathbf{c}\\ \mathbf{d}\\ \mathbf{e}\end{bmatrix} = \begin{bmatrix}\mathbf{G}\otimes\mathbf{A} & 0 & 0 & 0\\ 0 & \mathbf{P}\otimes\mathbf{I} & 0 & 0\\ 0 & 0 & \mathbf{J}\otimes\mathbf{I} & 0\\ 0 & 0 & 0 & \mathbf{R}\otimes\mathbf{I}\end{bmatrix},$$

where $\mathbf{G}$ was the genetic covariance matrix, $\mathbf{P}$ was the PE covariance matrix, $\mathbf{J}$ was the covariance matrix for HTD effects, and $\mathbf{R}$ was the residual covariance matrix. All covariance matrices were 2 × 2; $\mathbf{I}$ and $\mathbf{A}$ were as defined above; and ⊗ was the Kronecker product.

***Genetic Parameters for KET.*** A single-trait linear repeatability animal model was applied on data set 2 to estimate genetic parameters for KET. Threshold models are believed to be more appropriate to analyze binary traits, at least in theory. In previous studies on

NRF health data, multivariate threshold models were used (Heringstad et al., 2005); however, mixed linear models were applied in the current study. In their literature review, Pryce et al. (2016) indicated that linear models have performed equally well and are comparable to results from threshold models; several recent studies have also used linear models for ketosis (Koeck et al., 2014, 2016; Jamrozik et al., 2016). Moreover, genetic correlations are reported to be correct for binary traits using linear models (Negussie et al., 2008). In matrix notation, the following linear animal model was applied to the binary ketosis trait:

$$\mathbf{ket} = \mathbf{Xb} + \mathbf{Za} + \mathbf{Wc} + \mathbf{Hy} + \mathbf{e}, \qquad [2]$$

where **ket** was a vector of ketosis coded as 0 or 1 (per cow and lactation); **b** was a vector of systematic effects, including region × year× month of calving, parity × age at calving, and housing × milking system; **a** was a vector of random animal additive genetic effects; **c** was a vector of cows' random PE effects; **y** was a vector of random herd-year-month (**HY**) of calving effects; **e** was vector of random residual effects; and **X**, **Z**, **W**, and **H** were design matrices that relate records to the corresponding effects. Classes for age at calving were formed in the same way as for equation [1]. Assumptions of variance structures were also the same as in equation [1] for the single-trait analysis.

***Genetic Associations Among BHB, KET, and Milk Production Traits.*** For bivariate analysis of BHB traits with milk production traits, the same model as in equation [1] was applied using data set 1. Bivariate analyses of BHB traits or milk production traits with KET were done using data set 3. In the bivariate analysis of BHB traits or milk production traits with KET, 2 models were fitted: equation [1] for BHB traits or milk production traits, and equation [2] for KET using the AI-REML procedure in the DMU package (Madsen and Jensen, 2008).

## RESULTS AND DISCUSSION

### *Phenotypic Description of BHB*

Concentrations of blood BHB predicted from milk spectra ranged from −4.047 to 6.317 mmol/L, with an average of 1.192 mmol/L and a standard deviation of 0.289 mmol/L (Table 1). This was slightly higher than reference blood BHB values presented by Denis-Robichaud et al. (2014; average of 1.14 mmol/L, with values ranging from 0.2 to 6.3) and Belay et al. (2017), who found an average of 0.760 mmol/L with values ranging from 0.1 to 6.3 for Polish blood sample data. These values were also higher than predicted blood BHB from Polish dairy cows with the same calibration model as that used in the current study (Belay et al., 2017; average of 0.770 mmol/L, with values ranging from 0.492 to 3.95 mmol/L). Higher predicted blood BHB values in the current study than our previous study might have resulted from differences in spectral profile due to management (e.g., feeding), breed, or equipment used to produce spectra (Milkoscan Combifoss 6500 vs. MilkoScan FT6000). Unlike the reference methods, FT-MIR spectroscopy analysis may produce negative values; they account for about 0.004% of the observations in data set 1. These negative values may suggest very low concentrations of BHB in milk. Both the negative and nonnegative predicted blood BHB values were considered in the genetic analysis to be existing variation in the data.

Mean predicted blood BHB concentration at each DIM was calculated and is depicted in Figure 2. Generally, the mean predicted blood BHB concentration was higher in the beginning of lactation and decreased as DIM progressed. This result was in line with previous reports on milk BHB predicted from milk spectra (Koeck et al., 2014) and blood BHB measured with reference method (Belay et al., 2017). Predicted blood BHB decreased up to 20 DIM, and then increased again between 20 and 30 DIM. This is in agreement with a report of Oetzel (2007), who found an increase in predicted blood BHB from 20 to 35 DIM. The rise in BHB between 20 and 30 DIM might be due to ketosis type I that occurs between 3 and 6 wk postcalving, because cows are entering a peak lactation and these cows simply cannot keep up with energy demand mostly because of underfeeding (Oetzel, 2007).

We wanted to determine if models developed for Polish dairy cows could work and give some reasonable results with Norwegian data. We were surprised to find that the performance of the models with Norwegian

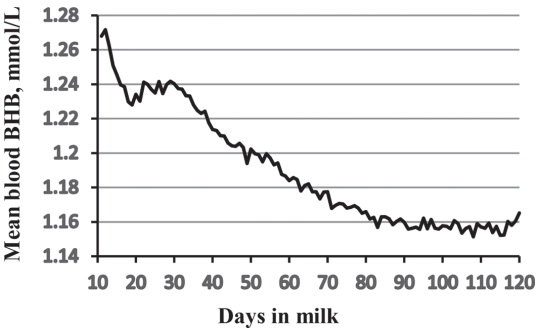



**Figure 2**. Mean blood BHB (mmol/L) predicted from milk spectra by DIM in first to fourth lactation Norwegian Red cows in data set 1.

data were reasonable. The phenotypic distribution of the predicted blood BHB, its heritability, and phenotypic and genetic associations with ketosis and milk traits were reasonable and in agreement with most of the published values. More interestingly, at a threshold ($\geq$1.2 mmol/L) that is commonly used to discriminate heathy from ketotic cows, the predicted blood BHB correctly classified more than 77% of the treated cows as ketotic cows. This is a practical validation of the model and does not mean that the models perfectly fitted the Norwegian data. The models might better fit data from the population that was used in the model development. Effects of different breeds and environments would be present, but might not hinder the use of the model in a different breed and environment where a calibration of milk spectra to blood BHB is not available.

### Frequency of KET

Mean frequency and standard deviation of veterinary-diagnosed KET across lactations based on data set 2 are given in Table 2. The mean frequency of KET across lactations that was 2.04% in data set 2 was reduced to 1.76% in data set 3, as not all animals diagnosed for KET might have recorded spectra or vice versa. Frequency of KET observed in the current study was lower than in previous studies on NRF cows. Using data from only first-crop daughters of NRF sires that were progeny tested from 1978 through 1998, Heringstad et al. (2005) reported mean frequency of KET ranging from 7.5% in first-lactation to 17.2% in the third-lactation cows.

### Phenotypic Associations of BHB with KET and Milk Traits

To investigate phenotypic associations of predicted blood BHB with KET and milk production traits, cows were grouped into 2 categories based on their predicted blood BHB test for risk of ketosis: negative (BHB $<$1.2 mmol/L), or positive (BHB $\geq$1.2 mmol/L). The threshold of $\geq$1.2 mmol/L was used because most authors have used that value (Rollin et al., 2010; van Knegsel et al., 2010; van der Drift et al., 2012a). Figure 3 shows phenotypic associations of cows with positive and negative predicted blood BHB test across early lactation (11–120 DIM) with test-day milk yield, milk fat, protein, and lactose contents, and KET. Mean milk yield for cows with positive predicted blood BHB test increased up to 42 DIM and rapidly decreased afterward, whereas the increase in milk yield lasted up to 60 DIM for cows with negative predicted blood BHB test ($<$1.2 mmol/L). During the first 2 mo of lactation, mean milk yield from cows with a positive test was higher than for cows with a negative test (Figure 3a). This indicates that high-yielding cows had higher blood BHB concentration and were possibly more prone to the risk of developing ketosis in early lactation compared with lower-yielding cows. The difference in mean milk yield between cows with negative and positive test decreased for DIM up to 68 DIM and almost overlapped afterward.

In addition, cows with a positive predicted blood BHB test had a higher milk fat content throughout the early lactation stage compared with cows with a negative test (Figure 3b). However, cows with a positive predicted blood BHB test had a lower content of milk protein throughout the lactation stages considered compared with cows with a negative test (Figure 3c). Mean milk lactose content for both cows with positive and negative BHB test increased up to 30 DIM and decreased afterward. Cows with positive predicted blood BHB test had lower milk lactose content throughout the lactation stage compared with cows with a negative test (Figure 3d). As expected, cows with positive test for predicted blood BHB had higher frequency of KET (3.41%) compared with cows with negative test (1.01%) (Figure 3e). This indicated that 77.15% of cows with KET were categorized with a positive predicted blood BHB test. Koeck et al. (2014) also observed higher frequency of KET among cows testing positive ($\geq$0.20 mmol/L of milk BHB, 10.8%), followed by cows classified as suspect ($\leq$0.15–0.20 mmol/L of milk BHB, 5.4%) and negative ($<$0.15 mmol/L of milk BHB, 2.3%).

### Phenotypic Associations of KET with BHB and Milk Traits

To study phenotypic associations of KET with predicted blood BHB and milk production traits, cows in data set 3 were grouped as nontreated (healthy) or treated (diseased) based on absence or presence of veterinary treatment for ketosis, respectively. The mean test day predicted blood BHB, milk yield, and milk fat, protein, and lactose contents between 11 and 120 DIM from cows treated by veterinarian for ketosis or not (KET) are given in Figure 4. Mean predicted blood BHB of cows with KET was higher in early lactation, but decreased rapidly with stage of lactation toward the level found in more healthy cows (Figure 4e). Means of predicted blood BHB traits at different DIM intervals were also calculated for cows with and without KET (Table 3). Cows with KET had higher means ($P < 0.05$) for the predicted blood BHB in all lactation stages. The means of predicted blood BHB both at each DIM and at the different DIM intervals found in the current study were in agreement with the

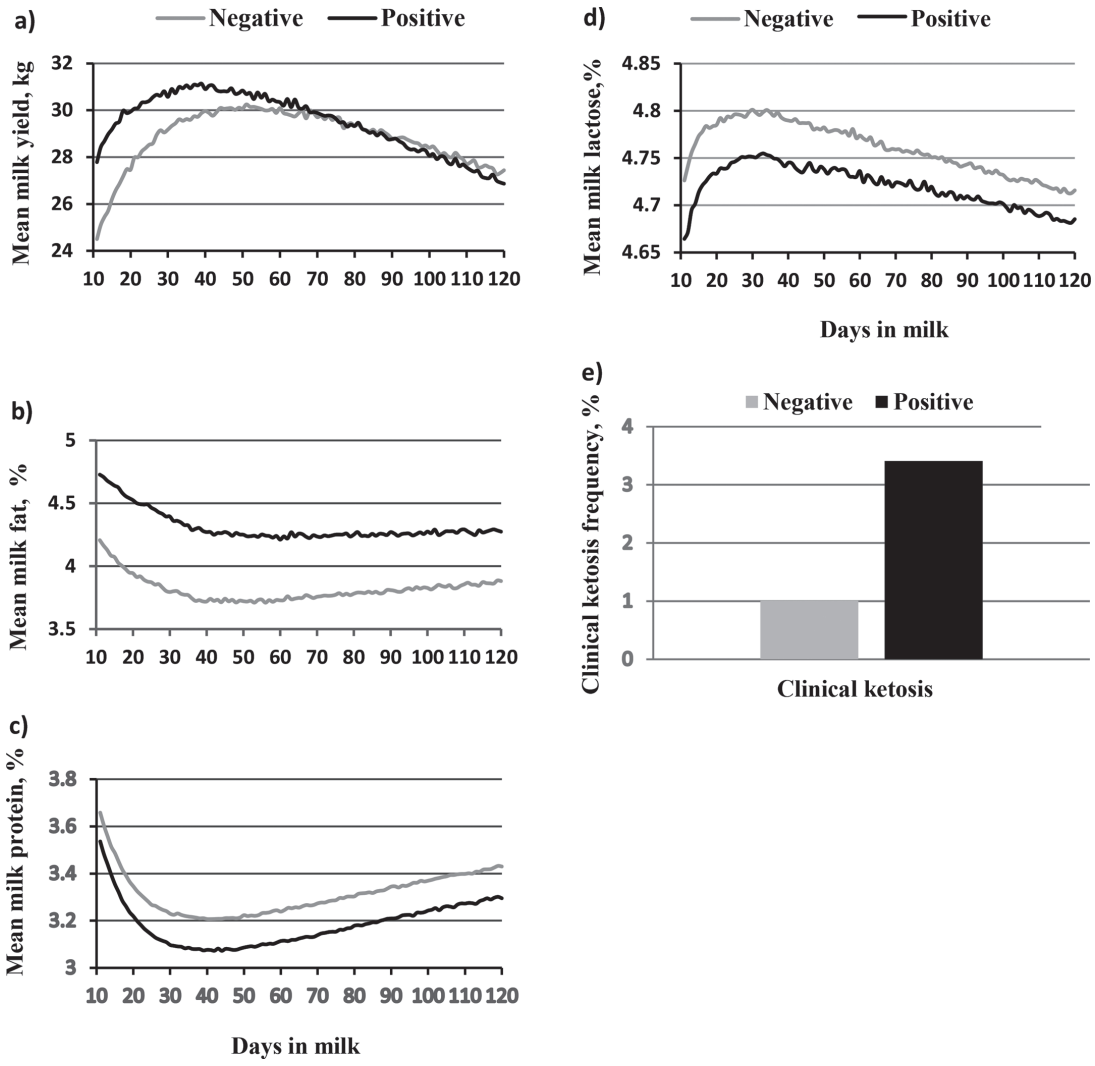



**Figure 3**. Mean (a) milk yield, (b) milk fat content, (c) protein content, (d) lactose content, and (e) frequency of clinical ketosis by DIM for cows with a negative (BHB <1.2 mmol/L) or positive (BHB ≥1.2 mmol/L) test result for risk of ketosis until 120 d after calving.

finding of Koeck et al. (2016) in Canadian dairy cows between 5 and 100 DIM.

Cows with KET had slightly higher mean milk yield in early lactation up to around 22 DIM, but slightly lower afterward up to around 60 DIM compared with healthy cows (Figure 4a). This again supported the idea that high-yielding cows are more prone to the risk of KET in early lactation compared with low-yielding cows. Contrary to the current finding, Koeck et al. (2013) reported that cows with KET had a slightly lower milk yield in early lactation compared with healthy cows in first lactation.

In our study, mean milk fat content in both cows with and without KET decreased in early lactation up to 60 DIM. Afterward, the mean milk fat content stabilized at an average value of around 4% (Figure 4b). The

mean milk fat content from cows with KET was higher in early lactation compared with healthy cows, whereas the mean milk protein content from cows with KET was slightly lower at every DIM compared with healthy cows (Figure 4c). In support of the current finding, Koeck et al. (2013) found higher mean milk fat content but a slightly lower milk protein content in cows with KET compared with healthy cows both at the first (5–30 DIM) and second test day (31–60 DIM). The mean milk lactose content from both cows with and without KET increased to about 30 DIM. Afterward, mean milk lactose content from healthy cows decreased

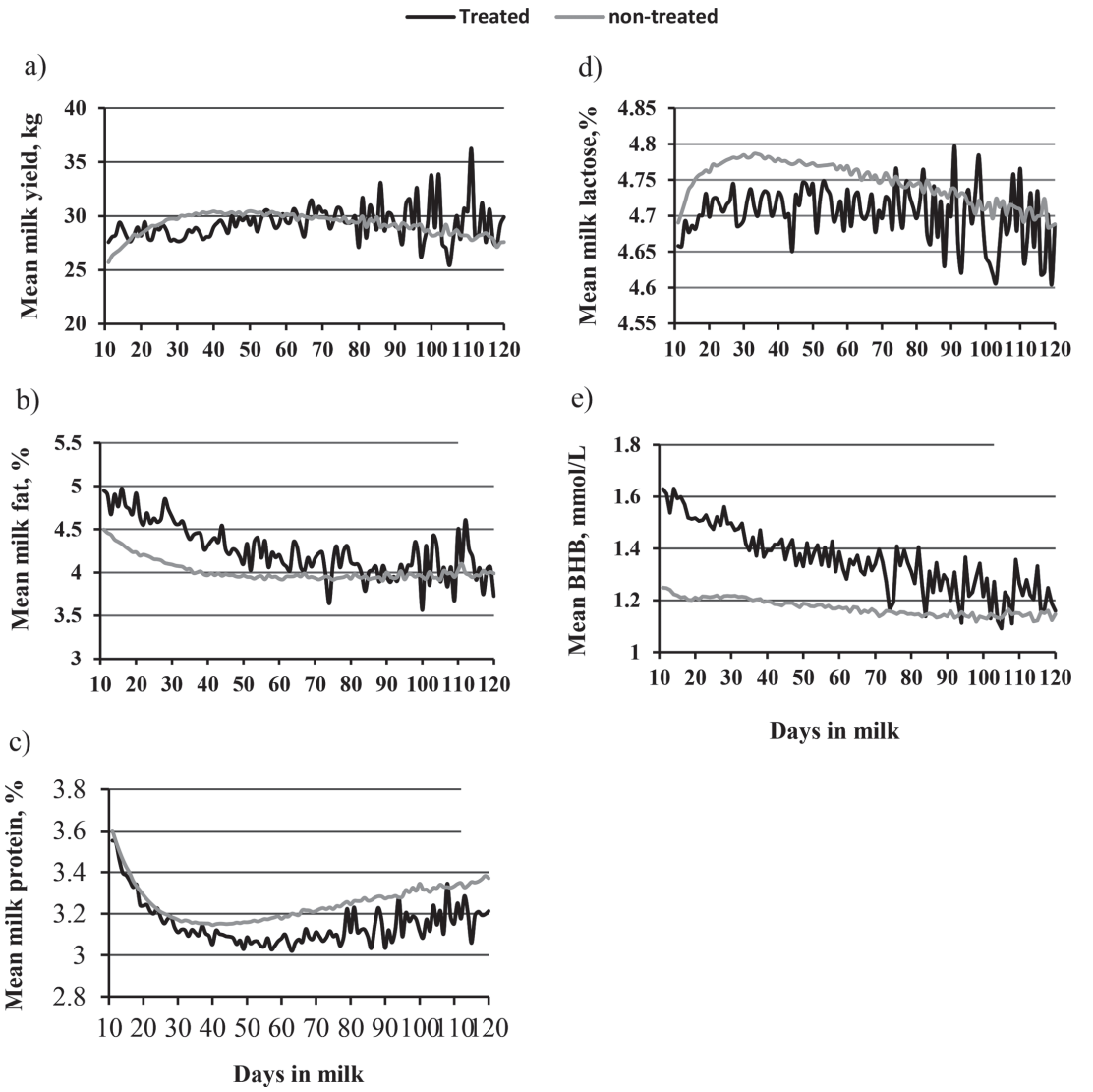



**Figure 4**. Mean (a) milk yield, (b) milk fat content, (c) milk protein content, (d) lactose content, and (e) blood BHB predicted from milk spectra by DIM for healthy cows (not treated for ketosis) and cows treated by veterinarian for clinical ketosis within the first 120 d after calving.

**Table 3**. Means (SD) of blood BHB predicted from milk spectra at 11 to 30 ($BHB_1$), 31 to 60 ($BHB_2$), 61 to 90 ($BHB_3$), and 91 to 120 DIM ($BHB_4$) and across early lactation stage (11 to 120 DIM; $BHB_{all}$) for nontreated cows or cows treated for clinical ketosis based on data set 3

| Trait | Period, DIM | Clinical ketosis | |
|---|---|---|---|
| | | Nontreated | Treated |
| $BHB_1$ | 11–30 | 1.218 (0.330)[b] | 1.540 (0.374)[a] |
| $BHB_2$ | 31–60 | 1.193 (0.291)[b] | 1.405 (0.342)[a] |
| $BHB_3$ | 61–90 | 1.155 (0.264)[b] | 1.315 (0.297)[a] |
| $BHB_4$ | 91–120 | 1.141 (0.258)[b] | 1.232 (0.295)[a] |
| $BHB_{all}$ | 11–120 | 1.195 (0.303)[b] | 1.434 (0.359)[a] |

[a,b]Values with different superscripts in the same row were significantly different at $P < 0.05$, which was tested by 2-sample $t$ test.

whereas that of treated cows stabilized at an average value of around 4.7% and slightly decreased in the last 20 DIM (Figure 4d).

Cows treated for KET showed much more variation of the means per DIM for all traits compared with nontreated cows (Figure 4). This is due to the smaller number of observations for treated cows (7,000) compared with nontreated cows (350,000). This is also evident in early lactation where, for treated cows, variation was smaller than in later lactation stages in all traits. This is because more animals were treated in early lactation. As DIM progressed the number of cows treated was reduced, resulting in larger variation in means per DIM.

### *Heritability*

Table 4 shows estimates of variance components and corresponding variance ratios of predicted blood BHB traits from univariate analyses. The estimated heritability of predicted blood BHB across lactation stages (11–120 DIM) was moderate (0.274); that of predicted blood BHB at different DIM intervals increased with lactation stage from 0.250 for $BHB_1$ to 0.365 for $BHB_4$, with standard errors ranging from 0.004 to 0.006 (Table 4). The heritability estimates of predicted blood BHB at different stages were similar in univariate and bivariate analyses [BHB at different DIM intervals with each other (Table 5) or with milk production traits]. In data set 3, however, heritability estimates of predicted blood BHB from bivariate analyses with KET were slightly lower (0.238 to 0.353), with relatively higher standard errors (0.006 to 0.008). The medium heritability of predicted blood BHB suggests that genetic selection for lower blood BHB concentrations in early lactation (<120 DIM) is feasible and may lower occurrences of ketosis. In agreement with the current study, Koeck et al. (2014) found an increasing trend in heritability for milk BHB with DIM, but Oikonomou et al. (2008) obtained a decreasing trend in heritability of measured serum BHB as DIM increased in primiparous cows. In the literature, estimates of heritabilities vary from 0.08 to 0.40 for measured blood BHB (Oikonomou et al., 2008; van der Drift et al., 2012b) and from 0.067 to 0.29 for milk BHB (van der Drift et al., 2012b; Koeck et al., 2014; Jamrozik et al., 2016). The heritability estimates of predicted blood BHB traits found in our study were generally in the range of published values, but slightly higher especially for BHB after 60 DIM. Among other factors, differences in periods of lactation (11–120 DIM vs. 5–60 or 100 DIM) and definitions of BHB traits might contribute to the differences in estimates.

Variance components and corresponding variance ratios of registered clinical KET across lactations from univariate analyses are given in Table 4. The heritability estimate of KET in univariate analysis based on data set 2 was 0.078 (Table 4), but was reduced to 0.002 when KET was analyzed bivariately either with predicted blood BHB traits (Table 7) or milk production traits using data set 3 (Table 8). As estimates of heritability for binary traits with linear models are frequency-dependent, the reduction in heritability of KET in bivariate analysis could be due to the low frequency of KET in data set 3 compared with that in data set 2. Estimates of heritability from univari-

**Table 4**. Estimates of additive genetic $\left(\sigma_a^2\right)$, permanent animal environment $\left(\sigma_{pe}^2\right)$, herd test day or herd year $\left(\sigma_{htd}^2\right)$ and residual $\left(\sigma_e^2\right)$ variances, as well as variance ratios for genetic ($h^2$), permanent environment ($c^2$), herd test day or herd year ($d^2$), and residual ($e^2$) effects for blood BHB predicted from milk spectra at 11 to 30 ($BHB_1$), 31 to 60 ($BHB_2$), 61 to 90 ($BHB_3$), 91 to 120 ($BHB_4$), and 11 to 120 DIM ($BHB_{all}$) and clinical ketosis (KET) from univariate analyses based on data set 1 for BHB and data set 2 for KET[1]

| Trait | $\sigma_a^2$ | $\sigma_{pe}^2$ | $\sigma_{htd}^2$ | $\sigma_e^2$ | $h^2$ | $c^2$ | $d^2$ | $e^2$ |
|---|---|---|---|---|---|---|---|---|
| $BHB_1$ | 0.023564 | 0.010366 | 0.015343 | 0.044974 | 0.250 | 0.110 | 0.163 | 0.477 |
| $BHB_2$ | 0.020551 | 0.007571 | 0.014075 | 0.031134 | 0.280 | 0.103 | 0.192 | 0.425 |
| $BHB_3$ | 0.020069 | 0.005571 | 0.012503 | 0.023387 | 0.326 | 0.091 | 0.203 | 0.380 |
| $BHB_4$ | 0.021047 | 0.004334 | 0.012248 | 0.020069 | 0.365 | 0.075 | 0.212 | 0.384 |
| $BHB_{all}$ | 0.019683 | 0.005257 | 0.013237 | 0.033719 | 0.274 | 0.073 | 0.184 | 0.469 |
| KET[2] | 0.001571 | 0.001017 | 0.001138 | 0.016472 | 0.078 | 0.050 | 0.056 | 0.816 |

[1]Standard error for $h^2 = 0.003$–$0.006$; for $c^2 = 0.002$–$0.006$; for $d^2 = 0.001$–$0.003$; and for $e^2 = 0.002$–$0.006$.

[2]For KET, $\sigma_{htd}^2$ and $d^2$ represent herd-year variance and the corresponding variance ratio, respectively.

**Table 5**. Heritabilities[1] (bold, on diagonal), genetic correlations (above diagonal), and phenotypic correlations (below diagonal) with SE in parentheses for BHB predicted from milk spectra at 11 to 30 ($BHB_1$), 31 to 60 ($BHB_2$), 61 to 90 ($BHB_3$), and 91 to 120 ($BHB_4$) from bivariate analyses of data set 1

| Trait | $BHB_1$ | $BHB_2$ | $BHB_3$ | $BHB_4$ |
|---|---|---|---|---|
| $BHB_1$ | **0.248 (0.005)** | 0.922 (0.007) | 0.830 (0.009) | 0.763 (0.011) |
| $BHB_2$ | 0.489 (0.003) | **0.274 (0.004)** | 0.975 (0.003) | 0.949 (0.004) |
| $BHB_3$ | 0.451 (0.003) | 0.574 (0.002) | **0.322 (0.005)** | 0.977 (0.003) |
| $BHB_4$ | 0.413 (0.003) | 0.541 (0.002) | 0.619 (0.002) | **0.360 (0.005)** |

[1]Heritability estimate and the corresponding SE in parentheses were calculated as the average estimate from all bivariate analyses containing a particular trait.

ate analysis for KET were in agreement with previous estimates from linear models (0.01 to 0.08; Jamrozik et al., 2016; Koeck et al., 2016; Pryce et al., 2016), but lower than the majority of estimates from threshold models (0.02 to 0.16; Pryce et al., 2016). Heringstad et al. (2005) obtained substantially higher estimates for KET on the same population (NRF) with threshold models, with estimates of 0.14, 0.15, and 0.16 for first, second, and third lactations, respectively. However, it has to be noted that linear model estimates of heritability for binary traits are frequency-dependent and therefore not directly comparable with estimates on the underlying liability scale (Pryce et al., 2016).

### *Proportion of Variance Attributed to PE, HTD, or HY*

The proportion of variance attributed to PE, HTD, and residual for predicted blood BHB traits are also given in Table 4. The HTD variance was lower than the additive genetic variance and the proportion of variance attributed to HTD increased as DIM progressed, from 16 to 21%. This indicates that events on test days (e.g., feeding and management) have less influence on the etiology of ketosis than genetic difference between cows. In contrast, van der Drift et al. (2012b) found a larger proportion of variance attributed to herd (not to HTD) than that of additive genetic variance for plasma BHB. Though effects of HTD were smaller than additive genetic effects, they had considerable influences (explained 16 to 21% of variation) on BHB traits. Therefore, prevention strategies for ketosis should include both feeding and management strategies at dairy farms, and genetic improvement through breeding programs, which was also concluded by van der Drift et al. (2012b). Similar to HTD effects, the proportion of variance attributed to PE was smaller than that of additive genetic variance, and it decreased as DIM progressed, from 11 to 7.5% of the total variance. In the study of Jamrozik et al. (2016), the permanent environmental effect captured 25% of the total variance in milk BHB in later lactations, which is more than 2 times the estimates found in the current study.

For KET, the proportion of variance attributed to HY, PE, and residual are presented in Table 4. The proportion of variance attributed to HY was slightly lower than that of additive genetic variance (5.6 vs. 7.8%) for univariate analysis. However, in bivariate analysis, the proportion of variance attributed to HY was much larger (87.4%) than that of additive genetic variance (results not shown in table). This result suggests that environmental factors (feeding and management) have more influence on the prevalence of KET in HY than the genetic difference between cows. Estimates of the HY effect obtained in the current study were higher than estimates reported for HY effects in Canadian Holsteins (Jamrozik et al., 2016). For KET and its indicator (predicted blood BHB), estimates of residual effects were lower than values reported in literature (Jamrozik et al., 2016), indicating that the models used in the current study fit well.

### *Genetic and Phenotypic Correlations Among BHB Traits*

Estimates of genetic and phenotypic correlations among predicted blood BHB traits are given in Table 5. The genetic correlations between the predicted blood BHB traits were higher between adjacent DIM intervals (0.92–0.98) and decreased as intervals were further apart (down to 0.76). This is in agreement with results found by Koeck et al. (2014) for milk BHB traits defined in a 20-d interval for cows between 5 and 60 DIM. Similar to the genetic correlations, the estimates of phenotypic correlations were higher between adjacent DIM intervals and in line with Koeck et al. (2014). Genetic correlations were strongest between $BHB_2$ and $BHB_3$ as well as $BHB_3$ and $BHB_4$, whereas phenotypic correlation was the strongest between $BHB_3$ and $BHB_4$ (0.62).

### *Genetic and Phenotypic Correlations Among BHB, Milk Traits, and KET*

Genetic correlations of predicted blood BHB traits with milk production traits from the same lactation

**Table 6.** Estimates of heritability[1] ($h^2$) and genetic ($r_g$) and phenotypic ($r_p$) correlations (SE) of blood BHB predicted from milk spectra at 11 to 30 ($BHB_1$), 31 to 60 ($BHB_2$), 61 to 90 ($BHB_3$), and 91 to 120 ($BHB_4$) with milk production traits[2] from the same lactation stage as BHB based on data set 1 in bivariate analyses

| | 11–30 DIM | | | 31–60 DIM | | | 61–90 DIM | | | 91–120 DIM | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Trait[2] | $h^2$ | $r_g$ | $r_p$ | $h^2$ | $r_g$ | $r_p$ | $h^2$ | $r_g$ | $r_p$ | $h^2$ | $r_g$ | $r_p$ |
| $BHB_{1\text{-}4}$ | 0.249 | | | 0.285 | | | 0.329 | | | 0.363 | | |
| | (0.006) | | | (0.005) | | | (0.005) | | | (0.005) | | |
| Milk, kg | 0.164 | 0.188 | 0.108 | 0.190 | 0.086 | 0.042 | 0.214 | 0.049 | 0.020 | 0.229 | 0.063 | −0.001 |
| | (0.006) | (0.023) | (0.003) | (0.005) | (0.017) | (0.002) | (0.005) | (0.016) | (0.003) | (0.006) | (0.015) | (0.003) |
| Fat, kg | 0.106 | 0.273 | 0.405 | 0.105 | 0.110 | 0.347 | 0.108 | 0.037 | 0.298 | 0.127 | 0.045 | 0.25 |
| | (0.005) | (0.024) | (0.002) | (0.004) | (0.020) | (0.002) | (0.004) | (0.020) | (0.002) | (0.005) | (0.019) | (0.002) |
| Prot, kg | 0.131 | 0.108 | 0.013 | 0.160 | −0.067 | −0.083 | 0.177 | −0.129 | −0.111 | 0.185 | −0.106 | −0.122 |
| | (0.006) | (0.025) | (0.003) | (0.005) | (0.018) | (0.002) | (0.005) | (0.017) | (0.003) | (0.005) | (0.016) | (0.003) |
| Lact, kg | 0.141 | 0.125 | 0.066 | 0.174 | 0.032 | 0.008 | 0.202 | 0.002 | −0.013 | 0.221 | 0.019 | −0.033 |
| | (0.006) | (0.024) | (0.003) | (0.005) | (0.018) | (0.002) | (0.005) | (0.016) | (0.003) | (0.006) | (0.016) | 0.003 |
| Fat, % | 0.096 | 0.168 | 0.445 | 0.105 | 0.062 | 0.421 | 0.122 | −0.006 | 0.380 | 0.167 | −0.030 | 0.351 |
| | (0.005) | (0.026) | (0.002) | (0.004) | (0.020) | (0.002) | (0.004) | (0.019) | (0.002) | (0.005) | (0.017) | (0.002) |
| Prot, % | 0.272 | −0.233 | −0.255 | 0.383 | −0.272 | −0.307 | 0.441 | −0.280 | −0.311 | 0.318 | −0.278 | −0.297 |
| | (0.007) | (0.018) | (0.003) | (0.005) | (0.012) | (0.002) | (0.005) | (0.011) | (0.002) | (0.002) | (0.004) | (0.003) |
| Lact, % | 0.406 | −0.234 | −0.200 | 0.451 | −0.172 | −0.167 | 0.458 | −0.159 | −0.163 | 0.455 | −0.154 | −0.167 |
| | (0.007) | (0.016) | (0.003) | (0.005) | (0.012) | (0.002) | (0.005) | (0.011) | (0.002) | (0.006) | (0.011) | (0.003) |

[1]Heritability estimates for blood BHB traits and the corresponding standard error in parentheses were calculated as the average estimate from all bivariate analyses containing a particular trait.

[2]Prot = protein yield (kg) or protein content (%); Lact = lactose yield (kg) or lactose content (%).

stage as BHB (Table 6) were slightly different from the correlation estimates with milk production traits across lactation stages (Table 7). For example, fat yield and fat and lactose contents at 11 to 30 DIM and protein content at 61 to 90 and 91 to 120 DIM from the same lactation stage as BHB had slightly stronger genetic correlations with BHB than corresponding estimates for those milk traits across lactation stage with BHB traits. Predicted blood BHB at the first test-day ($BHB_1$) was genetically moderately correlated with milk production traits both from the same lactation stage (Table 6) and across lactation stages (Table 7), except for milk fat content that was weakly correlated (Table 7). All predicted blood BHB traits were moderately negatively correlated with milk protein and lactose contents (Tables 6 and 7). The remaining milk traits showed negligible genetic correlation with predicted blood BHB in later lactation stages (31–120 DIM), except for milk and fat yields with $BHB_2$ and protein yield with $BHB_3$ and $BHB_4$ (Tables 6 and 7). Similar to genetic correlations, phenotypic correlations of BHB traits with milk production traits decreased as DIM progressed. Predicted blood BHB traits had moderate negative phenotypic correlations with milk protein and lactose contents from the same lactation stage (Table 6) and across lactation stages (Table 7). Unlike the genetic correlations, phenotypic correlations of BHB traits with milk traits from the same lactation stage as BHB were stronger than estimates across lactation stages. Predicted blood BHB traits were more strongly phenotypically correlated with fat content from the same lactation stage (0.351–0.445; Table 6) than across lactation stages (0.175–0.403; Table 7).

To the best of our knowledge, estimates of genetic and environmental correlations between milk or blood BHB and milk production traits are not available in the literature, except the work of Koeck et al. (2014), who reported correlations between EBV for milk BHB traits and routinely evaluated traits including yields of milk, fat, and protein. In agreement with the current study, Koeck et al. (2014) found moderate correlations (0.13–0.22) between EBV of milk BHB and milk yields, but found insignificant ($P > 0.05$) correlations between EBV of milk BHB and yields of fat and protein. The present results indicate higher genetic merits for milk, fat, protein, and lactose yields associated with higher BHB in early lactation and, therefore, a greater susceptibility to risk of ketosis. Phenotypic associations between predicted blood BHB and milk traits are given in Figure 3. They show, for example, that cows that had high milk yield in early lactation (up to 60 DIM) had positive test for risk of ketosis (high BHB values; Figure 3a). The opposite was true for cows with high milk protein content throughout all lactation stages considered

**Table 7**. Estimates of heritability[1] ($h^2$), genetic correlations ($r_g$), and phenotypic correlations ($r_p$) of blood BHB predicted from milk spectra at 11 to 30 ($BHB_1$), 31 to 60 ($BHB_2$), 61 to 90 ($BHB_3$), 91 to 120 ($BHB_4$), and 11 to 120 DIM ($BHB_{all}$) with clinical ketosis (KET) and milk production traits across lactation stage from bivariate analyses

| Parameter | Trait | $h^2$ | $BHB_1$ | $BHB_2$ | $BHB_3$ | $BHB_4$ | $BHB_{all}$ |
|---|---|---|---|---|---|---|---|
| $r_g$ | KET | 0.002 (0.0002) | 0.469 (0.050) | 0.310 (0.049) | 0.192 (0.049) | 0.179 (0.048) | 0.288 (0.043) |
| | Milk, kg | 0.195 (0.003) | 0.277 (0.016) | 0.118 (0.014) | 0.024 (0.013) | −0.031 (0.013) | 0.101 (0.010) |
| | Fat, kg | 0.107 (0.002) | 0.248 (0.016) | 0.126 (0.015) | 0.069 (0.015) | −0.011 (0.014) | 0.106 (0.011) |
| | Protein, kg | 0.165 (0.003) | 0.107 (0.017) | −0.047 (0.014) | −0.124 (0.014) | −0.166 (0.013) | −0.052 (0.011) |
| | Lactose, kg | 0.181 (0.003) | 0.229 (0.016) | 0.063 (0.013) | −0.028 (0.013) | −0.081 (0.016) | 0.051 (0.010) |
| | Fat, % | 0.116 (0.002) | 0.033 (0.016) | 0.054 (0.014) | 0.084 (0.014) | 0.042 (0.013) | 0.035 (0.011) |
| | Protein, % | 0.371 (0.003) | −0.367 (0.011) | −0.308 (0.009) | −0.252 (0.009) | −0.215 (0.009) | −0.288 (0.007) |
| | Lactose, % | 0.428 (0.003) | −0.189 (0.012) | −0.178 (0.009) | −0.173 (0.009) | −0.163 (0.009) | −0.174 (0.007) |
| $r_p$ | KET | | 0.065 | 0.047 | 0.041 | 0.023 | 0.047 |
| | Milk, kg | | 0.135 (0.002) | 0.033 (0.002) | −0.013 (0.002) | −0.038 (0.002) | 0.042 (0.002) |
| | Fat, kg | | 0.382 (0.002) | 0.231 (0.002) | 0.143 (0.002) | 0.102 (0.002) | 0.327 (0.001) |
| | Protein, kg | | 0.024 (0.003) | −0.065 (0.002) | −0.091 (0.002) | −0.107 (0.002) | −0.073 (0.001) |
| | Lactose, kg | | 0.096 (0.003) | 0.005 (0.002) | −0.038 (0.002) | −0.062 (0.002) | 0.009 (0.002) |
| | Fat, % | | 0.403 (0.002) | 0.281 (0.002) | 0.202 (0.002) | 0.175 (0.002) | 0.403 (0.001) |
| | Protein, % | | −0.285 (0.002) | −0.238 (0.002) | −0.186 (0.002) | −0.165 (0.002) | −0.292 (0.001) |
| | Lactose, % | | −0.188 (0.003) | −0.137 (0.002) | −0.126 (0.002) | −0.123 (0.002) | −0.173 (0.002) |

[1]Heritability estimate for ketosis and milk traits and the corresponding standard errors in parentheses were calculated as the average estimate from all bivariate analyses containing a particular trait.

(Figure 3c). The genetic associations observed between yield traits and BHB concentrations were expected, as genetic selection for high milk production would result in larger negative energy balance and require larger fat mobilization in early lactation (Veerkamp et al., 2003; Coffey et al., 2004), and hence a higher risk of ketosis.

Estimates of genetic correlations of BHB traits with KET are also presented in Table 7. All BHB traits were genetically moderately correlated with KET, with correlations between 0.18 and 0.47. Genetic relationship between KET and its indicator (predicted blood BHB) was reduced as DIM progressed. Estimates of genetic correlations between KET and its indicator at 11 to 30 and 31 to 60 DIM were in the range of that reported in literature (Koeck et al., 2014; Koeck, 2015; Jamrozik et al., 2016). Koeck et al. (2014) found an estimate of genetic correlation between KET and milk BHB at 5 to 40 DIM of 0.48, which is comparable with the 0.47 estimate observed between KET and blood BHB at 11 to 30 DIM in the current study. Jamrozik et al. (2016) found a genetic correlation of KET with milk BHB, with correlations of 0.63 and 0.37 for first and later lactations, respectively, and with a correlation of 0.25 for both KET and milk BHB in the later lactations. However, Koeck (2015) and Koeck et al. (2016) reported stronger genetic correlation between KET and milk BHB at 5 to 40 DIM, with correlations ranging from 0.70 to 0.75. Phenotypic correlations were low between KET and blood BHB traits, ranging from 0.023 (between KET and $BHB_4$) to 0.065 (between KET and $BHB_1$; Table 7), which is much lower than estimates reported for phenotypic correlations between KET and milk BHB at 5 to 40 DIM (0.150 to 0.186; Koeck et al., 2014, 2016).

**Table 8**. Estimates of heritability[1] ($h^2$), genetic ($r_g$), and phenotypic ($r_p$) correlations of clinical ketosis (KET) with milk production traits from bivariate analysis based on data set 3

| | | KET | |
|---|---|---|---|
| Trait | $h^2$ | $r_g$ | $r_p$ |
| KET | 0.002 (0.0002) | | |
| Milk, kg | 0.183 (0.003) | 0.178 (0.047) | −0.042 |
| Fat, kg | 0.103 (0.003) | 0.157 (0.046) | −0.010 |
| Protein, kg | 0.156 (0.004) | 0.002 (0.048) | −0.050 |
| Lactose, kg | 0.169 (0.004) | 0.161 (0.048) | −0.046 |
| Fat, % | 0.114 (0.003) | −0.023 (0.049) | 0.033 |
| Protein, % | 0.361 (0.003) | −0.333 (0.039) | −0.017 |
| Lactose, % | 0.412 (0.003) | −0.043 (0.037) | −0.025 |

[1]Heritability estimate for ketosis and the corresponding standard error in parentheses were calculated as the average estimate from all bivariate analyses containing a particular trait.

Table 8 presents the estimates of genetic and phenotypic correlations between KET and milk productions. Ketosis was significantly positively correlated with milk (0.178), fat (0.157), and lactose (0.161) yields, but not genetically correlated with protein yield (0.002). Previous studies on genetic correlations between KET and milk production traits in early lactation are scarce. Koeck et al. (2013) found that genetic correlations between milk yield at 5 to 30 and 31 to 60 DIM with KET were not different from zero. In their comprehensive review, Pryce et al. (2016) indicated that mean genetic correlation between 305-d milk yield and KET was unfavorable (positive), which is in line with the current study. In the same review, negative mean genetic correlations between KET and 305-d fat and protein yields

were reported (Pryce et al., 2016). In the current study, none of the genetic correlations of KET with milk fat (−0.023) or lactose (−0.043) content were significantly different from zero, but with milk protein content it was (−0.333). Contrary to the current finding, Koeck et al. (2013) reported a near zero genetic correlation of KET with milk protein content (−0.06 at 5–30 DIM and −0.09 at 31–60 DIM), but a medium positive genetic correlation with milk fat content at 5 to 30 DIM (0.33). The genetic relationships between KET and milk production traits observed in the current study indicated that a higher milk, fat, and lactose yields and a lower milk protein content would be associated with an increased risk of developing ketosis.

Phenotypic correlations between KET and milk production traits were low and negative (except for fat content that had a low positive genetic correlation with KET). This is in agreement with the phenotypic correlations of KET with milk yield and fat and protein contents reported by Koeck et al. (2013).

## CONCLUSIONS

Blood BHB predicted from milk spectra at different DIM intervals or across lactation stages is heritable, with heritability estimates ranging from 0.250 to 0.365. It seems sufficient to consider only $BHB_1$ and $BHB_2$ as indicator traits in a routine genetic evaluation for resistance to ketosis because $BHB_1$ and $BHB_2$ are available early in lactation and have higher genetic correlations with KET than later BHB traits ($BHB_3$ and $BHB_4$). Generally, predicted blood BHB was genetically moderately correlated with KET and milk production traits, and those correlations decreased as DIM progressed. Similarly, KET had moderate genetic correlations with milk, fat, and lactose yields and protein content. The moderate genetic correlations observed between BHB traits and KET indicate that selective breeding for lower BHB may contribute to lower susceptibility of cows to ketosis in early lactation. A lower BHB was genetically associated with higher milk protein and lactose contents, but with lower milk, fat, lactose, and protein yields. Blood BHB predicted from milk spectra can be routinely obtained from test-day milk samples and provides a practical alternative for breeding cows to have lower susceptibility to ketosis, even though correlations with KET are moderate. Before commencing genetic selection for a lower BHB in NRF dairy cattle, further studies are needed on genetic associations of BHB with health and fertility traits. The benefit of using FT-MIR predicted indicator trait (e.g., BHB) in addition to the directly observed ketosis also has to be studied.

## ACKNOWLEDGMENTS

The authors gratefully acknowledge TINE SA (Ås, Norway), the Norwegian dairy association that manage the central database for providing data for this work. The authors also thank Polish Federation of Cattle Breeders and Dairy Farmers (Warsaw, Poland) for allowing us to use the calibration model developed based on blood BHB and milk spectra from Polish cows.

## REFERENCES

Belay, T. K., B. S. Dagnachew, Z. M. Kowalski, and T. Ådnøy. 2017. An attempt at predicting blood β-hydroxybutyrate from Fourier-transform mid-infrared spectra of milk using multivariate mixed models in Polish dairy cattle. J. Dairy Sci. 100:XXXX–XXXX. https://doi.org/10.3168/jds.2016-12252.

Coffey, M. P., G. Simm, J. Oldham, W. Hill, and S. Brotherstone. 2004. Genotype and diet effects on energy balance in the first three lactations of dairy cows. J. Dairy Sci. 87:4318–4326.

de Roos, A., H. Van Den Bijgaart, J. Hørlyk, and G. De Jong. 2007. Screening for subclinical ketosis in dairy cattle by Fourier transform infrared spectrometry. J. Dairy Sci. 90:1761–1766.

Denis-Robichaud, J., J. Dubuc, D. Lefebvre, and L. DesCôteaux. 2014. Accuracy of milk ketone bodies from flow-injection analysis for the diagnosis of hyperketonemia in dairy cows. J. Dairy Sci. 97:3364–3370.

Duffield, T. F., K. Lissemore, B. McBride, and K. Leslie. 2009. Impact of hyperketonemia in early lactation dairy cows on health and production. J J. Dairy Sci. 92:571–580.

Grelet, C., C. Bastin, M. Gelé, J.-B. Davière, M. Johan, A. Werner, R. Reding, J. F. Pierna, F. Colinet, and P. Dardenne. 2016. Development of Fourier transform mid-infrared calibrations to predict acetone, β-hydroxybutyrate, and citrate contents in bovine milk through a European dairy network. J. Dairy Sci. 99:4816–4825.

Heringstad, B., Y. Chang, D. Gianola, and G. Klemetsdal. 2005. Genetic analysis of clinical mastitis, milk fever, ketosis, and retained placenta in three lactations of Norwegian red cows. J. Dairy Sci. 88:3273–3281.

Heringstad, B., G. Klemetsdal, and J. Ruane. 2000. Selection for mastitis resistance in dairy cattle: A review with focus on the situation in the Nordic countries. Livest. Prod. Sci. 64:95–106.

Jamrozik, J., A. Koeck, G. Kistemaker, and F. Miglior. 2016. Multiple-trait estimates of genetic parameters for metabolic disease traits, fertility disorders, and their predictors in Canadian Holsteins. J. Dairy Sci. 99:1990–1998.

Koeck, A. 2015. Development of genetic evaluations for metabolic disease traits for Canadian dairy cattle. Interbull Bull. 49:76–79.

Koeck, A., J. Jamrozik, G. Kistemaker, F. Schenkel, R. Moore, D. Lefebvre, D. Kelton, and F. Miglior. 2016. Genetic and phenotypic associations of milk β-hydroxybutyrate with ketosis in Canadian Holsteins. Can. J. Anim. Sci. 96:302–305.

Koeck, A., J. Jamrozik, F. Schenkel, R. Moore, D. Lefebvre, D. Kelton, and F. Miglior. 2014. Genetic analysis of milk β-hydroxybutyrate and its association with fat-to-protein ratio, body condition score, clinical ketosis, and displaced abomasum in early first lactation of Canadian Holsteins. J. Dairy Sci. 97:7286–7292.

Koeck, A., F. Miglior, J. Jamrozik, D. Kelton, and F. Schenkel. 2013. Genetic associations of ketosis and displaced abomasum with milk production traits in early first lactation of Canadian Holsteins. J. Dairy Sci. 96:4688–4696.

Madsen, P., and J. Jensen. 2008. DMU: A user's guide. A package for analysing multivariate mixed models. Version 6, release 4.7. DJF, Foulum, Denmark.

Meyer, K. 2007. WOMBAT—A tool for mixed model analyses in quantitative genetics by restricted maximum likelihood (REML). J. Zhejiang Univ. Sci. B 8:815–821.

Negussie, E., I. Strandén, and E. A. Mäntysaari. 2008. Genetic analysis of liability to clinical mastitis, with somatic cell score and production traits using bivariate threshold–linear and linear–linear models. Livest. Sci. 117:52–59.

Oetzel, G. R. 2007. Herd-level ketosis–diagnosis and risk factors. Pages 67–91 in Proceedings of the 40th Annual Conference of American Association of Bovine Practitioners, September 19, 2007, Vancouver, Canada. American Association of Bovine Practitioners, Auburn, AL.

Oikonomou, G., G. Valergakis, G. Arsenos, N. Roubies, and G. Banos. 2008. Genetic profile of body energy and blood metabolic traits across lactation in primiparous Holstein cows. J. Dairy Sci. 91:2814–2822.

Østerås, O., H. Solbu, A. Refsdal, T. Roalkvam, O. Filseth, and A. Minsaas. 2007. Results and evaluation of thirty years of health recordings in the Norwegian dairy cattle population. J. Dairy Sci. 90:4483–4497.

Pryce, J. E., K. P. Gaddis, A. Koeck, C. Bastin, M. Abdelsayed, N. Gengler, F. Miglior, B. Heringstad, C. Egger-Danner, and K. Stock. 2016. Invited review: Opportunities for genetic improvement of metabolic diseases. J. Dairy Sci. 99:6855–6873.

Rollin, E., R. Berghaus, P. Rapnicki, S. Godden, and M. Overton. 2010. The effect of injectable butaphosphan and cyanocobalamin on postpartum serum β-hydroxybutyrate, calcium, and phosphorus concentrations in dairy cattle. J. Dairy Sci. 93:978–987.

Schwarzenbacher, H., W. Obritzhauser, B. Fuerst-Waltl, A. Koeck, and C. Egger-Danner. 2010. Health monitoring system in Austrian dual purpose Fleckvieh cattle: incidences and prevalences. Page 145 in Proc. Book of Abstracts of the 61th Ann. Meeting of the EAAP, Heraklion, Greece. Wageningen Academic Publishers, Wageningen, the Netherlands.

van der Drift, S., R. Jorritsma, J. Schonewille, H. Knijn, and J. Stegeman. 2012a. Routine detection of hyperketonemia in dairy cows using Fourier transform infrared spectroscopy analysis of β-hydroxybutyrate and acetone in milk in combination with test-day information. J. Dairy Sci. 95:4886–4898.

van der Drift, S., K. Van Hulzen, T. Teweldemedhn, R. Jorritsma, M. Nielen, and H. Heuven. 2012b. Genetic and nongenetic variation in plasma and milk β-hydroxybutyrate and milk acetone concentrations of early-lactation dairy cows. J. Dairy Sci. 95:6781–6787.

van Knegsel, A., S. van der Drift, M. Horneman, A. de Roos, B. Kemp, and E. Graat. 2010. Short communication: Ketone body concentration in milk determined by Fourier transform infrared spectroscopy: Value for the detection of hyperketonemia in dairy cows. J. Dairy Sci. 93:3065–3069.

Veerkamp, R., B. Beerda, and T. Van der Lende. 2003. Effects of genetic selection for milk yield on energy balance, levels of hormones, and metabolites in lactating cattle, and possible links to reduced fertility. Livest. Prod. Sci. 83:257–275.

# Paper IV

**Prediction accuracy of direct and indirect approaches, and their relationships with prediction ability of calibration models**

T.K. Belay, B.S. Dagnachew, S.A. Boison and T. Ådnøy

*Submitted to Journal of Dairy Science*

**Prediction accuracy of direct and indirect approaches, and their relationships with prediction ability of calibration models**

**T.K. Belay,*[1] B.S. Dagnachew,* S.A. Boison, † T. Ådnøy***

*Department of Animal and Aquacultural Sciences, Norwegian University of Life Sciences, P. O. Box 5003, 1432 Ås, Norway

† Nofima: The Norwegian Institute of Food, Fisheries and Aquaculture Research, Osloveien 1, 1430 Ås, Norway

**ABSTRACT**

Milk infrared spectra are routinely used for phenotyping traits of interest through links developed between the traits and spectra. Predicted individual traits are then used in genetic analyses for breeding value estimations (EBV) or for phenotypic predictions using a single-trait mixed model. This approach is referred to as indirect prediction (IP). An alternative approach (direct prediction – DP) is a direct genetic analysis of (a reduced dimension of) the spectra using a multi-trait model to predict multivariate EBV of the spectral components and ultimately also to predict the univariate EBV or phenotype for the traits of interest. In this study, we simulated three traits under different genetic (low: 0.10 to high: 0.90) and residual (zero to high: ± 0.90) correlation scenarios between the traits and assumed the first trait is a linear combination of the other two traits. The aim was to compare the IP and DP approaches for predictions of EBV and phenotypes under the different correlation scenarios. We also evaluated relationships between performances of the two approaches and the accuracy of calibration equations. Moreover, effect of using different regression coefficients estimated from simulated phenotypes ($\beta_p$), true breeding values ($\beta_g$) and residuals ($\beta_r$) on performance of the two approaches were evaluated. The simulated data contained

[1] Corresponding author's: name, Tesfaye Kebede Belay; e-mail: tesfaye.kebede.belay@nmbu.no

2,100 parents (100 sires and 2000 'cows') and 8,000 offspring (four offspring per 'cow'). Of the 8,000 observations, 2,000 were randomly selected and used to develop links between the first and the other two traits using partial least square (**PLS**) regression analysis. The different PLS regression coefficients such as $\beta_p$, $\beta_g$ and $\beta_r$ were used in subsequent predictions following the IP and DP approaches. BLUP analyses were done on the remaining 6,000 observations using the 'true' (co)variance components that had been used for the simulation. Accuracy of prediction (of EBV and phenotype) was calculated as a correlation between predicted and true values from the simulations. The result showed that accuracies of EBV prediction were higher in the DP than in the IP approach. The reverse was true for accuracy of phenotypic prediction ($\beta_p$), but not when using $\beta_g$ and $\beta_r$, where accuracy of phenotypic prediction in the DP was slightly higher than in the IP approach. Within the DP approach, accuracies of EBV when using $\beta_g$ were higher than when using $\beta_p$, especially at the low genetic correlation scenario. However, there were no differences in EBV prediction accuracy between the $\beta_p$ and $\beta_g$ in the IP approach. Accuracy of the calibration models increased with increase in genetic and residual correlations between the traits. Performance of both approaches increased with increase in accuracy of the calibration models. In conclusion, the DP approach is a good strategy for EBV prediction, but not for phenotypic prediction, where the classical PLS regression based equations or the IP approach provided better results.

**Key words:** indirect prediction, direct prediction, breeding value, phenotype

## INTRODUCTION

Fourier transform mid-infrared (**FT-MIR**) spectrometry is a potential tool for collection of data at population level for phenotypic and genetic analyses of milk components (or other derived traits). An individuals' phenotype for a trait is predicted from the FT-MIR spectra. This prediction is dependent on availability of links between the trait of interest and milk spectra. The predicted trait together with pedigree information and variance component estimates are used to calculate breeding values (**EBV**) and other random components included in the model based on a single-trait best linear unbiased prediction (**BLUP**) approach. Dagnachew et al. (2013b) referred to such an approach as indirect prediction (**IP**) because the multi-trait spectral information is not directly used in EBV prediction procedures. Alternatively, genetic analyses can be applied directly on the milk spectral variables or on their factor scores (latent traits). BLUP predictions of the random components of the model (EBV, herd-test*day, permanent environment and residual) for the traits of interest are then predicted as correlated traits to the corresponding random components of the spectra. Dagnachew et al. (2013b) referred to such an approach as direct prediction (**DP**). Given the strong correlations among milk FT-MIR spectral variables (Soyeurt et al., 2010;Dagnachew et al., 2013a), direct genetic analyses on such correlated spectral variables may result in better accuracy of genetic evaluations (Dagnachew et al., 2013b).

The IP and DP approaches have been used to predict EBV for major milk contents (fat, protein and lactose) in goats (Dagnachew et al., 2013b) and for traits related to fine milk compositions and technological properties of milk in cows (Bonfatti et al., 2017a). The former authors showed that the DP approach performed better than the IP approach i.e. relative genetic gain was improved by 3-5% in the DP compared to the IP approach. Dagnachew et al. (2013b) also reported high rank correlation coefficients (0.93 to 0.96) between EBV predicted using the IP and the DP. However,

Bonfatti et al. (2017) reported rank correlations ranging from 0.07 to 0.96, but with < 0.5 rank correlations for most traits investigated in that study. Belay et al. (2017) adopted the two approaches to predict phenotype for β-hydroxybutyrate in blood from milk spectra and reported a slightly better phenotypic prediction by the IP than the DP approach.

Based on studies done so far, it is difficult to make a conclusive remark on whether the DP approach is better than the IP approach for EBV or phenotype prediction. Each of the studies cited in the preceding paragraph has their limitation. For example, independent chemical analyses (reference values) for the milk contents were not available in the study of Dagnachew et al. (2013b) i.e. the study used phenotypes predicted from the same spectra as reference values for both model calibration and evaluation. Possibly for the reason above, the coefficients of determination (**$R^2$**) were very high (> 0.96). Moreover, the accuracies of EBV were estimated based on coefficient matrices of the mixed model equations in that study. In the study of Bonfatti et al. (2017a) reference values measured independently of the spectra were used to develop prediction equations that had medium (0.35) to high (0.86) $R^2$ values. However, it is difficult to distinguish the approaches that performed better based on that study because the IP and DP approaches were evaluated based on rank correlations. In an attempt to predict phenotypes with IP and DP (Belay et al., 2017), the $R^2$ were low, and datasets used for the model validation as well as for evaluation of the two approaches were small.

Furthermore, in the three studies, covariance components of the latent traits estimated by the DP approach were converted to variance components to be used in the IP approach using links (regression coefficients) estimated based on phenotypes ($\beta_p$). Similarly, EBV of the latent traits were converted into single-trait EBV using phenotype based links. Utilization of a partial least square (**PLS**) regression coefficient estimated from phenotypes ($\beta_p$), to convert EBV of latent traits

into EBV of trait of interest does not seem appropriate. This might have an impact on the performance of the approaches. Parameters estimated at one level (e.g. at phenotypic level) were used at another level (e.g. at genetic level). Therefore, the effect of using appropriate conversion parameters (e.g., $\beta_g$: estimated from true breeding values) to convert multi-trait structures to single-trait structures on performances of the two approaches is unknown and needs to be studied. Moreover, relationship between performance of the two approaches and accuracy of calibration models is unclear.

Therefore, objectives of this study were: 1) to evaluate performance of the IP and DP approaches for prediction of EBV and phenotype under different genetic and residual structures between traits; 2) to evaluate effect of using different PLS regression coefficients (e.g. $\beta_p$, $\beta_g$ etc.) for converting covariance components or EBV of latent traits into univariate structure on performance of the two approaches; and 3) to study relationship between performance of the two approaches and accuracy of calibration models.

## MATERIALS AND METHODS

### *Simulation*

A simulation program written in R (R Core Team, 2016) to make single- and multi-trait datasets based on pedigree was used. The R codes used for the simulation can be found here: https://github.com/soloboan/Multi-trait_simulations. A base population consisting of 100 sires and 2,000 cows with three traits under different genetic and residual correlation scenarios were simulated. Then, subsequently, two generations of data were simulated, with 2,100 parents (100 sires and 2,000 'cows') in each generation. It was assumed that a 'cow' would have four offspring per generation, resulting in 8,000 offspring per generation, from which parents for the next

generation were selected. Sex ratio of offspring was fixed at 50%. Parents were randomly selected, and the selected animals were randomly mated by random union of gametes leading to pseudo-overlapping generations as is mostly used in cattle breeding.

Variance components and the corresponding heritabilities used for simulation of the three traits are given in Table 1, while the different genetic and residual correlation scenarios are presented in Table 2. The first trait (hereafter referred to as the focal trait) was assumed to be milk protein content, and is a linear combination of the two other traits, which were assumed to be latent traits earlier derived from spectral variables. Mean, genetic variance and heritability of the focal trait were taken from a previous study (Belay et al., unpublished data) while those of the other two traits were based on estimates already reported (Belay et al., 2015). Residual variances for the three traits were calculated from the genetic variance and heritability of each trait. Twelve possible pairs of genetic (3 scenarios) and residual (4 scenarios) correlation scenarios were simulated. The genetic correlations between the three traits were assumed to be either low (0.10 to 0.25), medium (0.50 to 0.70) or high (0.80 to 0.90) while the residual correlations were grouped as zero, low, medium and high. Simulations were replicated 100 times for each scenario. The simulated data consisted of generation number, pedigree, sex, true breeding values (TBV), residuals and true phenotype values (TPV) for each trait.

***Multivariate Calibration Models***

The link between the focal trait ('milk protein content') and the other two traits were developed by PLS regression using the 'pls package' (Mevik and Wehrens, 2007) implemented in R (R Core Team, 2016). The PLS regression analyses were done on 2,000 observations randomly sampled without replacement from the 8,000 offspring population in the second generation. The

calibrations were done for each replication within a scenario, and average values of calibration outputs reported. The PLS regression analysis was undertaken using phenotypes (TBVs + errors), as well as the TBV and residuals of the focal trait as dependent variables. In these analyses, the two other traits were used as predictor variables. The PLS regression coefficients (β) from phenotype (β-phenotypic, $\boldsymbol{\beta_p}$), TBV (β-genetic, $\boldsymbol{\beta_g}$), and error (β-residual, $\boldsymbol{\beta_r}$) were used to convert multivariate structures into univariate in the subsequent predictions/calculations. Prediction of EBV and phenotypes were performed following the DP or IP approaches. The PLS coefficient of determinations ($R^2$) for phenotype ($R^2_p$), and TBV ($R^2_g$) were plotted against prediction accuracy of IP and DP approaches to evaluate relationship between accuracy of calibration model and the two approaches. Figure 1 shows a schematic representation of the two prediction approaches.

***Direct Prediction (DP)***

In this approach, procedures were similar to those described in previous studies for prediction of EBV (Dagnachew et al., 2013b; Bonfatti et al., 2017) or phenotypes (Belay et al., 2017). In the current study, however, the steps for dimension reduction of spectral information into few latent variables and variance component estimation for those latent variables were bypassed. Trait 2 and 3 in our simulations were assumed to be the latent variables obtained after dimensional reduction of spectral information. The genetic variances and heritabilities used in simulating trait 2 and 3 were based on our previous study (Belay et al., 2015). BLUP estimates (e.g., EBV) for trait 2 and 3 were obtained by fitting bivariate animal models using Wombat (Meyer, 2007). True (co)variance components were used to predict EBV and other model components including predicted values for trait 2 and 3. The model in matrix notation was as follows:

$$\mathbf{t} = \mathbf{Xb} + \mathbf{Za} + \mathbf{e}$$

Here, **t** is a vector of simulated phenotypes of trait 2 and 3; **b** is a mean (fixed effect); **a** is a vector of random additive genetic effects; **e** is a vector of random residual effects; **X** is a column of ones and **Z** is a design matrix that relates records to the corresponding effects.

The following (co)variance structure for the latent traits was assumed:

$$\text{var}\begin{bmatrix}\mathbf{a}\\ \mathbf{e}\end{bmatrix} = \begin{bmatrix}\mathbf{G}\otimes\mathbf{A} & \mathbf{0}\\ \mathbf{0} & \mathbf{R}\otimes\mathbf{I}\end{bmatrix},$$

where **G** is genetic (co)variance matrix for trait 2 and 3, and **R** is the residual (co)variance matrix. All (co)variance matrices were 2×2. **I** and **A** are identity and additive relationship matrices, respectively, and $\otimes$ is the Kronecker product.

***EBV of Focal Trait from EBV of Trait 2 and 3.*** The predicted EBV ($\hat{a}_{2x2}$) of trait 2 and 3 were directly transformed into EBV of focal trait ($\widehat{EBV}_{DP}$) through either $\beta_p$ or $\beta_g$ i.e. $\widehat{EBV}_{DP} = \hat{\mathbf{a}}_{2x2}\hat{\beta}_{i(2x1)}$, where i=p: phenotypic beta ($\beta_p$) or i=g: genetic beta ($\beta_g$).

***Phenotypes for Focal Trait from Predicted Trait 2 and 3.*** In addition to prediction of the random effects ($\hat{\mathbf{a}}_{2x2}$ and $\hat{\mathbf{e}}_{nx2}$), predicted phenotypes for trait 2 and 3 ($\hat{\mathbf{T}}_{nx2}$) were also computed in the BLUP analysis. Predicted phenotypes not adding the residual term ($\hat{\mathbf{e}}_{nx2}$) for the focal trait ($\widehat{\text{PNE}}_{\text{DP}}$) were computed from the predicted phenotypes of trait 2 and 3 ($\hat{\mathbf{T}}_{nx2}$). The $\hat{\mathbf{T}}_{nx2}$ were converted into predicted focal trait ($\widehat{\text{PNE}}_{\text{DP}}$) either through $\beta_p$ or $\beta_g$. Mathematically $\widehat{\text{PNE}}_{\text{DP}} = \hat{\mathbf{T}}_{nx2}\hat{\beta}_{i(2x1)}$ where i is as defined above and n is number of observations. This predicted focal trait $\widehat{\text{PNE}}_{\text{DP}}$ did not contain residual effects ($\hat{\mathbf{e}}_{nx2}$). Alternatively, the residual effects ($\hat{\mathbf{e}}_{nx2}$) were transformed into univariate form through $\beta_r$, and then added to the $\widehat{\text{PNE}}_{\text{DP}}$ to obtain phenotypes that contained residual effects $\left(\widehat{\text{PWE}}_{\text{DP}}\right)$. Mathematically $\widehat{\text{PWE}}_{\text{DP}} = \hat{\mathbf{T}}_{nx2}\hat{\beta}_{i(2x1)} + \hat{\mathbf{e}}_{nx2}\hat{\beta}_{r(2x1)}$.

***Indirect Prediction (IP)***

In this approach, the focal trait was predicted from the other two simulated traits by classical PLS regression using the $\beta_p$ estimated above ($\widehat{P}_{PLS} = \mathbf{TPV}_{nx2}\widehat{\beta}_{p(2x1)}$), and then genetic analysis (BLUP) was conducted on the predicted phenotypes ($\widehat{P}_{PLS}$). The true (co)variance components (those used in the simulation) of the two other traits were converted into variance components through $\beta_p$ or $\beta_g$ ($\sigma_a^2 = \widehat{\beta}'_{i(1x2)}\mathbf{G}_{2x2}\widehat{\beta}_{i(2x1)}$ and $\sigma_e^2 = \widehat{\beta}'_{i(1x2)}\mathbf{R}_{2x2}\widehat{\beta}_{i(2x1)}$), where i is as defined above, $\sigma_a^2$ is additive genetic variance, and $\sigma_e^2$ is residual variance. These variance components were then used in single-trait BLUP analysis. The same animal models as in the DP, but with a single-trait were fitted. In this case, we assumed $\text{var(a)} = \mathbf{A}\sigma_a^2$, and $\text{var(e)} = \mathbf{I}\sigma_e^2$ , where $\sigma_a^2$ , $\sigma_e^2$, **I** and **A** were as defined above.

Predicted focal trait without residual effects ($\widehat{PNE}_{IP}$) and EBV ($\widehat{EBV}_{IP}$) and solutions for random residuals ($\hat{\mathbf{e}}$) were directly obtained from the BLUP analysis. BLUP analysis of PLS predicted traits ($\widehat{P}_{PLS}$) for prediction of itself (the same phenotype) may be superfluous, but done to conform to the phenotype predicted in the DP ($\widehat{PNE}_{DP}$). Similar to the DP, the residual effects ($\hat{\mathbf{e}}$) from BLUP were added to the $\widehat{PNE}_{IP}$ to obtain a phenotype that contained error (i.e., $\widehat{PWE}_{IP} = \widehat{PNE}_{DP} + \hat{e}$). So, in addition to the $\widehat{P}_{PLS}$, we got two vectors of the predicted focal trait under the IP approach: $\widehat{PNE}_{IP}$ and $\widehat{PWE}_{IP}$.

***Evaluation of the IP and DP Approaches***

Performance of the two approaches were evaluated based on accuracy of EBV or phenotype prediction. Accuracy of EBV prediction was defined as the correlation between TBV and EBV. Pearson's correlation coefficients between DP predicted EBV and TBV of the focal trait were computed and compared with correlation between the IP predicted EBV and TBV. In a similar

manner, accuracy of phenotypic prediction was defined as the correlation between predicted phenotypes and simulated phenotypes. The predicted phenotypes such as $\widehat{\mathrm{PNE}}_{\mathrm{DP}}$ and $\widehat{\mathrm{PWE}}_{\mathrm{DP}}$ from the DP approach, $\widehat{\mathrm{PNE}}_{\mathrm{IP}}$ and $\widehat{\mathrm{PWE}}_{\mathrm{IP}}$ from the IP approach, and $\hat{\mathrm{P}}_{\mathrm{PLS}}$ from classical PLS, are different predictions of the same trait (e.g. milk protein content). Pearson's correlations between those predicted phenotypes and the simulated phenotypes of the focal trait (TPV) were also computed and compared.

## RESULTS AND DISCUSSION

### *Accuracy of EBV Prediction under the DP and IP Approaches*

Accuracy of EBV prediction for sires as well as for all animals with records were computed. However, only sire evaluations are presented here, because the trends for the IP and DP approaches were similar. Accuracy of sire EBV predicted using the DP and IP approaches are presented in Table 3. In the DP approach, EBV of trait 2 and 3 were converted into EBV of the focal trait ('milk protein content') using PLS regression coefficients estimated based on TBV ($\beta_g$) or TPV – true phenotypic values – ($\beta_p$). In the IP approach, the covariance components of trait 2 and 3 were converted into variance of the focal trait using the $\beta_g$ or $\beta_p$. Those variances were used in univariate BLUP analyses for EBV or phenotype prediction. In both DP and IP approaches, the accuracy of sire EBV predictions increased as the genetic correlations between all the traits increased regardless of residual correlations. At a given genetic correlation scenario, EBV accuracy also increased with increase in residual correlation, except at low genetic correlation scenario where no clear trend was observed. Increasing in residual correlations should not increase accuracy of EBVs except that residual correlations between traits could lead to artificial resemblance between EBV because the phenotypes are correlated.

In all possible combination of genetic and residual correlation scenarios, regardless of the types of βs used, accuracies of EBV prediction using the DP approach were higher than the corresponding predictions in the IP approach. The exception was at zero residual correlations, where there was little to no difference in accuracy between the two approaches (Table 3). In agreement with the current study, Dagnachew et al. (2013b) reported reduction in prediction error variance (and hence increase in accuracy) for using DP approach instead of IP approach. We have also evaluated performance of the two approaches when there was no or little correlations between trait 2 and 3, but with varying correlations between the focal trait and the other two traits, and found similar EBV accuracy prediction in both IP and DP approaches (results not shown). This indicates that there should be some correlations between the other two traits, in addition to the correlations to the focal trait for the DP to perform better than the IP approach. This is because the DP approach utilizes covariance structure between the other two traits to predict EBV for the focal trait. If there is no covariance between the other two traits, there is no mathematical explanation to expect better performance for the DP than the IP approach. For better EBV accuracy prediction in the DP approach, existence of correlations between focal trait and the other two traits as well as between the other two traits are a prerequisite.

Comparison between the effects of using $\beta_p$ and $\beta_g$ on accuracy of EBV were made within each approach. In the DP approach, accuracy of EBV when using $\beta_g$ were higher than the corresponding predictions using $\beta_p$, especially in the low genetic scenario. This suggested that if the genetic correlations between the three traits are low, an appropriate regression coefficient (e.g. $\beta_g$) should be used for better EBV accuracy in the DP approach. However, there were no difference in EBV accuracy between using the $\beta_p$ and $\beta_g$ in the IP approach. This is interesting as the IP approach is

the commonly used method in genetic analyses of traits predicted from milk FT-MIR spectra. It might be difficult to estimate $\beta_g$ from real data, so that it is a challenge.

Both in the current study that depends on simulated data and a previous study that used real data (Dagnachew et al., 2013b), the DP approach has improved accuracy of EBV prediction. In the DP approach, EBV are predicted once for latent variables derived from spectra and later combined into EBV of focal traits without a need of first predicting phenotypes for the focal traits from milk spectra. This is particularly important when FT-MIR calibration equations are available for a high number of traits (Bonfatti et al., 2017). Such use of the DP approach for spectra would be possible when the spectral dimension is reduced by principal component analysis **(PCA)**, and not as feasible with PLS. Parameter estimates for the latent variables from PCA are population parameters that characterize any information available in the milk spectra. This is not the case, for example, for factor scores from the PLS regression as they mainly contain information related to the particular trait used in the calibration. However, the retained latent variables from PCA might not contain all information about the focal trait (Soyeurt et al., 2010; Dagnachew et al., 2013b; Belay et al., 2015; Bonfatti et al., 2017).

***Relationships between Accuracy of EBV Prediction and Calibration Equations***

Coefficient of determinations in calibration models estimated from simulated phenotypes ($R_p^2$) ranged from 0.004 (for low genetic and low residual correlation scenario) to 0.787 (for high genetic and high residual correlation scenario). The $R_p^2$ increased with the increase in either genetic or residual or both correlations (Figures 2, 4 and 6). The corresponding estimates from true breeding values ($R_g^2$) ranged from 0.070 (for low genetic correlation) to 0.814 (for high genetic correlation). The $R_g^2$ increased with increase in genetic correlations, but did not change as residual correlations

increased (Figures 3, 5 and 7). At most of the correlation scenarios considered, estimates for the $R_g^2$ were higher the corresponding values for the $R_p^2$.

The relationships between performance of the IP and DP approaches in predicting individual EBV and predictive ability of calibration models ($R^2$) are depicted in Figure 2 for $\beta_p$ and $R_p^2$ and in Figure 3 for $\beta_g$ and $R_g^2$. When using the $\beta_p$, prediction accuracy of the IP and DP approaches increased with increase in predictive ability of the calibration models ($R_p^2$). The exception was at low genetic correlation (with zero to high residual correlation scenarios), where EBV accuracy generally decreased as the $R_p^2$ increased. This indicated that improvement in residual structure for lowly genetically correlated traits has no contribution in improving their EBV prediction accuracy. In other words, calibration models with higher $R^2$ do not necessarily result in better EBV prediction accuracies when genetic correlations between traits are low. What really matters for better accuracy of EBV prediction is the genetic correlation compared to residual correlation. For example, at zero (Figure 2a) or low (Figure 2b) residual structure, $R_p^2$ of the calibration models slightly increased with increase in genetic correlations (low to high), but EBV accuracy greatly improved compared to the $R_p^2$. At high residual correlation (Figure 2d) as well as at low residual correlation (Figure 2b) with medium to high genetic correlation scenarios, the difference in prediction accuracy between the IP and DP approaches were more visible than at either zero (Figure 2a) or medium (Figure 2c) residual correlations. This was more clear for EBV of all animals with records than for EBV of sires i.e. as the predictive ability of the calibration models increased the difference in prediction accuracy between the IP and DP approaches became more apparent (results not shown).

Similar to when using the $\beta_p$, prediction accuracy of the IP and DP approaches in using the $\beta_g$ increased with increase in predictive ability ($R_g^2$) of the calibration equations (Figure 3). As expected, the residual structure (Figure 3a-d) in this case had no effect on predictive ability of the calibration models, but on accuracy of EBV prediction, as the predicted EBV was derived from a phenotype that contained both residual and genetic information. Increase in genetic correlations between traits increased both accuracy of the two approaches (IP and DP) and predictive ability of the calibration models regardless of the residual structures. Except at zero residual correlation (Figure 3a), difference in performance of the two approaches became visible as the residual correlations increased (Figure 3b-d), especially at low genetic scenarios. However, such clear differences were not observed when using $\beta_p$ (Figure 2), indicating that performance of the DP approach is sensitive to the type of PLS regression coefficients used.

The better performance of DP when using $\beta_g$, especially at low genetic correlation might be due to the simultaneous increase in genetic correlations between the focal trait and the other two traits (trait 2 and 3) as well as between trait 2 and 3. As the genetic correlation between the traits increases, their information content becomes similar (i.e. one trait provides more information about the other traits) and hence simultaneous analysis of such traits would be slightly different from analyzing them separately. The better performance of DP in using $\beta_g$ at low genetic correlation scenario might also be related to predictive ability of calibration models from which $\beta_g$ was estimated. At low genetic correlation, predictive ability of calibration models was low (<0.10) and $\beta_g$ captures little information about the intended trait. Hence, EBV of trait 2 and 3 were predicted more accurately due to utilization of covariance between them (DP). However, the corresponding EBV prediction in IP was relatively inaccurate as the phenotypes from which the EBV derived were poorly predicted through a $\beta_g$ that contained little information. At high predictive ability of

calibration models, the IP approach gives accurate prediction and not much gain from the DP that utilizes little extra information not utilized by the IP approach, resulting in small difference in performance between the approaches. That means performance of the IP approach is approaching that of the DP approach as accuracy of the calibration models increased.

In our previous study, we have suggested that predictive ability of calibration model could affect performance of the two approaches. The suggestion was made based on the work of Dagnachew et al. (2013b) where high $R^2$ was associated with better performance in DP than in IP and Belay et al. (2017) where low $R^2$ was associated with inferior performance in the DP compared to the IP approach. However, Bonfatti et al. (2017) reported absence of relationships between rank correlations (between EBV obtained by the IP and the DP approach) and predictive ability of calibration equations. In the current study, we have shown that performance of the two approaches increased with increase in predictive ability of calibration models, but it is not necessary to have calibration equation with high $R^2$ for the DP approach to perform better the IP approach for EBV prediction. This might not be the case for phenotypic prediction (details in next sections).

***Accuracy of Phenotypic Prediction ($\widehat{PNE}_{IP}$ and $\widehat{PNE}_{DP}$)***

Similar to accuracies of EBV, accuracies of predicted phenotypes generally increased with increase in correlations (genetic and residual) between the focal trait and the other two traits. However, no clear trend was observed for DP compared to IP performance when the residual correlation increased at low genetic scenario when using $\beta_g$ (Table 4). Accuracies of predicted phenotypes in IP approach were higher than the corresponding estimates in the DP approach both when using $\beta_g$ and $\beta_p$. This is in agreement with our previous study (Belay et al., 2017) where slightly better prediction of blood BHB was found in the IP than in the DP approach. However, it is in contrast with EBV accuracy observed in the current (Table 3 or Figures 2 and 3) and previous

(Dagnachew et al., 2013b) studies. Performance of the IP over the DP increased with increase both in genetic and residual correlations among traits, except at the zero residual correlation scenario where performance of the two approaches were similar. The differences in performance between the two approaches were larger especially at medium or high residual correlation scenarios. For the DP approach, no or little difference in performance was observed in using $\beta_g$ or $\beta_p$, except at low genetic scenario where using $\beta_p$ generally gave higher accuracies (Table 4). However, within the IP approach, accuracies of predicted phenotypes were generally slightly higher when using $\beta_g$ than when using $\beta_p$. This is contrary to the EBV accuracy comparison made between using the $\beta_g$ and $\beta_p$ within either the DP or IP approach. Similarly, at low genetic correlation, accuracy of phenotypic prediction was higher when using $\beta_p$ in DP, but accuracy of EBV prediction was higher in using the $\beta_g$ in the DP approach.

***Accuracy of the $\widehat{PNE}_{IP}$ and $\widehat{PNE}_{DP}$ versus Accuracy of Calibration Equations***

The relationships between phenotype prediction accuracy of the IP and DP approaches with the predictive ability of calibration models ($R^2$) were evaluated and are depicted in Figure 4 when using $\beta_p$ and $R_p^2$ and in Figure 5 in using $\beta_g$ and $R_g^2$. Similar to the accuracy of EBV, phenotypic prediction accuracy increased with increase in predictive ability of calibration models. Both at zero (Figure 4a) and low (Figure 4b) residual correlation scenarios, predictive ability of calibration model was very low ($R_p^2$ <0.1) and most increment in accuracy of phenotypic prediction in the two approaches were due to increase in genetic correlations. At medium residual correlation scenario (Figure 4c), the $R_p^2$ increased to a maximum of 0.56 while the corresponding values at high residual correlation scenario (Figure 4d) was 0.79. The corresponding maximum value in accuracy of phenotype was 0.55 for DP or 0.65 for IP at the medium residual correlation scenario and was 0.56 for DP or 0.74 at the high residual correlation scenario. As indicated above, the IP approaches

perform better than the DP for predicting the phenotypes, and the difference in prediction accuracy between the two approaches become clearer as predictive ability of calibration models ($R_p^2$) increased (Figure 4).

In using the $\beta_g$, phenotypic prediction accuracy of the IP and DP approaches also increased with increase in predictive ability of the calibration models ($R_g^2$). Increase in genetic correlations between traits increased both performance of the two approaches and the calibration models regardless of the residual structures (Figure 5). At zero residual correlation (Figure 5a), there was no difference in performance between the two approaches. However, at the remaining residual correlation scenarios (Figure 5b-d), the IP approach outperformed the DP approach with increase in $R_g^2$ of calibration models, and the highest difference in performance between the two approaches was observed at high residual correlation scenario (Figure 5d). Unlike when using $\beta_p$, it is possible to distinguish between performance of the two approaches when using $\beta_g$ at low $R_g^2$ (< 0.1), where genetic and residual correlations between traits were low.

***Accuracy of Predicted Phenotypes ($\widehat{PWE}_{IP}$ and $\widehat{PWE}_{DP}$)***

In this study, phenotype for the focal trait was predicted in two ways: 1) phenotypic prediction obtained directly from BLUP when such phenotypes did not contain residual effects, and 2) those phenotypes from BLUP + residual effects from BLUP. The former was described in the preceding section while the latter phenotypes are presented in this section. In the DP approach, BLUP solutions for residuals of trait 2 and 3 were converted into residual of focal trait using $\beta_r$ that were estimates from residual part of simulated phenotypes. This was done assuming that use of appropriate PLS regression coefficients for the different model components would result in better prediction than using only regression coefficient estimated from phenotype. Accordingly,

predicted phenotypes of trait 2 and 3 were converted into single-trait predicted phenotypes ($\widehat{\mathrm{PNE}}_{\mathrm{DP}}$) using either $\beta_p$ or $\beta_g$ (as described above), whereas $\beta_r$ was used to convert predicted residual of trait 2 and 3 into single-trait residual ($\hat{e}$). This single-trait residual ($\hat{e}$) was later added to the $\widehat{\mathrm{PNE}}_{\mathrm{DP}}$ to obtain predicted phenotypes for the focal trait ($\widehat{\mathrm{PWE}}_{\mathrm{DP}} = \widehat{\mathrm{PNE}}_{\mathrm{DP}} + \hat{e}$). In the IP approach, predicted phenotypes (i.e., phenotypes corrected for residual effects:$\widehat{\mathrm{PNE}}_{\mathrm{IP}}$) and residual effects directly obtained from BLUP were added together to generate new predicted phenotypes for the focal trait ($\widehat{\mathrm{PWE}}_{\mathrm{IP}} = \widehat{\mathrm{PNE}}_{\mathrm{IP}} + \hat{e}$). In addition to the DP and IP predicted phenotypes ($\widehat{\mathrm{PWE}}_{\mathrm{IP}}$ and $\widehat{\mathrm{PWE}}_{\mathrm{DP}}$), phenotypes predicted by classical PLS ($\hat{\mathrm{P}}_{\mathrm{PLS}}$) were also computed and compared with the $\widehat{\mathrm{PWE}}_{\mathrm{IP}}$ and $\widehat{\mathrm{PWE}}_{\mathrm{DP}}$.

Similar to accuracies of phenotypes corrected for residual effects (i.e. $\widehat{\mathrm{PNE}}_{\mathrm{IP}}$ and $\widehat{\mathrm{PNE}}_{\mathrm{DP}}$,), accuracies of phenotypes that contained residual effects (i.e. $\widehat{\mathrm{PWE}}_{\mathrm{IP}}$ and $\widehat{\mathrm{PWE}}_{\mathrm{DP}}$) increased with increase in genetic and residual correlations (Table 5). In using $\beta_p$, accuracy of phenotypic prediction in the IP approach was generally similar to the accuracy in the DP approach. This is contrary to accuracy of the $\widehat{\mathrm{PNE}}_{\mathrm{IP}}$ and $\widehat{\mathrm{PNE}}_{\mathrm{DP}}$ reported in the current (Table 4) and previous (Belay et al., 2017) studies. The exceptions were at zero residual correlation with medium and high genetic correlation scenarios, where the DP performed better than the IP approach, but with relatively high standard errors. At low residual correlation with medium and high genetic correlation scenarios, however, IP performed better than the DP approach. In using $\beta_g$, however, accuracy of phenotypic prediction in the DP approach was slightly higher than accuracy in the IP approach. In addition, accuracy of phenotypic prediction in the DP approach was higher when using $\beta_g$ than using $\beta_p$, especially at zero and low residual correlations. This gives a clue to the importance of using appropriate regression coefficients ($\beta_g$ and $\beta_r$) in the DP approach for converting the multi-trait into single-trait structure. Within the IP approach, accuracy of phenotypic prediction using $\beta_p$ was

equal to the one predicted in using $\beta_g$. These phenotypic prediction accuracies were also equal to the ones predicted by PLS (results not shown in table); indicating unnecessity of doing single-trait BLUP (IP) on PLS predicted traits for ultimate prediction of phenotypes.

***Accuracy of $\widehat{PWE}_{IP}$ and $\widehat{PWE}_{DP}$ Prediction versus Accuracy of Calibration Equations***

The relationships between accuracy of the $\widehat{\text{PWE}}_{\text{IP}}$ or $\widehat{\text{PWE}}_{\text{DP}}$ and the predictive ability of calibration models ($R^2$) are depicted in Figure 6 for $\beta_p$ and $R_p^2$ and in Figure 7 for $\beta_g$ and $R_g^2$. When using the $\beta_p$, accuracy of phenotypic prediction increased with increase in $R_p^2$ (Figure 6). This is similar to accuracy of EBV and $\widehat{\text{PNE}}_{\text{IP}}$ or $\widehat{\text{PNE}}_{\text{DP}}$ described above in the current study. Differences in performance between the two approaches were clearer at low $R_p^2$ (<0.1; Figure 6a-b) than at medium to high $R_p^2$ (0.277-0.787; Figure 6c-d). This is contrary to performance of the two approaches for predicting accuracy of EBV and $\widehat{\text{PNE}}_{\text{IP}}$ or $\widehat{\text{PNE}}_{\text{DP}}$, where performance of the two approaches overlap at low $R_p^2$ (< 0.1; Figures 2 and 4).

In using the $\beta_g$, the prediction accuracy of the IP and DP approaches also increased with increase in $R_g^2$ of calibration models (Figure 7). Similar to accuracy of EBV, $\widehat{\text{PNE}}_{\text{IP}}$, or $\widehat{\text{PNE}}_{\text{DP}}$, the residual structures had no effect on the $R_g^2$ of calibration models, but on the accuracy of phenotypic prediction by the two approaches (Figure 7). Increase in genetic correlations between traits increased both performance of the two approaches and $R_g^2$ of calibration models regardless of the residual structures. At zero (Figure 7a) and low (Figure 7b) residual correlation scenarios, differences in performances between the two approaches increased with increase in $R_g^2$ of calibration models or with increase in genetic correlations. At medium (Figure 7c) and high (Figure 7d) residual correlation scenarios, however, performance of the two approaches were similar regardless of increase in $R_g^2$ or in genetic correlations. These results are contrary to

performance of the two approaches in predicting accuracy for the $\widehat{\text{PNE}}_{\text{IP}}$ and $\widehat{\text{PNE}}_{\text{DP}}$ phenotypes (Figures 4 and 5). This indicated that for traits with low link to predictor variables better phenotypic prediction would be found with the DP in using both $\beta_g$ and $\beta_r$ together.

## CONCLUSIONS

In this study, performance of the IP and DP approaches under different genetic and residual correlation scenarios were evaluated. In addition, effects of using different regression coefficients ($\beta_g$, $\beta_p$ or $\beta_r$) on accuracy of prediction (EBV and phenotype) were investigated. The relationships between performance of the IP and DP approaches and accuracy of calibration models (for phenotype, genetic and residual values of the focal trait) also studied. Accuracies of EBV were higher in the DP approach than in the IP approach, whereas the reverse was true for accuracy of phenotypic prediction (i.e. accuracy of $\widehat{\text{PNE}}_{\text{IP}} > \widehat{\text{PNE}}_{\text{DP}}$). The exception was when using $\beta_g$ and $\beta_r$, where accuracy of phenotypic prediction in the DP approach was slightly higher than that in the IP approach, especially at the zero and low residual correlation scenarios (i.e. accuracy of $\widehat{\text{PWE}}_{\text{IP}} < \widehat{\text{PWE}}_{\text{DP}}$). Predictive ability of the calibration models increased with improvement in genetic and residual structures between traits. Performance of both IP and DP increased with increase in predictive ability of the calibration models. The exceptions were when using $\beta_g$ (where performance of the two approaches were not affected by predictive ability of the calibration models at a given genetic scenario) and at low genetic correlation scenario (where accuracy of EBV prediction of the two approaches slightly decreases as the $R_p^2$ was increased). Therefore, it is not a good strategy to use the DP approach for phenotypic prediction, except when the $\beta_g$ and $\beta_r$ that are difficult to estimate using real data would be available. Use of the DP approach for prediction of EBV seems useful while the IP or PLS based prediction equations are a method of choice for phenotypic prediction.

## ACKNOWLEDGEMENTS

The Department of Animal and Aquacultural Sciences, Norwegian University of Life Sciences financially supported this study.

## REFERENCES


Belay, T., B. Dagnachew, Z. Kowalski, and T. Ådnøy. 2017. An attempt at predicting blood β-hydroxybutyrate from Fourier-transform mid-infrared spectra of milk using multivariate mixed models in Polish dairy cattle. Journal of Dairy Science 100(8):(in press).

Belay, T. K., K. Słoniewski, Z. Kowalski, and T. Adnoy. 2015. Predicting ketosis from milk mid infrared (MIR) spectra using multivariate mixed models. In Proc. Third DairyCare Conference, Croatia, Zadar. Accessed Mar. 16, 2017. www.dairycareaction.org/uploads/2/4/2/6/24266896/belay_tuesday.pdf.

Bonfatti, V., D. Vicario, L. Degano, A. Lugo, and P. Carnier. 2017a. Comparison between direct and indirect methods for exploiting Fourier transform spectral information in estimation of breeding values for fine composition and technological properties of milk. Journal of Dairy Science 100(3):2057-2067.

Dagnachew, B. S., A. Kohler, and T. Ådnøy. 2013a. Genetic and environmental information in goat milk Fourier transform infrared spectra. Journal of Dairy Science 96(6):3973-3985.

Dagnachew, B. S., T. H. E. Meuwissen, and T. Ådnøy. 2013b. Genetic components of milk Fourier-transform infrared spectra used to predict breeding values for milk composition and quality traits in dairy goats. Journal of Dairy Science 96(9):5933-5942.

Mevik, B.-H. and R. Wehrens. 2007. The pls package: principal component and partial least squares regression in R. Journal of Statistical software 18(2):1-24.

Meyer, K. 2007. WOMBAT—A tool for mixed model analyses in quantitative genetics by restricted maximum likelihood (REML). Journal of Zhejiang University Science B 8(11):815-821.

Soyeurt, H., I. Misztal, and N. Gengler. 2010. Genetic variability of milk components based on mid-infrared spectral data. Journal of dairy science 93(4):1722-1728.

R Core Team (2016). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

## FIGURES

**Figure 1**. Schematic representation of the indirect (IP) and direct (DP) prediction approaches. For the IP method, phenotype for focal trait was predicted ($\widehat{P}_{PLS}$) from trait 2 and 3 using regression coefficient ($\beta_p$) estimated from true phenotype values (TPV). Then, BLUP was applied to this predicted trait using the true genetic and residual covariance of trait 2 and 3 after converting into respective variances through $\beta_p$ or $\beta_g$ (regression coefficient from true breeding values-TBV) for EBV ($\widehat{EBV}_{IP}$) and phenotypic ($\widehat{PNE}_{IP}$) prediction. The $\widehat{PNE}_{IP}$that obtained directly from BLUP do not contain residual effects (ê) whose addition to the $\widehat{PNE}_{IP}$ resulted in phenotype that contained error ($\widehat{PWE}_{IP}$). For the DP method, trait 2 and 3 were analyzed multivariately with a genetic model before predicted model components are combined through the $\beta_p$ or $\beta_g$ to eventually predict phenotype ($\widehat{PNE}_{DP}$) and $EBV_{DP}$. Residual part of predicted trait 2 and 3 ($\hat{e}_t$) was converted into single-trait residual (ê) through the $\beta_r$ (regression coefficient estimated from the residual part of simulated phenotypes) and the ê added to the $\widehat{PNE}_{DP}$ to obtain predicted phenotype that contained residual effects ($\widehat{PWE}_{DP}$).

**Figure 2**. Determination coefficients of calibration models estimated based on phenotype ($R_p^2$) and mean EBV accuracy of sire predicted using the direct (DP; black circle) and indirect (IP; black triangle) prediction approaches. The mean EBV accuracy predicted under low (LG), medium (MG) and high (HG) genetic correlations with zero (a), low (b), medium (c) and high (d) residual correlation scenarios. The gray shades indicate genetic correlation scenarios. The $\beta_p$ was used to convert covariance components of trait 2 and 3 into variance components in IP or EBV of trait 2 and 3 into EBV of the intended trait in DP.

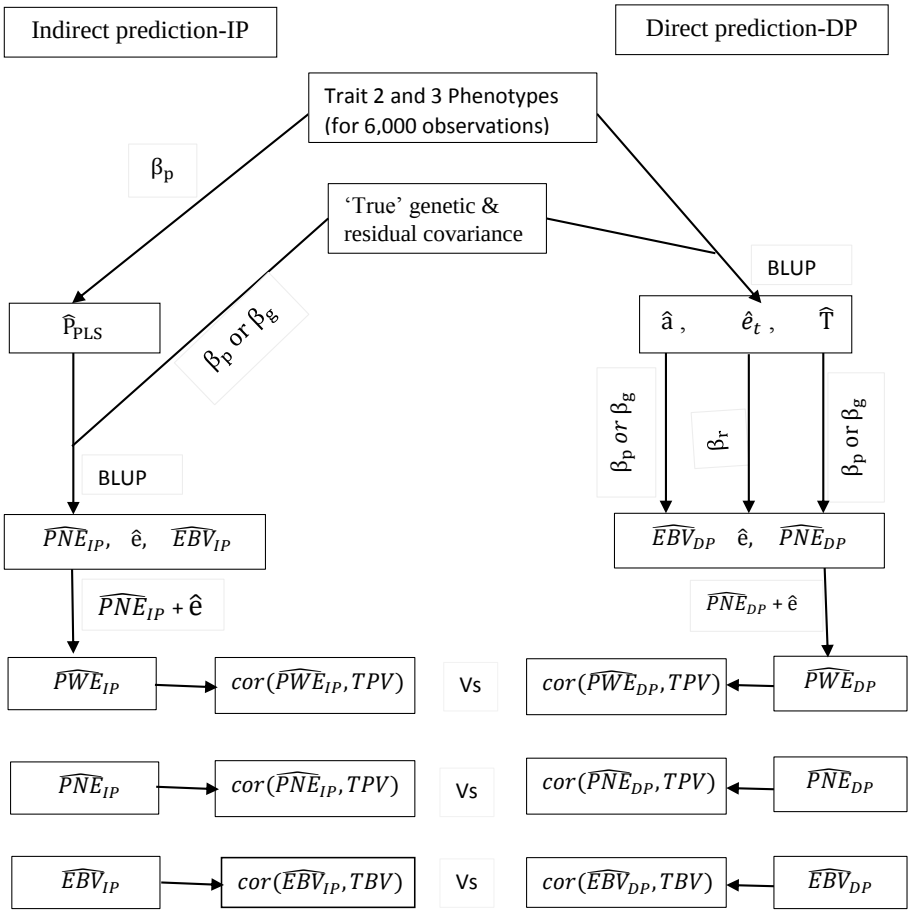
**Figure 3.** Determination coefficients of calibration models estimated based on true breeding values ($R_g^2$) and mean EBV accuracy of sire predicted using the direct (DP; black circle) and indirect (IP; black triangle) prediction approaches. The mean EBV accuracy predicted under low (LG), medium (MG) and high (HG) genetic correlations with zero (a), low (b), medium (c) and high (d) residual correlation scenarios. The gray shades indicate genetic correlation scenarios. The $\beta_g$ was used to convert covariance components of trait 2 and 3 into variance components in IP or EBV of trait 2 and 3 into EBV of the intended trait in DP.

**Figure 4.** Determination coefficients of calibration models estimated based on phenotype ($R_p^2$) and mean accuracy of predicted phenotypes (without residual effects) using the direct (DP; black circle) and indirect (IP; black triangle) prediction approaches. The mean phenotypic accuracy predicted under low (LG), medium (MG) and high (HG) genetic correlations with zero (a), low (b), medium (c) and high (d) residual correlation scenarios. The gray shades indicate genetic correlation scenarios. The $\beta_p$ was used to convert covariance components of trait 2 and 3 into variance components to be used in IP or predicted phenotypes of trait 2 and 3 into predicted phenotypes of the intended trait in DP.

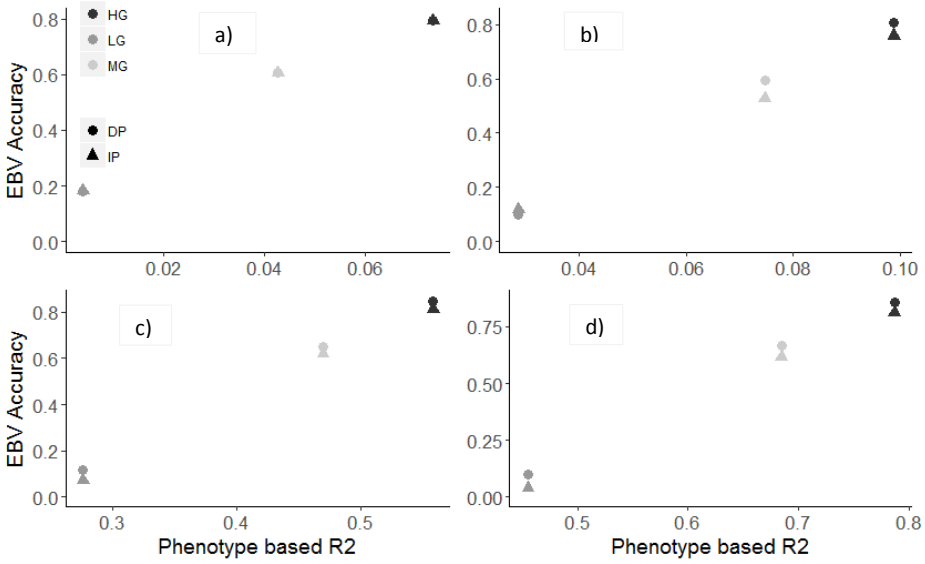
**Figure 5**. Determination coefficients of calibration models estimated based on true breeding values ($R_g^2$) and mean accuracy of predicted phenotypes (without residual effects) using the direct (DP; black circle) and indirect (IP; black triangle) prediction approaches. The mean phenotypic accuracy predicted under low (LG), medium (MG) and high (HG) genetic correlations with zero (a), low (b), medium (c) and high (d) residual correlation scenarios. The gray shades indicate genetic correlation scenarios. The $\beta_g$ was used to convert covariance components of trait 2 and 3 into variance components to be used in IP or predicted phenotypes of trait 2 and 3 into predicted phenotypes of the intended trait in DP.

**Figure 6.** Determination coefficients of calibration models estimated based on phenotypes ($R_p^2$) and mean accuracy of predicted phenotypes (including residual effects) using the direct (DP; black circle) and indirect (IP; black triangle) prediction approaches. The mean phenotypic accuracy predicted under low (LG), medium (MG) and high (HG) genetic correlations with zero (a), low (b), medium (c) and high (d) residual correlation scenarios. The gray shades indicate genetic correlation scenarios. The $\beta_p$ was used to convert covariance components of trait 2 and 3 into variance components to be used in IP or predicted phenotypes of trait 2 and 3 into predicted phenotypes of the intended trait in DP.

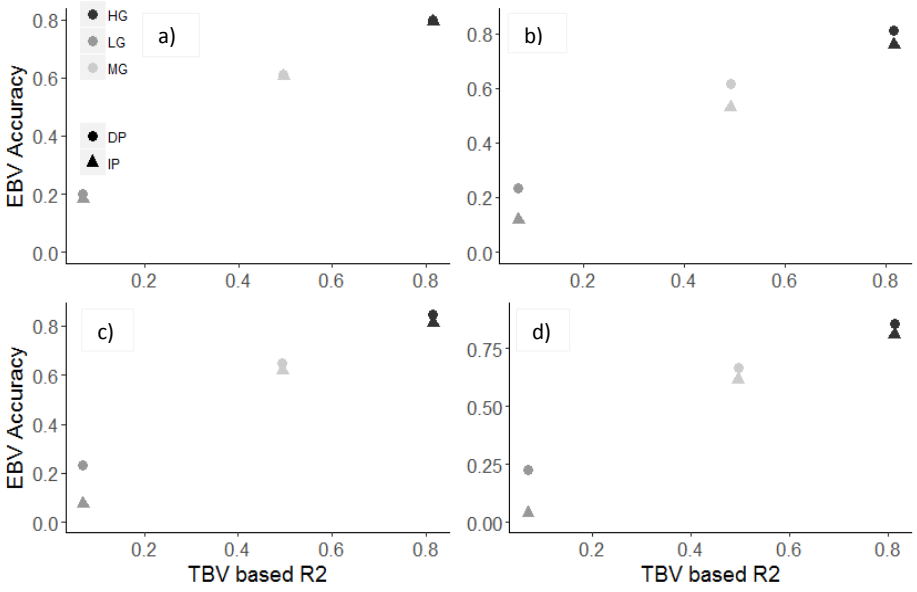
**Figure 7.** Determination coefficients of calibration models estimated based on true breeding values ($R_g^2$) and mean accuracy of predicted phenotypes (including residual effects) using the direct (DP; black circle) and indirect (IP; black triangle) prediction approaches. The mean phenotypic accuracy predicted under low (LG), medium (MG) and high (HG) genetic correlations with zero (a), low (b), medium (c) and high (d) residual correlation scenarios. The gray shades indicate genetic correlation scenarios. The $\beta_g$ was used to convert covariance components of trait 2 and 3 into variance components to be used in IP or predicted phenotypes of trait 2 and 3 into predicted phenotypes of the intended trait in DP.



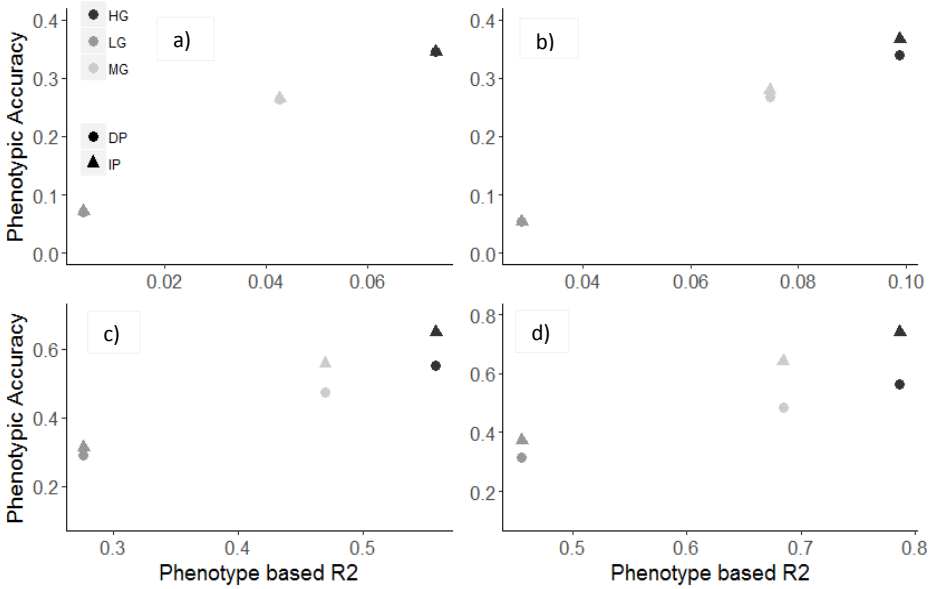

**Belay et al. Figure 1**



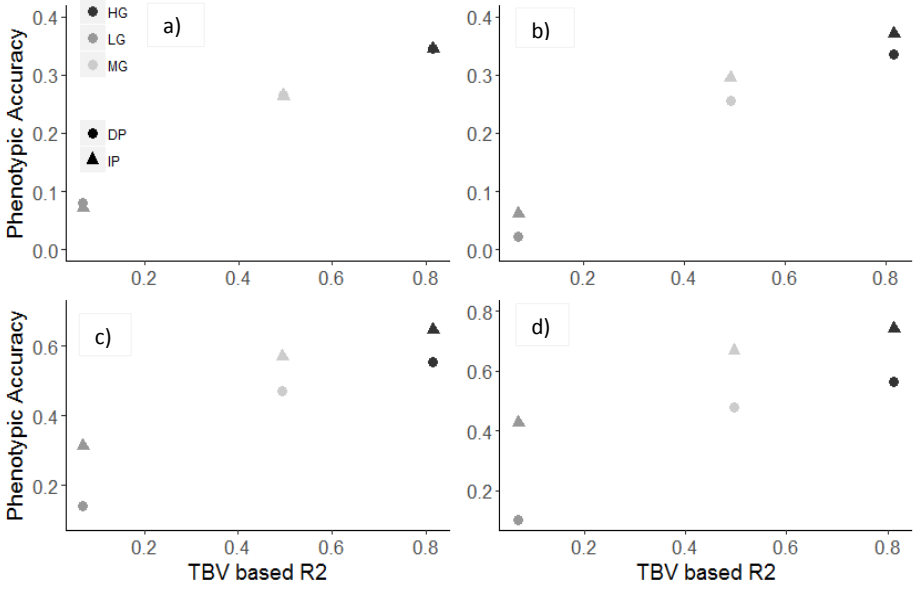

**Belay et al. Figure 2**



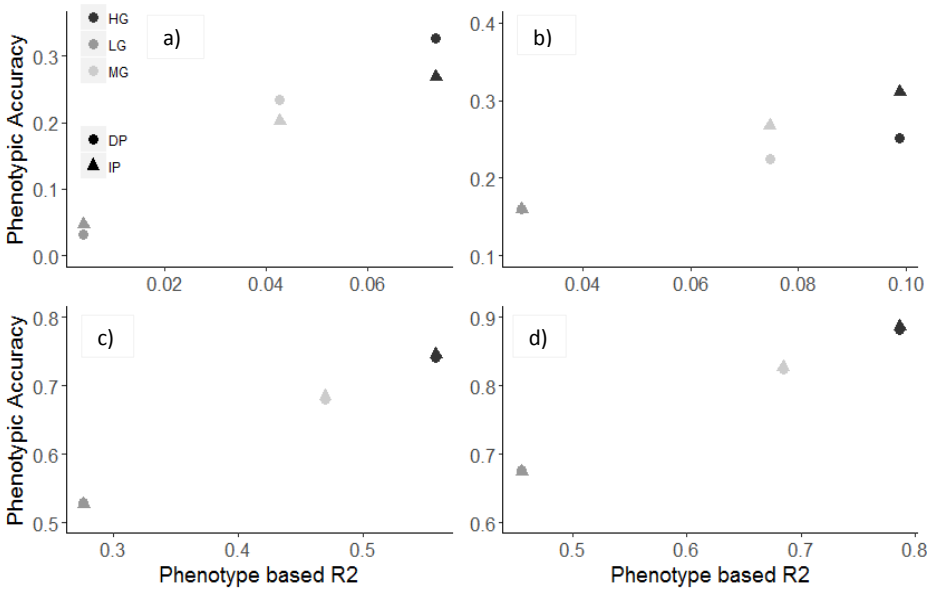

**Belay et al. Figure 3**



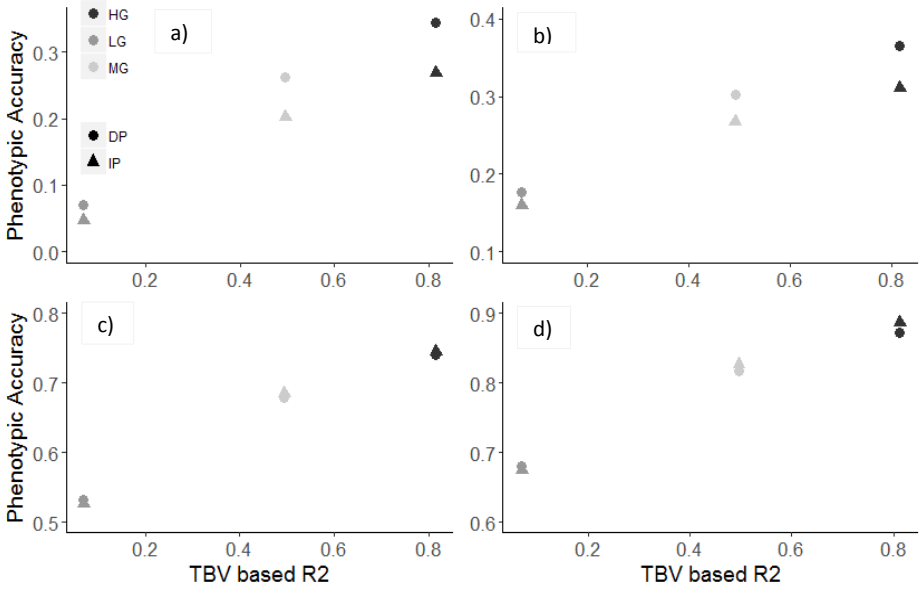

**Belay et al. Figure 4**




**Belay et al. Figure 5**




**Belay et al. Figure 6**




**Belay et al. Figure 7**

## Tables

**Table 1.** Overall means, genetic and residual variances and heritabilities of the simulated traits*

| Trait | Mean | Genetic variance | Residual variance | Heritability |
|---|---|---|---|---|
| Trait1 | 3.450 | 0.027 | 0.044 | 0.377 |
| Trait2 | -0.030 | 0.089 | 0.839 | 0.096 |
| Trait3 | -0.050 | 0.158 | 0.686 | 0.187 |

*Except residual variance, all other parameters were estimates from previous studies (estimates for trait1 that represent protein percent are from Belay et al., unpublished data) while estimates for the other two traits (trait2 and trait3) represent latent traits derived from milk spectra in Belay et al. (2015). Residual variance was calculated given the heritability and genetic variance of each trait.

**Table 2.** Scenarios for genetic and residual correlations between the three traits used in simulations

| | Scenarios | | | | | | |
|---|---|---|---|---|---|---|---|
| | Genetic | | | Residual | | | |
| Correlation(r)* | High | Medium | Low | High | Medium | Low | Zero |
| $r_{12}$ | 0.80 | 0.50 | 0.25 | -0.80 | -0.50 | -0.25 | 0 |
| $r_{13}$ | 0.90 | 0.70 | 0.10 | 0.90 | 0.70 | 0.10 | 0 |
| $r_{23}$ | 0.85 | 0.65 | 0.15 | -0.85 | -0.60 | -0.15 | 0 |

*$r_{12}$, $r_{13}$, and $r_{23}$ are correlations between traits 1 and 2, 1 and 3, and 2 and 3, respectively. Trait 1 is the focal trait ('protein percent') and trait 2 and 3 are assumed to be latent traits after dimensionality reduction of spectra information.

**Table 3.** Mean accuracy*(SE) of predicted breeding values of sires using the direct (DP) and indirect (IP) prediction approaches under different correlation scenarios using regression coefficients estimated based on true breeding values ($\beta_g$) and true phenotype ($\beta_p$)

| Correlation | | DP | | IP | |
|---|---|---|---|---|---|
| Genetic | Residual | $\beta_p$ | $\beta_g$ | $\beta_p$ | $\beta_g$ |
| Low: | Zero | 0.179 (0.123) | 0.199 (0.115) | 0.182 (0.125) | 0.182 (0.125) |
| | Low | 0.099 (0.105) | 0.236 (0.108) | 0.117 (0.105) | 0.117 (0.105) |
| | Medium | 0.118 (0.110) | 0.233 (0.101) | 0.074 (0.111) | 0.073 (0.112) |
| | High | 0.100 (0.103) | 0.227 (0.096) | 0.040 (0.106) | 0.040 (0.106) |
| Medium: | Zero | 0.607 (0.072) | 0.612 (0.070) | 0.608 (0.072) | 0.608 (0.072) |
| | Low | 0.596 (0.077) | 0.617 (0.074) | 0.528 (0.089) | 0.528 (0.088) |
| | Medium | 0.650 (0.057) | 0.650 (0.057) | 0.619 (0.061) | 0.619 (0.061) |
| | High | 0.665 (0.064) | 0.666 (0.064) | 0.618 (0.070) | 0.617 (0.070) |
| High: | Zero | 0.796 (0.035) | 0.798 (0.035) | 0.796 (0.035) | 0.796 (0.035) |
| | Low | 0.809 (0.036) | 0.811 (0.035) | 0.759 (0.049) | 0.758 (0.049) |
| | Medium | 0.846 (0.029) | 0.846 (0.029) | 0.813 (0.033) | 0.813 (0.033) |
| | High | 0.855 (0.033) | 0.855 (0.033) | 0.810 (0.041) | 0.810 (0.041) |

*Accuracy was defined as a correlation between predicted EBV and simulated 'true' breeding values (TBV). Average of 100 replicates per scenario is reported and the SE were calculated as the standard deviation of the 100 accuracies for each scenario.

**Table 4.** Mean accuracy*(SE) of predicted phenotypes (corrected for residual effects) using the direct (DP) and indirect (IP) prediction approaches under different correlation scenarios using regression coefficients that estimated based on true breeding values ($\beta_g$) and true phenotype ($\beta_p$)

| Correlation | | DP | | IP | |
|---|---|---|---|---|---|
| Genetic | Residual | $\beta_p$ | $\beta_g$ | $\beta_p$ | $\beta_g$ |
| Low | Zero | 0.069 (0.038) | 0.079 (0.034) | 0.072 (0.039) | 0.072 (0.040) |
| | Low | 0.055 (0.032) | 0.022 (0.036) | 0.053 (0.033) | 0.061 (0.035) |
| | Medium | 0.290 (0.028) | 0.140 (0.050) | 0.314 (0.034) | 0.312 (0.057) |
| | High | 0.314 (0.024) | 0.102 (0.051) | 0.372 (0.035) | 0.426 (0.073) |
| Medium | Zero | 0.265 (0.024) | 0.266 (0.023) | 0.265 (0.024) | 0.264 (0.025) |
| | Low | 0.267 (0.028) | 0.256 (0.027) | 0.280 (0.032) | 0.295 (0.028) |
| | Medium | 0.474 (0.016) | 0.471 (0.016) | 0.557 (0.016) | 0.569 (0.019) |
| | High | 0.483 (0.015) | 0.479 (0.016) | 0.640 (0.017) | 0.667 (0.024) |
| High | Zero | 0.345 (0.020) | 0.346 (0.020) | 0.345 (0.020) | 0.344 (0.020) |
| | Low | 0.339 (0.022) | 0.336 (0.022) | 0.366 (0.023) | 0.372 (0.022) |
| | Medium | 0.553 (0.012) | 0.553 (0.012) | 0.648 (0.010) | 0.646 (0.015) |
| | High | 0.565 (0.011) | 0.565 (0.011) | 0.739 (0.011) | 0.742 (0.020) |

*Accuracy was defined as a correlation between predicted phenotype and simulated 'true' phenotype values (TPV). Average of 100 replicates per scenario is reported and the SE were calculated as the standard deviation of the 100 phenotypic accuracies for each scenario.

**Table 5**. Mean accuracy* (SE) of predicted phenotypic values (including residual effect in the prediction) of animals with records using the direct (DP) and indirect (IP) approaches when PLS regression coefficients based on true breeding values ($\beta_g$) and true phenotypic values ($\beta_p$) were used

| Correlation | | DP | | IP | |
|---|---|---|---|---|---|
| Genetic | Residual | $\beta_p$ | $\beta_g$ | $\beta_p$ | $\beta_g$ |
| Low: | Zero | 0.032 (0.042) | 0.069 (0.035) | 0.048 (0.010) | 0.048 (0.010) |
| | Low | 0.160 (0.014) | 0.176 (0.014) | 0.160 (0.015) | 0.160 (0.015) |
| | Medium | 0.529 (0.011) | 0.533 (0.009) | 0.527 (0.012) | 0.527 (0.012) |
| | High | 0.677 (0.008) | 0.680 (0.007) | 0.675 (0.009) | 0.675 (0.009) |
| Medium: | Zero | 0.235 (0.050) | 0.263 (0.025) | 0.203 (0.015) | 0.203 (0.015) |
| | Low | 0.225 (0.022) | 0.303 (0.021) | 0.268 (0.014) | 0.268 (0.014) |
| | Medium | 0.681 (0.008) | 0.679 (0.008) | 0.684 (0.007) | 0.684 (0.007) |
| | High | 0.824 (0.005) | 0.818 (0.005) | 0.827 (0.005) | 0.827 (0.005) |
| High: | Zero | 0.327 (0.031) | 0.344 (0.020) | 0.268 (0.015) | 0.268 (0.015) |
| | Low | 0.252 (0.025) | 0.365 (0.020) | 0.311 (0.014) | 0.311 (0.014) |
| | Medium | 0.740 (0.007) | 0.740 (0.007) | 0.746 (0.007) | 0.746 (0.007) |
| | High | 0.882 (0.003) | 0.872 (0.003) | 0.887 (0.003) | 0.887 (0.003) |

*Accuracy was defined as a correlation between predicted phenotype and simulated 'true' phenotype values (TPV). Average of 100 replicates per scenario is reported and the SE were calculated as the standard deviation of the 100 phenotypic accuracies for each scenario.

ISBN: 978-82-575-1462-4
ISSN: 1894-6402

Norwegian University
of Life Sciences

Postboks 5003
NO-1432 Ås, Norway
+47 67 23 00 00
www.nmbu.no