THE UNIVERSITY of EDINBURGH

Edinburgh Research Explorer

Prediction of metabolic clusters in early lactation dairy cows using models based on 2 milk biomarkers

OPEN ACCESS

Prediction of metabolic clusters in early lactation dairy cows using models based on milk biomarkers

SCHOLARONE™
Manuscripts

1   **TITLE:** Prediction of metabolic clusters in early lactation dairy cows using models based on

2   milk biomarkers

3   **FIRST AUTHOR:** J. De Koster

4   **INTERPRETIVE SUMMARY**

5   Early lactating dairy cows were grouped into clusters at 15 and 37 DIM based on k-means

6   clustering of 4 blood metabolites. These metabolic clusters were used to identify (IM)-

7   BALANCED metabolic profiles. Subsequently, phenotypic production parameters was

8   modelled showing significant differences in dry matter intake and energy balance between

9   clusters. Finally, 3 sets of milk biomarkers were compared according to their predictive

10   accuracy for the observed metabolic profile. Accuracy was highest using Fourier transformed

11   mid-infrared spectra and milk metabolites & enzymes. The metabolic profiles can be used as

12   novel trait for genetic selection and identification of (IM)BALANCED early lactating dairy

13   cows.

14

15   METABOLIC CLUSTERING IN DAIRY COWS

16   **Prediction of metabolic clusters in early lactation dairy cows using models based on milk**

17   **biomarkers**

18   **J. De Koster*, M. Salavati†, C. Grelet‡, M. A. Crowe#, E. Matthews#, R. O'Flaherty#, G.**

19   **Opsomer*, L. Foldager‖¤, the GplusE consortium and M. Hostens*[1]**

20   * Department of Reproduction, Obstetrics and Herd Health, Ghent University, Merelbeke,

21   Belgium.

22   † Royal Veterinary College, London, United Kingdom.

23   ‡ Walloon Agricultural Research Center, Valorisation of Agricultural Products Department,

24   Gembloux, Belgium.

25   ‖ Department of Animal Science, Aarhus University, Tjele, Denmark.

26   ¤ Bioinformatics Research Centre, Aarhus University, Aarhus, Denmark.

27   # University College Dublin, Dublin, Ireland.

28   [1] Corresponding author: Miel Hostens, Salisburylaan 133, 9820 Merelbeke, Belgium. Email

29   miel.hostens@ugent.be

30

31

32   **ABSTRACT**

33   The aim of this study was to describe metabolism of early lactation dairy cows by clustering

34   cows based on glucose, insulin like growth factor I (IGF-I), free fatty acid (FFA), and β-

35   hydroxybutyrate (BHB) using the k-means method. Predictive models for metabolic clusters

36   were created and validated using three sets of milk biomarkers (milk metabolites and enzymes,

37   glycans on the immuno-gamma globulin (IgG) fraction of milk, and Fourier transformed mid-

38   infrared spectra (FT-MIR) of milk). Metabolic clusters are used to identify dairy cows with a(n)

39   (im)balanced metabolic profile. Around 14 and 35 days in milk, serum or plasma concentrations

40   of BHB, FFA, glucose and IGF-I, were determined. Cows with a favorable metabolic profile

41   were grouped in BALANCED (n=43) and compared with OTHER$_{BAL}$ (n=64). Cows with an

42   unfavorable metabolic profile were grouped in IMBALANCED (n=19) and compared with

43   OTHER$_{IMBAL}$ (n=88). Glucose and IGF-I were higher in BALANCED compared with

44   OTHER$_{BAL}$. FFA and BHB were lower in BALANCED compared with OTHER$_{BAL}$. Glucose

45   and IGF-I were lower in IMBALANCED compared with OTHER$_{IMBAL}$. FFA and BHB were

46   higher in IMBALANCED. Metabolic clusters were related to production parameters. There was

47   a trend for a higher daily increase in fat and protein corrected milk yield (FPCM) in

48   BALANCED while FPCM of IMBALANCED was higher. Dry matter intake (DMI) and the

49   daily increase in DMI were higher in BALANCED and lower in IMBALANCED. Energy

50   balance was continuously higher in BALANCED and lower in IMBALANCED. Weekly or bi-

51   weekly milk samples were taken and milk metabolites and enzymes (milk glucose, glucose-6-

52   phosphate, BHB, lactate dehydrogenase, N-acetyl-β-D-glucosaminidase, isocitrate), IgG

53   glycans (19 peaks) and FT-MIR (1,060 wavelengths reduced to 15 principal components)

54   determined. Milk biomarkers with or without additional cow information (DIM, parity, milk

55   yield features) were used to create predictive models for the metabolic clusters. Accuracy for

56   prediction of BALANCED (80%) and IMBALANCED (88%) was highest using milk

57   metabolites and enzymes combined with DIM and parity. The results and models of the present

58   study are part of the GplusE project and identify novel milk based phenotypes that may be used

59   as predictors for metabolic and performance traits in early lactation dairy cows.

60   **Key words: metabolic clustering, dairy cows, prediction, milk biomarkers**

61   **INTRODUCTION**

62   Physiological adaptations of dairy cows in the transition period are well described (Bell and

63   Bauman, 1997, Drackley et al., 2005). As parturition approaches, growth hormone (**GH**)

64   concentrations start to increase while insulin and insulin like growth factor I (**IGF-I**)

65   concentrations decrease (Lucy et al., 2009). Growth hormone enhances milk production

66   through partitioning of nutrients towards the mammary gland (Lucy, 2008), lipolysis increases,

67   lipogenesis is almost completely downregulated and hepatic gluconeogenesis is stimulated

68   (Bauman, 1999, Lucy, 2008). Growth hormone induces the production of IGF-I by the liver

69   which inhibits GH secretion by the hypophysis (Lucy, 2008). Homeorhetic adaptation

70   mechanisms in the periparturient period decrease the expression of GH receptors by the liver

71   which downregulate hepatic IGF-I production, known as uncoupling of the GH-IGF-I axis

72   (Lucy, 2008). Decreased IGF-I concentrations in the periparturient period relieves negative

73   feedback on pituitary GH production (Lucy, 2008). As lactation progresses, recoupling of the

74   GH-IGF-I axis is characterized by repleted IGF-I levels and effectuated by slightly elevated

75   insulin levels (Lucy, 2008).

76   Homeorhetic mechanisms commonly observed during the transition period supports the

77   metabolic prioritization of the lactating mammary gland leading to a sharp increase in milk

78   production (Bauman and Currie, 1980). The increase in DMI in postpartum cows lags behind

79   the increased milk production which results in a period of negative energy balance (**EBAL**)

80   (Grummer et al., 2004). Severe negative EBAL is a risk factor for metabolic, infectious and

81   reproductive disorders (Moyes et al., 2013). Metabolic indicators for the severity of the negative

82   EBAL are used to detect individual cows, which are unable to cope with the altered metabolic

83   challenge of lactation. Elevated free fatty acids (**FFA**) and BHB, and decreased glucose and

84   IGF-I are reported as biomarkers of metabolic imbalanced cows which are more at risk for an

85   unsuccessful transition from the dry period to lactation (Ingvartsen et al., 2003, Puppel and

86   Kuczynska, 2016). Cut-off values for FFA and BHB have been determined to predict negative

87   health consequences in the periparturient period (McArt et al., 2013, Ospina et al., 2013).

88   However, the metabolic profile of individual cows within a herd shows a high level of

89   variability (Ingvartsen et al., 2003). Especially the correlation between FFA and BHB is

90   difficult to explain. While some cows with normal BHB levels have high FFA levels, vice versa

91   other cows with normal FFA levels have high BHB levels, demonstrating unexplained inter-

92   individual variability within one group of cows (Ospina et al., 2013, McCarthy et al., 2015).

93   This inter-individual variability represents inter-individual ability of cows to adapt to the altered

94   metabolic challenge of lactation and demonstrates the limitations of monitoring metabolism

95   based on a single metabolic indicator (Ingvartsen et al., 2003, Bjerre-Harpoth et al., 2012,

96   Moyes et al., 2013). Moreover, monitoring early lactation metabolism requires early

97   identifiable biomarkers. Due to better accessibility and the ease of automated repeat sampling

98   owing to recently developed 'in-line' sampling and analytical technologies, milk is the

99   preferred medium in which biomarkers can be measured (Ingvartsen and Friggens, 2005,

100   Nielsen et al., 2005a, Egger-Danner et al., 2015). It is highly probable that certain glycan

101   structures (Zhao and Keating, 2007), milk metabolites and enzymes (Weekes et al., 1983,

102   Wallace and Matthews, 2002, Bjerre-Harpoth et al., 2012) or mid-infra-red spectra (Voelker

103   and Allen, 2003, Maury et al., 2007, Soyeurt et al., 2011) of bovine milk can serve as

104   biomarkers to monitor early lactation metabolism and performance but comparable prediction

105   methodologies are lacking.

106   The aim of this study was to combine concentrations of blood metabolites to describe

107   metabolism of dairy cows in early and peak lactation by clustering cows based on

108   concentrations of glucose, IGF-I, FFA, and BHB. Metabolic clusters are used to identify dairy

109   cows with a(n) (im)balanced metabolic profile. Furthermore, the relationship between

110   metabolic clusters and production parameters (DMI, fat-protein-corrected milk production

111   (**FPCM**), <mark>BW</mark>, EBAL and <mark>BCS</mark>) are determined. Finally, three sets of milk biomarkers (milk

112   metabolites and enzymes (<mark>milk biomarker 1, **MBM1**</mark>), glycans on the <mark>immunoglobulin-gamma</mark>

113   (**IgG**) fraction of milk (<mark>milk biomarker 2, **MBM2**</mark>), and Fourier transformed mid-infrared

114   spectra (**FT-MIR**) of milk (<mark>milk biomarker 3, **MBM3**</mark>)) are used to create and validate

115   predictive models for the different metabolic clusters.

MATERIALS AND METHODS

117   <mark>The experiments were carried out in accordance with the standards recommended by the EU</mark>

118   <mark>Directive 2010/63/EU for animal experiments.</mark> Detailed description of the experiments,

119   laboratory analysis of blood and milk, production data etc. are given in Foldager et al. (2018)

120   and described in brief below.

Animals and Sampling

122   Samples and data between calving and 50 days post calving (1-50 <mark>DIM</mark>) were obtained from

123   130 Holstein Friesian cows (parity 2: n = 42; parity 3: n = 51; parity ≥ 4 : n = 37) in four research

124   herds: AU (Aarhus University, Denmark), UCD (UCD Lyons Research farm, University

125   College Dublin, Ireland), AFBI (Agri-Food and Biosciences Institute, Northern Ireland, UK),

126   and FBN (Leibniz Institute for Farm Animal Biology, Germany). <mark>Cows were milked twice</mark>

127   <mark>daily.</mark>

Analyses of Glucose, IGF-I, FFA and BHB

129   Blood samples were taken around <mark>14 DIM (15 ± 0.1 DIM, **D14**)</mark> and 35 DIM (37 ± 0.1 DIM,

130   <mark>**D35**</mark>) in serum and heparin tubes by jugular or coccygeal venipuncture. Blood plasma glucose

131   concentrations were determined using an enzymatic method (ADVIA 1800 Clinical Chemistry

132   System, Siemens Diagnostics). Plasma FFA concentrations were determined using an

133   enzymatic method (NEFA C ACS-ACOD assay method, Wako). Plasma BHB concentrations

134   were determined by measuring absorbance at 340 nm due to the production of NADH at

135 alkaline pH in the presence of BHB dehydrogenase. Serum concentrations of IGF-I were

136 determined using a radioimmunoassay following acid–ethanol extraction using the method

137 previously described by Beltman et al. (2010).

138 ***Analyses of Milk Metabolites and Enzymes, Milk Glycans and FT-MIR***

139   ***Metabolites and Enzymes.*** Weekly, two AM milk samples were collected starting from

140 the first week in milk until 50 DIM. At each day of sampling, two separate samples of

141 approximately 8 mL were obtained and stored at -18°C. Fluorometric assays were used to

142 determine milk glucose and glucose-6-phosphate (Larsen, 2015), BHB (Larsen and Nielsen,

143 2005), lactate dehydrogenase (Larsen, 2005), N-acetyl-β-D-glucosaminidase (Larsen et al.,

144 2010) and isocitrate (Larsen, 2014). Urea was determined by spectrophotometry (Nielsen et al.,

145 2005b). Finally, six milk metabolites and enzymes were available as the first set of milk

146 biomarker (MBM1).

147   ***IgG Glycans.*** Weekly, two AM milk samples were collected starting from the first week

148 in milk until 50 DIM. Milk samples were centrifuged at 4,000 g for 30 min at 4°C. Five mL of

149 whole milk internatant was recovered and frozen at -20°C. In duplicate, 300 µL of thawed

150 sample was filtered through a 1 µm glass fiber filter plate (Acroprep, VWR International Ltd,

151 Radnor, PA, USA) at 3,000 g for 10 min and collected in a 96-well greiner plate (Cruinn

152 Diagnostics, Dublin, Ireland). All processes relating to IgG purification and IgG-glycan release

153 were carried out on a Hamilton Robotics StarLet liquid-handling platform (Hamilton Robotics,

154 Reno, NV, USA) using a protocol adapted from Stöckmann et al. (2013). The N-glycans were

155 separated on a Waters Acquity UPLC instrument (Waters, Milford, MA, USA) and analyzed

156 using Empower V3 (Waters 2010, Milford).

157 IgG was captured by passing 290 µL of filtered sample through a Protein G matrix (Phytip G;

158 200 µl column; 20 µl resin bed. Phynexus, San Jose, CA, USA). The column tips were washed

159   five times with a 0.1 M sodium phosphate binding buffer (pH 7.4), eluted with a 0.2 M glycine-

160   hydrochloride buffer (pH 2.5) and neutralized with 1 M Tris-hydrochloride buffer (pH 9.0). The

161   purified protein was enriched by pooling from two identical plates.

162   One hundred μL of the enriched IgG was transferred to a 96-well ultrafiltration plate (Acroprep,

163   Omega membrane, 10 kDa, VWR International Ltd) and centrifuged at 3,000 g for 30 min at

164   room temperature. Fifty μL of a dithiothreitol denaturing buffer was added to each well,

165   mixed and left at room temperature for 10 min. The sample was then transferred to a 96-well

166   thermobalanced denaturation plate (Armadillo, High Performance 96-well PCR Plate, Thermo

167   Scientific, Waltham, MS, US) incubated at 95°C for 10 min and cooled at room temperature for

168   20 min. The denatured sample was transferred back onto the ultrafiltration plate and incubated

169   in 20 μL of 1M Iodoacetamide buffer for 10 min. The sample was washed with 20 μL of 25

170   mM sodium bicarbonate and filtered. 0.4 μL of PNGase F (2.5 U/mL) (New England Biolabs,

171   Ipswich, UK) was added to each well and incubated for 30 min at 38°C with agitation. The

172   released N-glycans were recovered by centrifugation at 3,000 g through a 10 kDa filter for 10

173   min.

174   Eight μL of the N-glycan sample was incubated with 12 μL of the fluorescent tag 6-

175   aminoquinolyl-N-hydroxysuccinimidyl carbamate (3 mg/mL MeCN) at room temperature.

176   Sixteen μL of each sample was separated using hydrophilic separation chromatography (Glycan

177   BEH Amide 130Å Column, Waters) and 19 peaks were manually identified and integrated.

178   Each peak's percentage of the total area under the 19 peaks was used as the IgG glycan measure

179   for the statistical analyses. This set of 19 was used as second set of milk biomarker (MBM2)

180   ***Fourier Transformed Mid-Infra-Red spectra.*** Twice weekly, AM and PM milk

181   samples were collected starting from the first week in milk until 50 DIM and preserved with

182   bronopol 0.02% and stored at 4°C. Analyses were done locally on FT2 and FT6000

183   spectrometers (Foss, Hillerød, Denmark) or at CRA-W (Belgium) by a Standard Lactoscope

184   FT-MIR automatic (Delta Instruments, Drachten, The Netherlands). The AM and PM FT-MIR

185   spectra were combined into a daily spectrum by a weighted average taking into account the AM

186   and PM milk yields. The FT-MIR spectra of the different instruments were standardized and

187   merged into a common dataset following the procedure described in Grelet et al. (2015).

188   Finally, absorbance values at 1,060 wavenumbers were available as third set of milk biomarker

189   (MBM3). Samples were analyzed locally for fat, protein and lactose content by Fourier

190   transform infrared spectroscopy with FT2 and FT6000 spectrometers (Foss, Hillerød,

191   Denmark) or a Standard Lactoscope FT-MIR automatic (Delta Instruments, Drachten, The

192   Netherlands) and using the predictive models provided by the manufacturers.

193   ***FPCM, DMI, EBAL and BW Measurements***

194   All cows were milked twice daily and daily yields were recorded. Milk samples were collected

195   twice weekly until 50 DIM and analyzed for composition of protein, fat and lactose by mid-

196   infrared analysis, and for somatic cell count by flow cytometry. Bronopol (0.02%, Sigma-

197   Aldrich) was added as a preservative to all samples. Milk weighted average for milk

198   components were used in the subsequent analysis. Fat-protein-corrected milk production

199   (FPCM) was calculated as [(0.337 + 0.116 x milk fat % + 0.06 x milk protein %) x kg of milk]

200   (CVB, 2007).

201   Body condition score was recorded at the moment of blood sampling using a 1 to 5 scale, with

202   0.25-point increments (Edmonson et al., 1989). Live weights were recorded twice weekly using

203   weight scales until 50 DIM. Daily individual DMI were recorded using electronic feeding

204   system (Insentec, Markneesse, Netherlands). Over the course of the sampling period, weekly

205   ration samples were collected, dried and shipped for NE_L analysis in a single run (Cumberland

206   Valley Agricultural Services, Maryland). The EBAL was calculated by calculating the daily

207   difference between energy input and output. The energy input was calculated by multiplying

208   the weekly NE_L density of the ration with the daily DMI of the animal. The energy output was

209 calculated according to the National Research Council calculation correcting for NE_M using the

210 last recorded BW of the animal and most recent fat, protein and lactose content together with

211 the daily milk yield for the NE_L (NRC, 2001).

Statistical analyses

213 *Metabolic Clustering.* Due to missing observations, technical and logistical limitations,

214 only 107 cows were included in the final statistical analyses. The k-means clustering method

215 was used to group cows based on the log-transformed and standardized concentrations (mean

216 = 0 and STDEV = 1) of glucose, IGF-I, FFA and BHB on D14 and D35. Sum of squares plots

217 were used to determine the optimal number of clusters. Three clusters were created in both

218 periods: A (n = 51), B (n = 51), and C (n = 5) on D14 and D (n = 58), E (n = 25) and F (n = 24)

219 on D35. Pairwise comparisons of glucose, IGF-I, FFA and BHB between A, B, and C on D14

220 and D, E and F on D35 were done using an ANOVA. The clustering of the individual cows on

221 D14 and D35 was compared. Cows in cluster A on D14 and cluster D on D35 were considered

222 metabolically balanced cows (n = 43, **BALANCED**). Cows not in cluster A on D14 and cluster

223 D D35 were grouped together (n = 64, **OTHER$_{BAL}$**). Equally, cows in cluster B on D14 and

224 cluster F on D35 (n = 14) together with cows in cluster C on D14 and cluster F on D35 (n = 5)

225 were considered metabolically imbalanced cows (n = 19, **IMBALANCED**). Cows not in these

226 clusters on D14 or D35 were grouped together (n = 88, **OTHER$_{IMBAL}$**).

227 *Metabolic and Production Performance of Metabolic Clusters.* Pairwise comparisons

228 between BALANCED vs OTHER$_{BAL}$ and IMBALANCED vs OTHER$_{IMBAL}$ were done using

229 a linear mixed effect model after log transformation of the metabolite and hormone

230 concentrations with period (D14 and D35) as repeated observation within the random factor

231 cow. Linear mixed effect models were constructed for the FPCM, DMI, EBAL and BW of the

232 cows in the metabolic cluster (BALANCED vs OTHER$_{BAL}$ and IMBALANCED vs

233 OTHER$_{IMBAL}$) with day post calving as repeated observation within the random factor cow. All

234   pairwise comparisons were done using the Tukey's post hoc test. Residuals of the models were

235   checked and found to be normally distributed. Interaction effects were removed from the model

236   if non-significant ($P < 0.05$). Significance and tendency were declared at $P < 0.05$ and $0.05 \leq$

237     $P < 0.10$, respectively. Results are presented as LSMEANS ± SEM unless otherwise stated.

238   **_Prediction of Metabolic Clusters_**

239   **_Feature Preparation of FT-MIR._** Due the nature of the FT-MIR data, a dimension

240   reduction step was needed to reduce the high number of variables (n = 1,060) in contrast with

241   glycan (n = 19) or blood metabolites (n = 6). First, wavenumbers were removed known to be

242   non-informative due to the water component in milk (Grelet et al., 2016). Next, FT-MIR spectra

243   were reduced using a principal component analysis. Exploration of the variance plot revealed

244   15 principal components (**PC**) that contributed most to the entire variation in FT-MIR data.

245   After the filter and reduction step, 1,060 wavenumbers were reduced to 15 PC which were used

246   in the remainder of the analysis.

247   **_Biomarker Preparation._** Each set of biomarkers was subsequently split into 3. One set

248   of the biomarkers without additional cow information, one set of the biomarkers including the

249   DIM at sampling and the parity of the animal, and one set of the biomarkers including the DIM

250   at sampling, the parity of the animal, and milk yield features (number of milkings, the

251   minimum, maximum, mean, stdev, sum of milk yield in the period up to the sampling).

252   **_Random Sampler._** Three sets of milk biomarkers (MBM1, MBM2, and MBM3) were

253   available as predictors to classify the animals according to their metabolic cluster. However,

254   for each of the biomarkers and according to the research protocol, multiple samples at varying

255   DIM were available. A random sampler was created to select one sample per animal from 1 to

256   50 DIM. As such, the sampler works as if an official milk recording organization entered each

257   of the participating research herds and sampled all cows at random stages in lactation on a given

258   day.

259   *Random Forest.* In the next step, 10 models were created using the aforementioned

260   random milk recording sampler as dataset and a random forest (**RF**) classifier to predict

261   BALANCED and IMBALANCED cows in each run. Each RF used 1,000 trees and a maximum

262   depth of 5. One third of the features were selected to be used as candidates for splitting at each

263   tree node. A separate training and test dataset was created for BALANCED and

264   IMBALANCED cows by randomly splitting the animals by a ratio of 75/25. All features were

265   standardized before entering the model to have a mean of 0 and standard deviation of 1. The

266   accuracy (% of cows with a correctly predicted metabolic cluster) was selected as the evaluation

267   metric to rank the performance of each of the models. The final minimum, maximum and

268   average accuracy of the models were reported. A schematic overview of the different steps in

269   the data preparation, model creation and validation and statistical analysis is given in Figure 1.

270   **RESULTS**

271   *Metabolic Clusters*

272   Metabolic clustering based on concentrations of glucose, IGF-I, FFA and BHB using the k-

273   means method resulted in 3 distinct metabolic clusters on D14 and D35. On D14, glucose

274   concentrations were highest in A, intermediate in B, and lowest in C, IGF-I concentrations were

275   highest in A, intermediate in C and lowest in B, while FFA and BHB concentrations were lowest

276   in A, intermediate in B and highest in C (Table 1). On D35, glucose and IGF-I concentrations

277   were higher in D compared with E and F, FFA concentrations were higher in F compared with

278   D and E, and BHB concentrations were lowest in D, intermediate in E and highest in F (Table

279   1).

280   *Production and Metabolic Performance in BALANCED vs OTHER$_{BAL}$*

281   By comparing the metabolic clusters of individual cows on D14 and D35, 43 cows were

282   identified to have a balanced metabolic profile (cluster A on D14 and cluster D on D35, Table

283   2) during the postpartum period and were grouped together in BALANCED. All other cows (n

284   = 64) were grouped in OTHER$_{BAL}$. Glucose and IGF-I concentrations increased towards D35

285   in both groups and were higher in BALANCED compared with OTHER$_{BAL}$ (Table 3). FFA

286   concentrations decreased towards D35 in both groups and were lower in BALANCED

287   compared with OTHER$_{BAL}$ (Table 3). BHB concentrations were lower in BALANCED

288   compared with OTHER$_{BAL}$ but did not change in between periods (Table 3). There was no

289   difference in FPCM yield in BALANCED compared with OTHER$_{BAL}$ (Figure 2, Table 4). DMI

290   and the daily increase in DMI were higher in BALANCED (Figure 2, Table 4). The decrease

291   in BW was more pronounced in OTHER$_{BAL}$. Energy balance was continuously higher in

292   BALANCED compared with OTHER$_{BAL}$ (Figure 2, Table 4). The BCS of OTHER$_{BAL}$

293   decreased from D14 to D35, but not in BALANCED (Table 3).

Production and Metabolic Performance in IMBALANCED vs OTHER$_{IMBAL}$

295   By comparing the metabolic clusters of individual cows on D14 and D35, 19 cows were

296   identified to have an imbalanced metabolic profile (5 cows in cluster C on D14 and cluster F

297   on D35, and 14 cows in cluster B on D14 and cluster F on D35, Table 2) during the postpartum

298   period and were grouped together in IMBALANCED. All other cows (n = 88) were grouped in

299   OTHER$_{IMBAL}$. Glucose concentrations were lower in IMBALANCED cows and were not

300   different on D14 compared with D35 (Table 5). IGF-I concentrations were lower in

301   IMBALANCED compared with OTHER$_{IMBAL}$ and increased in both groups on D35 (Table 5).

302   FFA and BHB concentrations were higher in IMBALANCED in both periods. While FFA

303   concentrations decreased on D35 compared with D14, BHB concentrations were not different

304   in between periods (Table 5). FPCM yield of IMBALANCED cows was higher compared with

305   OTHER$_{IMBAL}$ (Figure 3, Table 6). DMI and the daily increase in DMI were lower in

306 IMBALANCED (Figure 3, Table 6). The decrease in BW was more pronounced in

307 IMBALANCED. Energy balance was continuously lower in IMBALANCED (Figure 3, Table

308 6). IMBALANCED cows had a higher BCS and the decrease in BCS towards D35 was more

309 pronounced when compared with OTHER$_{IMBAL}$ (Table 5).

310 ***Prediction of Metabolic Clusters***

311 A total of 5,400 cross-validation models were summarized in Table 7 and 8. The average

312 prediction accuracy of BALANCED cows was highest (76%) for FT-MIR spectra. The average

313 prediction of BALANCED cows outperformed the prediction of the separate metabolic clusters

314 on D14 and D35. Overall, the highest accuracy was found at 80% accuracy using milk

315 metabolites and enzymes combined with DIM and parity followed by FT-MIR spectra

316 combined with DIM, parity and milk yield features (79% accuracy). The lowest CV in accuracy

317 between the different runs was found in FT-MIR Spectra predictions with DIM and parity.

318 The average prediction accuracy of IMBALANCED cows was highest (81%) for metabolites

319 and enzymes with DIM and parity. In contrast to the BALANCED cows, the prediction of

320 IMBALANCED cows outperformed the prediction of the separate metabolic clusters on D14

321 and D35 except for milk metabolites and enzymes combined with DIM, parity and milk yield

322 features, IgG glycans combined with DIM, parity and milk yield features and predictions using

323 FT-MIR spectra. The highest accuracy was found at 88% accuracy using Metabolites and

324 enzymes in combination with DIM and parity or FT-MIR Spectra (87% accuracy).

325   **DISCUSSION**

326 This research is part of an EU funded project "Genotype plus Environment" (GplusE) aiming

327 to identify novel milk based phenotypes that may be used as predictors for health traits in dairy

328 cows (Crowe et al., 2018). The present study describes how milk based phenotypes are linked

329   with clustering of cows in metabolic balance and imbalance, and the relationship with

330   production performance parameters.

331   Different metabolites and hormones can be used to characterize the metabolism of cows in early

332   lactation. The altered metabolic environment in the periparturient period increases the rate of

333   adipose tissue lipolysis. A certain level of FFA is considered normal and necessary. However,

334   when lipolysis is excessive, FFA concentrations increase and this is associated with a

335   detrimental impact on immunity, metabolism and milk production (Roche et al., 2013,

336   Contreras et al., 2018). A part of the circulating FFA are converted into ketone bodies by the

337   liver. Excessive production of ketone bodies may lead to (sub)clinical ketosis (Roche et al.,

338   2013). Glucose concentrations are tightly controlled by homeostatic mechanisms (De Koster

339   and Opsomer, 2013). Cows suffering from severe negative EBAL have lower glucose

340   concentrations (Wathes et al., 2011). The usefulness of glucose as single indicator of metabolic

341   imbalance has been questioned (Mulligan et al., 2006). However, studies by Bjerre-Harpoth et

342   al. (2012) and Moyes et al. (2013) identified glucose as an important metabolite to be included

343   in an index of metabolic imbalance together with FFA and BHB. Late pregnancy and early

344   lactation are marked by a decrease in the concentration of IGF-I. The nadir IGF-I concentration

345   is reached in the first week after calving (Butler et al., 2003, Radcliff et al., 2003) and after the

346   second week, IGF-I concentrations are markedly influenced by the energy status of the cows

347   (Fenwick et al., 2008, Wathes et al., 2011). Animals with severe negative EBAL have lower

348   IGF-I concentrations (Fenwick et al., 2008) and IGF-I has been suggested to be an indicator of

349   the nutritional status (Cohick, 1998, Zulu et al., 2002).

350   Metabolic imbalance is defined as 'a condition where the regulating mechanisms are

351   insufficient for the animals to function optimally leading to a high risk of a complex of

352   digestive, metabolic and infectious problems' (Ingvartsen, 2006). Metabolic clustering of dairy

353   cows based on concomitant changes in the concentration of different metabolites improves the

354   identification of metabolically balanced cows compared with the use of a single indicator

355   (Bjerre-Harpoth et al., 2012, Moyes et al., 2013, Grelet et al., 2018). Metabolic clustering using

356   the k-means method in the present study resulted in three distinct metabolic groups in early and

357   peak lactation. Based on the comparison of the glucose, IGF-I, FFA and BHB concentrations

358   between the metabolic groups, group A and D show a balanced metabolic profile on D14 and

359   D35, respectively (high glucose, high IGF-I, low FFA and low BHB). While the metabolic

360   profile of group B and C on D14 and group F on D35 show an imbalanced metabolic profile

361   (low glucose, low IGF-I, high FFA or high BHB). Dairy cows in cluster A on D14 and cluster

362   D on D35, were identified as metabolically BALANCED throughout the study and had higher

363   glucose and IGF-I, and lower FFA and BHB compared with other cows in both early and peak

364   lactation. Dairy cows in cluster C on D14 and cluster F on D35 together with cows in cluster B

365   on D14 and cluster F on D35, were identified as metabolically IMBALANCED throughout the

366   study and had lower glucose and IGF-I, and higher FFA and BHB compared with other cows

367   on D14 and D35.

368   The different metabolic clusters were characterized by differences in FPCM, DMI, BW, BCS

369   and EBAL in the postpartum period. The relationship between metabolic clusters and FPCM

370   was weak. The FPCM yield was not different in BALANCED cows. Metabolically

371   IMBALANCED cows had higher FPCM yield. The weak relationship between milk yield,

372   metabolic balance and disorders has been described before by Ingvartsen et al. (2003),

373   Ingvartsen (2006) and Bjerre-Harpoth et al. (2012). Ingvartsen (2006) stated that it is not milk

374   yield as such that is the cause of metabolic imbalance or disorders but the individual cow's

375   inability to cope with the metabolic challenges of early lactation. Dry matter intake was

376   consistently higher in metabolically BALANCED cows and lower in IMBALANCED cows.

377   Daily increase in DMI was higher in cows with a favorable metabolic profile (BALANCED

378   and OTHER$_{IMBAL}$). The relationship between DMI and metabolism may be explained by the

379   fact that certain metabolites (FFA) may regulate feed intake in ruminants by the hepatic

380   oxidation of these metabolites thereby causing a satiety signal and depressing feed intake

381   (Ingvartsen and Andersen, 2000, Allen et al., 2009). Energy balance is mainly influenced by

382   energy intake and less by milk production (Santos et al., 2010). The improved nutritional status

383   in BALANCED and OTHER$_{IMBAL}$ cows has a positive effect on the EBAL of the animals.

384   While IMBALANCED cows were in severe negative EBAL throughout the study. Body weight

385   is affected by different factors in postpartum cows: frame size, DMI, stage of lactation and

386   EBAL. As such, BW is not a good predictor for the severity of the negative EBAL or the

387   mobilization of energy (Schröder and Staufenbiel, 2006). However, automated daily

388   measurement of BW has been used to assess the energy status of dairy cows (van Straten et al.,

389   2008, Thorup et al., 2012). The more pronounced decrease in BW in OTHER$_{BAL}$ and

390   IMBALANCED compared with BALANCED and OTHER$_{IMBAL}$, respectively, can be

391   explained by a higher degree of mobilization of energy reserves due to the severe negative

392   EBAL or a lower gut fill due to the decreased DMI or a combination of both factors. BCS is a

393   subjective indicator for the mobilization of subcutaneous adipose tissue. Especially the

394   postpartum decrease in BCS is associated with metabolic and infectious disorders (Roche et al.,

395   2013). Cows with an unfavorable metabolic profile (OTHER$_{BAL}$ and IMBALANCED) had a

396   more pronounced decrease in BCS from D14 to D35, indicative for a higher degree of body fat

397   mobilization.

398   The current study furthermore focused on comparing multiple biomarkers to predict the

399   metabolic clusters as novel phenotypic trait in early lactating dairy cows. To our knowledge,

400   this is the first study positioning the different sets of milk biomarkers relatively across one

401   another according to their predictive accuracy. The current scope of the study focused on

402   comparing different milk biomarkers rather than the individual fine tuning of the individual

403   biomarkers to increase the predictive accuracy as described by Grelet et al. (2018) for FT-MIR

404 and metabolic clustering. From our results, it can be noted that within each set of biomarkers,

405 adding DIM, parity and milk yield features did not improve the predictive accuracy much.

406 Hence, it is debatable whether collecting such information is worth the effort compared to the

407 relatively small gain in predictive accuracy. Overall, we were able to predict BALANCED and

408 IMBALANCED animals in early lactation with varying prediction accuracy across the 3 sets

409 of biomarkers. Prediction of IMBALANCED was more accurate across all milk biomarkers

410 compared to BALANCED where FT-MIR outperformed the other milk biomarkers. Further

411 efforts are made within the GplusE project for industry wide application of the metabolic

412 clustering technique. External validation of the BALANCED cows can lead to establishment

413 of a novel phenotypic trait for genetic selection as suggested by Egger-Danner et al. (2015) and

414 Crowe et al. (2018). External validation of the IMBALANCED cows can help identifying cows

415 for specific management strategies such as elective propylenic glycol treatment versus group

416 treatment as proposed by others (Lomander et al., 2012, Jenkins et al., 2015).

417   **CONCLUSION**

418 The k-means clustering of blood metabolites was found to effectively identify BALANCED

419 and IMBALANCED cows across the participating countries within the GplusE project.

420 Furthermore, production parameters revealed marked differences in dry matter intake and

421 energy balance, underlining the phenotypic validity of metabolic clusters. Finally, prediction

422 using both FT-MIR or milk metabolites and enzymes allows implementation of metabolic

423 clusters across larger cow numbers as novel trait for genetic selection or identification of

424 imbalanced early lactating dairy cows.

425   **ACKNOWLEDGMENTS**

428   The views expressed in this publication are the sole responsibility of the authors and do not

429   necessarily reflect the views of the European Commission.

430   List of authors within the GplusE consortium: Niamh McLoughlin, Alan Fahey, Elizabeth

431   Matthews, Andreia Santoro, Colin Byrne, Pauline Rudd, Roisin O'Flaherty, Sinead Hallinan,

432   Claire Wathes, Zhangrui Cheng, Ali Fouladi, Geoff Pollott, Dirk Werling, Beatriz Sanz

433   Bernardo, Alistair Wylie, Matt Bell, Mieke Vaneetvelde, Kristof Hermans, Geert Opsomer,

434   Sander Moerman, Jenne De Koster, Hannes Bogaert, Jan Vandepitte, Leila Vandevelde, Bonny

435   Vanranst, Johanna Hoglund, Susanne Dahl, Soren Ostergaard, Janne Rothmann, Mogens

436   Krogh, Else Meyer, Charlotte Gaillard, Jehan Ettema, Tine Rousing, Federica Signorelli,

437   Francesco Napolitano, Bianca Moioli, Alessandra Crisà, Luca Buttazzoni, Jennifer McClure,

438   Daragh Matthews, Francis Kearney, Andrew Cromie, Matt McClure, Shujun Zhang, Xing

439   Chen, Huanchun Chen, Junlong Zhao, Liguo Yang, Guohua Hua, Chen Tan, Guiqiang Wang,

440   Michel Bonneau, Andrea Pompozzi, Armin Pearn, Arnold Evertson, Linda Kosten, Anders

441   Fogh, Thomas Andersen, Matthew Lucey, Chris Elsik, Gavin Conant, Jerry Taylor, Nicolas

442   Gengler, Michel Georges, Frédéric Colinet, Marilou Ramos Pamplona, Hedi Hammami,

443   Catherine Bastin, Haruko Takeda, Aurelie Laine, Anne-Sophie Van Laere, Martin Schulze,

444   Sergio Palma Vera.

445

REFERENCES

446

447 Allen, M. S., B. J. Bradford, and M. Oba. 2009. Board-Invited Review: The hepatic oxidation

448 theory of the control of feed intake and its application to ruminants. J. Anim. Sci. 87(10):3317-

449 3334.

450 Bauman, D. E. 1999. Bovine somatotropin and lactation: from basic science to commercial

451 application. Domest. Anim. Endocrinol. 17(2-3):101-116.

452 Bauman, D. E. and W. B. Currie. 1980. Partitioning of nutrients during pregnancy and lactation

453 - a review of mechanisms involving homeostasis and homeorhesis. J. Dairy Sci. 63(9):1514-

454 1529.

455 Bell, A. W. and D. E. Bauman. 1997. Adaptations of glucose metabolism during pregnancy and

456 lactation. J. Mammary Gland Biol. Neoplasia 2(3):265-278.

457 Beltman, M. E., N. Forde, P. Furney, F. Carter, J. F. Roche, P. Lonergan, and M. A. Crowe.

458 2010. Characterisation of endometrial gene expression and metabolic parameters in beef heifers

459 yielding viable or non-viable embryos on Day 7 after insemination. Reprod. Fertil. Dev.

460 22(6):987-999.

461 Bjerre-Harpoth, V., N. C. Friggens, V. M. Thorup, T. Larsen, B. M. Damgaard, K. L.

462 Ingvartsen, and K. M. Moyes. 2012. Metabolic and production profiles of dairy cows in

463 response to decreased nutrient density to increase physiological imbalance at different stages

464 of lactation. J. Dairy Sci. 95(5):2362-2380.

465 Butler, S. T., A. L. Marr, S. H. Pelton, R. P. Radcliff, M. C. Lucy, and W. R. Butler. 2003.

466 Insulin restores GH responsiveness during lactation-induced negative energy balance in dairy

467 cattle: effects on expression of IGF-I and GH receptor 1A. J. Endocrinol. 176(2):205-217.

468 Cohick, W. S. 1998. Role of the insulin-like growth factors and their binding proteins in

469 lactation. J. Dairy Sci. 81(6):1769-1777.

470   Contreras, G. A., C. Strieder-Barboza, and J. De Koster. 2018. Symposium review: Modulating

471   adipose tissue lipolysis and remodeling to improve immune function during the transition

472   period and early lactation of dairy cows. J. Dairy Sci. 101(3):2737-2752.

473   Crowe, M. A., M. Hostens, and G. Opsomer. 2018. Reproductive management in dairy cows -

474   the future. Ir Vet J 71:1.

475   CVB. 2007. Tabellenboek veevoeding 2007. Centraal Veevoederbureau Lelystad, the

476   Netherlands.

477   De Koster, J. D. and G. Opsomer. 2013. Insulin resistance in dairy cows. Vet. Clin. North Am.

478   Food Anim. Pract. 29(2):299-322.

479   Drackley, J. K., H. M. Dann, G. N. Douglas, N. A. J. Guretzky, N. B. Litherland, J. P.

480   Underwood, and J. J. Loor. 2005. Physiological and pathological adaptations in dairy cows that

481   may increase susceptibility to periparturient diseases and disorders. Ital. J. Anim. Sci. 4(4):323-

482   344.

483   Edmonson, A. J., I. J. Lean, L. D. Weaver, T. Farver, and G. Webster. 1989. A Body Condition

484   Scoring Chart for Holstein Dairy Cows. J. Dairy Sci. 72(1):68-78.

485   Egger-Danner, C., J. B. Cole, J. E. Pryce, N. Gengler, B. Heringstad, A. Bradley, and K. F.

486   Stock. 2015. Invited review: overview of new traits and phenotyping strategies in dairy cattle

487   with a focus on functional traits. Animal 9(2):191-207.

488   Fenwick, M. A., R. Fitzpatrick, D. A. Kenny, M. G. Diskin, J. Patton, J. J. Murphy, and D. C.

489   Wathes. 2008. Interrelationships between negative energy balance (NEB) and IGF regulation

490   in liver of lactating dairy cows. Domest. Anim. Endocrinol. 34(1):31-44.

491   Grelet, C., C. Bastin, M. Gele, J. B. Daviere, M. Johan, A. Werner, R. Reding, J. A. F. Pierna,

492   F. G. Colinet, P. Dardenne, N. Gengler, H. Soyeurt, and F. Dehareng. 2016. Development of

493   Fourier transform mid-infrared calibrations to predict acetone, beta-hydroxybutyrate, and

494   citrate contents in bovine milk through a European dairy network. J. Dairy Sci. 99(6):4816-

495   4825.

496   Grelet, C., J. A. F. Pierna, P. Dardenne, V. Baeten, and F. Dehareng. 2015. Standardization of

497   milk mid-infrared spectra from a European dairy network. J. Dairy Sci. 98(4):2150-2160.

498   Grelet, C., A. Vanlierde, M. Hostens, L. Foldager, M. Salavati, K. Ingvartsen, M. Crowe, M.

499   Sorensen, E. Froidmont, and C. Ferris. 2018. Potential of milk mid-IR spectra to predict

500   metabolic status of cows through blood components and an innovative clustering approach.

501   Animal:1-10.

502   Grummer, R. R., D. G. Mashek, and A. Hayirli. 2004. Dry matter intake and energy balance in

503   the transition period. Vet. Clin. North Am. Food Anim. Pract. 20(3):447-470.

504   Ingvartsen, K. L. 2006. Feeding- and management-related diseases in the transition cow -

505   Physiological adaptations around calving and strategies to reduce feeding-related diseases.

506   Anim. Feed Sci. Tech. 126(3-4):175-213.

507   Ingvartsen, K. L. and J. B. Andersen. 2000. Integration of metabolism and intake regulation: A

508   review focusing on periparturient animals. J. Dairy Sci. 83(7):1573-1597.

509   Ingvartsen, K. L., R. J. Dewhurst, and N. C. Friggens. 2003. On the relationship between

510   lactational performance and health: is it yield or metabolic imbalance that cause production

511   diseases in dairy cattle? A position paper. Livest. Prod. Sci. 83(2-3):277-308.

512   Ingvartsen, K. L. and N. C. Friggens. 2005. To what extent do variabilities in hormones,

513   metabolites and energy intake explain variability in milk yield? Domest. Anim. Endocrinol.

514   29(2):294-304.

515   Jenkins, N. T., G. Pena, C. Risco, C. C. Barbosa, A. Vieira-Neto, and K. N. Galvao. 2015.

516   Utility of inline milk fat and protein ratio to diagnose subclinical ketosis and to assign propylene

517   glycol treatment in lactating dairy cows. Can. Vet. J. 56(8):850-854.

518   Larsen, T. 2005. Determination of lactate dehydrogenase (LDH) activity in milk by a

519   fluorometric assay. J. Dairy Res. 72(2):209-216.

520   Larsen, T. 2014. Fluorometric determination of free and total isocitrate in bovine milk. J. Dairy

521   Sci. 97(12):7498-7504.

522   Larsen, T. 2015. Fluorometric determination of free glucose and glucose 6-phosphate in cows'

523   milk and other opaque matrices. Food Chem 166:283-286.

524   Larsen, T. and N. I. Nielsen. 2005. Fluorometric determination of beta-hydroxybutyrate in milk

525   and blood plasma. J. Dairy Sci. 88(6):2004-2009.

526   Larsen, T., C. M. Rontved, K. L. Ingvartsen, L. Vels, and M. Bjerring. 2010. Enzyme activity

527   and acute phase proteins in milk utilized as indicators of acute clinical E. coli LPS-induced

528   mastitis. Animal 4(10):1672-1679.

529   Lomander, H., J. Frossling, K. L. Ingvartsen, H. Gustafsson, and C. Svensson. 2012.

530   Supplemental feeding with glycerol or propylene glycol of dairy cows in early lactation--effects

531   on metabolic status, body condition, and milk yield. J. Dairy Sci. 95(5):2397-2408.

532   Lucy, M. C. 2008. Functional differences in the growth hormone and insulin-like growth factor

533   axis in cattle and pigs: Implications for post-partum nutrition and reproduction. Reprod.

534   Domest. Anim. 43:31-39.

535   Lucy, M. C., G. A. Verkerk, B. E. Whyte, K. A. Macdonald, L. Burton, R. T. Cursons, J. R.

536   Roche, and C. W. Holmes. 2009. Somatotropic axis components and nutrient partitioning in

537   genetically diverse dairy cows managed under different feed allowances in a pasture system. J.

538   Dairy Sci. 92(2):526-539.

539   Maury, E., K. Ehala-Aleksejev, Y. Guiot, R. Detry, A. Vandenhooft, and S. M. Brichard. 2007.

540   Adipokines oversecreted by omental adipose tissue in human obesity. Am. J. Physiol.

541   Endocrinol. Metab. 293(3):E656-665.

542   McArt, J. A. A., D. V. Nydam, G. R. Oetzel, T. R. Overton, and P. A. Ospina. 2013. Elevated

543   non-esterified fatty acids and β-hydroxybutyrate and their association with transition dairy cow

544   performance. Vet. J. 198(3):560-570.

545   McCarthy, M. M., S. Mann, D. V. Nydam, T. R. Overton, and J. A. A. McArt. 2015. Short

546   communication: Concentrations of nonesterified fatty acids and β-hydroxybutyrate in dairy

547   cows are not well correlated during the transition period. J. Dairy Sci. 98(9):6284-6290.

548   Moyes, K. M., T. Larsen, and K. L. Ingvartsen. 2013. Generation of an index for physiological

549   imbalance and its use as a predictor of primary disease in dairy cows during early lactation. J.

550   Dairy Sci. 96(4):2161-2170.

551   Mulligan, F. J., L. O'Grady, D. A. Rice, and M. L. Doherty. 2006. A herd health approach to

552   dairy cow nutrition and production diseases of the transition cow. Anim. Reprod. Sci.

553   96(3):331-353.

554   Nielsen, N. I., N. C. Friggens, M. G. Chagunda, and K. L. Ingvartsen. 2005a. Predicting risk of

555   ketosis in dairy cows using in-line measurements of beta-hydroxybutyrate: a biological model.

556   J. Dairy Sci. 88(7):2441-2453.

557   Nielsen, N. I., T. Larsen, M. Bjerring, and K. L. Ingvartsen. 2005b. Quarter health, milking

558   interval, and sampling time during milking affect the concentration of milk constituents. J.

559   Dairy Sci. 88(9):3186-3200.

560   NRC. 2001. Nutrient requirements of dairy cattle. Vol. 1. 7th ed. Natl. Acad. Press.,

561   Washington, DC.

562   Ospina, P., J. McArt, T. Overton, T. Stokol, and D. Nydam. 2013. Using nonesterified fatty

563   acids and β-hydroxybutyrate concentrations during the transition period for herd-level

564   monitoring of increased risk of disease and decreased reproductive and milking performance.

565   Vet. Clin. North Am. Food Anim. Pract. 29(2):387-412.

566 Puppel, K. and B. Kuczynska. 2016. Metabolic profiles of cow's blood; a review. J. Sci. Food

567 Agric. 96(13):4321-4328.

568 Radcliff, R. P., B. L. McCormack, B. A. Crooker, and M. C. Lucy. 2003. Plasma hormones and

569 expression of growth hormone receptor and insulin-like growth factor-I mRNA in hepatic tissue

570 of periparturient dairy cows. J. Dairy Sci. 86(12):3920-3926.

571 Roche, J. R., J. K. Kay, N. C. Friggens, J. J. Loor, and D. P. Berry. 2013. Assessing and

572 Managing Body Condition Score for the Prevention of Metabolic Disease in Dairy Cows. Vet.

573 Clin. North Am. Food Anim. Pract. 29(2):323-336.

574 Santos, J. E., R. S. Bisinotto, E. S. Ribeiro, F. S. Lima, L. F. Greco, C. R. Staples, and W. W.

575 Thatcher. 2010. Applying nutrition and physiology to improve reproduction in dairy cattle. Soc.

576 Reprod. Fertil. Suppl. 67:387-403.

577 Schröder, U. and R. Staufenbiel. 2006. Invited Review: Methods to Determine Body Fat

578 Reserves in the Dairy Cow with Special Regard to Ultrasonographic Measurement of Backfat

579 Thickness. J. Dairy Sci. 89(1):1-14.

580 Soyeurt, H., F. Dehareng, N. Gengler, S. McParland, E. Wall, D. P. Berry, M. Coffey, and P.

581 Dardenne. 2011. Mid-infrared prediction of bovine milk fatty acids across multiple breeds,

582 production systems, and countries. J. Dairy Sci. 94(4):1657-1667.

583 Stockmann, H., B. Adamczyk, J. Hayes, and P. M. Rudd. 2013. Automated, high-throughput

584 IgG-antibody glycoprofiling platform. Anal. Chem. 85(18):8841-8849.

585 Thorup, V. M., D. Edwards, and N. C. Friggens. 2012. On-farm estimation of energy balance

586 in dairy cows using only frequent body weight measurements and body condition score. J. Dairy

587 Sci. 95(4):1784-1793.

588 van Straten, M., N. Y. Shpigel, and M. Friger. 2008. Analysis of daily body weight of high-

589 producing dairy cows in the first one hundred twenty days of lactation and associations with

590 ovarian inactivity. J. Dairy Sci. 91(9):3353-3362.

591   Voelker, J. A. and M. S. Allen. 2003. Pelleted beet pulp substituted for high-moisture corn: 3.

592   Effects on ruminal fermentation, pH, and microbial protein efficiency in lactating dairy cows.

593   J. Dairy Sci. 86(11):3562-3570.

594   Wallace, T. M. and D. R. Matthews. 2002. The assessment of insulin resistance in man. Diabet.

595   Med. 19(7):527-534.

596   Wathes, D. C., Z. Cheng, M. A. Fenwick, R. Fitzpatrick, and J. Patton. 2011. Influence of

597   energy balance on the somatotrophic axis and matrix metalloproteinase expression in the

598   endometrium of the postpartum dairy cow. Reproduction 141(2):269-281.

599   Weekes, T. E., Y. Sasaki, and T. Tsuda. 1983. Enhanced responsiveness to insulin in sheep

600   exposed to cold. Am. J. Physiol. 244(4):335-345.

601   Zhao, F. Q. and A. F. Keating. 2007. Expression and regulation of glucose transporters in the

602   bovine mammary gland. J. Dairy Sci. 90 Suppl 1:E76-86.

603   Zulu, V. C., T. Nakao, and Y. Sawamukai. 2002. Insulin-like growth factor-I as a possible

604   hormonal mediator of nutritional regulation of reproduction in cattle. J. Vet. Med. Sci.

605   64(8):657-665.

606

608 **TABLES**

Table 1. Glucose, insulin like growth factor I (IGF-I), free fatty acid (FFA) and β-hydroxybutyrate (BHB) concentrations in metabolic clusters[1] around 14 DIM (D14, clusters A, B and C) and 35 DIM (D35, clusters D, E and F). Data are presented as LSMEANS ± SEM.

	D14			D35		
Cluster[1]	A	B	C	D	E	F
Glucose (mM)	3.59 ± 0.04 [a]	3.36 ± 0.04 [b]	2.60 ± 0.10 [c]	3.83 ± 0.04 [a]	3.46 ± 0.06 [b]	3.35 ± 0.06 [b]
IGF-I (ng/mL) [2]	115.35 ± 6.66 [a]	43.26 ± 2.52 [b]	63.22 ± 10.95 [b]	136.48 ± 6.77 [a]	49.44 ± 3.77 [b]	59.12 ± 4.60 [b]
FFA (mM) [3]	0.46 ± 0.03 [a]	0.87 ± 0.06 [b]	1.47 ± 0.32 [b]	0.33 ± 0.03 [a]	0.26 ± 0.03 [a]	0.77 ± 0.11 [b]
BHB (mM)	0.46 ± 0.02 [a]	0.65 ± 0.03 [b]	2.19 ± 0.32 [c]	0.43 ± 0.02 [a]	0.52 ± 0.03 [b]	1.14 ± 0.07 [c]

[abc] LSMEANS with different superscript within the same time period differ ($P < 0.05$).

[1] Clusters A, B and C around 14 DIM and clusters D, E and F around 35 DIM were created using the k-means clustering method based on the log-transformed and standardized concentrations of glucose, IGF-I, FFA and BHB.

[2] Trend for a difference between B and C ($P = 0.10$).

[3] Trend for a difference between B and C ($P = 0.06$).

609

Table 2. Number of cows in metabolic cluster[1] A, B, and C around 14 DIM (D14) and metabolic cluster D, E and F around 35 DIM (D35).

		D35			
	Cluster[1]	D	E	F	Total
D14	A	43	3	5	51
	B	15	22	14	51
	C	0	0	5	5
	Total	58	25	24	107

[1] Clusters A, B and C around 14 DIM and clusters D, E and F around 35 DIM were created using the k-means clustering method based on the log-transformed and standardized concentrations of glucose, IGF-I, FFA and BHB.

610

Table 3. Glucose, insulin like growth factor I (IGF-I), free fatty acid (FFA) and β-hydroxybutyrate (BHB) concentrations in BALANCED and OTHER$_{BAL}$ metabolic cluster around 14 DIM (D14) and 35 DIM (D35). Data are presented as LSMEANS ± SEM.

	D14		D35		*P*-values		
	BALANCED	OTHER$_{BAL}$	BALANCED	OTHER$_{BAL}$	Cluster	Period	Cluster*Period
Glucose (mM) [1]	3.62 ± 0.05 [a]	3.28 ± 0.04 [b]	3.84 ± 0.05 [a]	3.49 ± 0.04 [b]	< 0.0001	< 0.0001	NS
IGF-I (ng/mL) [1]	121.66 ± 7.60 [a]	49.63 ± 2.55 [b]	152.93 ± 9.55 [a]	62.39 ± 3.24 [b]	< 0.0001	< 0.0001	NS
FFA (mM) [2]	0.52 ± 0.05 [a]	0.81 ± 0.06 [b]	0.29 ± 0.03 [a]	0.45 ± 0.03 [b]	< 0.001	< 0.0001	NS
BHB (mM)	0.44 ± 0.03 [a]	0.71 ± 0.04 [b]	0.42 ± 0.03 [a]	0.69 ± 0.04 [b]	< 0.0001	0.42	NS
BCS [3]	2.75 ± 0.06	2.87 ± 0.05	2.71 ± 0.05	2.71 ± 0.05	0.43	< 0.0001	< 0.05

[ab] LSMEANS with different superscript within the same period differ (*P* < 0.05).

[1] Period effect indicates an increase in glucose and IGF-I concentrations on D35 compared with D14 in both clusters.

[2] Period effect indicates a decrease in FFA concentrations on D35 compared with D14 in both clusters.

[3] Cluster * Period interaction effect indicates a decrease in BCS on D35 compared with D14 in OTHER$_{BAL}$ but not in BALANCED.

611

Table 4. Fat and protein corrected milk yield (FPCM), dry matter intake (DMI), body weight (BW) and energy balance (EBAL) in BALANCED

and OTHER$_{BAL}$ metabolic cluster and the effect of day after calving (slope DIM) in the postpartum period.

		Cluster		P-values		
		BALANCED	OTHER$_{BAL}$	DIM	Cluster	DIM*Cluster
FPCM (kg/d)	LSMEANS ± SEM	38.42 ± 0.99	37.76 ± 0.80	< 0.05	0.60	NS
	Slope DIM ± SEM	0.043 ± 0.018	0.043 ± 0.018			
DMI (kg/d)	LSMEANS ± SEM	22.12 ± 0.48 [a]	18.72 ± 0.39 [b]	< 0.0001	< 0.01	< 0.01
	Slope DIM ± SEM	0.11 ± 0.01 [a]	0.06 ± 0.01 [b]			
BW (kg)	LSMEANS ± SEM	647.97 ± 9.56	650.51 ± 7.66	< 0.0001	< 0.05	< 0.0001
	Slope DIM ± SEM	-0.34 ± 0.14 [a]	-1.41 ± 0.11 [b]			
EBAL (MCal/d)	LSMEANS ± SEM	-2.40 ± 0.71 [a]	-7.59 ± 0.61 [b]	< 0.0001	< 0.0001	NS
	Slope DIM ± SEM	0.11 ± 0.01	0.11 ± 0.01			

[ab] LSMEANS and slope with different superscript differ ($P < 0.05$).

Table 5. Glucose, insulin like growth factor I (IGF-I), free fatty acid (FFA) and β-hydroxybutyrate (BHB) concentrations in IMBALANCED and OTHER$_{IMBAL}$ metabolic cluster around 14 DIM (D14) and 35 DIM (D35). Data are presented as LSMEANS ± SEM.

	D14		D35		P-values		
	IMBALANCED	OTHER$_{IMBAL}$	IMBALANCED	OTHER$_{IMBAL}$	Cluster	Period	Cluster*Period
Glucose (mM) [1]	3.19 ± 0.06 [a]	3.49 ± 0.04 [b]	3.38 ± 0.07 [a]	3.69 ± 0.04 [b]	< 0.0001	< 0.0001	NS
IGF-I (ng/mL) [2]	43.61 ± 5.53 [a]	78.56 ± 4.73 [b]	55.07 ± 6.98 [a]	99.21 ± 6.01 [b]	< 0.0001	< 0.0001	NS
FFA (mM) [3]	1.32 ± 0.17 [a]	0.56 ± 0.04 [b]	0.76 ± 0.10 [a]	0.32 ± 0.02 [b]	< 0.0001	< 0.0001	NS
BHB (mM)	1.15 ± 0.09 [a]	0.50 ± 0.02 [b]	1.11 ± 0.09 [a]	0.49 ± 0.02 [b]	< 0.0001	0.39	NS
BCS [4]	3.16 ± 0.09 [a]	2.75 ± 0.04 [b]	2.89 ± 0.09 [a]	2.66 ± 0.04 [b]	< 0.0001	< 0.0001	< 0.001

[a,b] LSMEANS with different superscript within the same period differ ($P < 0.05$).

[1] Glucose concentrations were not different between periods for IMBALANCED and increased on D35 for OTHER$_{IMBAL}$.

[2] Period effect indicates an increase in IGF concentrations on D35 compared with D14 in both clusters.

[3] Period effect indicates a decrease in FFA concentrations on D35 compared with D14 in both clusters.

[4] Period effect indicates a decrease in BCS on D35 compared with D14 in both clusters. There was a more pronounced decrease in BCS from D14 to D35 in IMBAL compared with OTHER$_{IMBAL}$.

Table 6. Fat and protein corrected milk yield (FPCM), dry matter intake (DMI), body weight (BW) and energy balance (EBAL) in IMBALANCED and OTHER$_{IMBAL}$ metabolic cluster and the effect of day after calving (slope DIM) in the postpartum period.

		Cluster		P-values		
		IMBALANCED	OTHER$_{IMBAL}$	DIM	Cluster	DIM*Cluster
FPCM (kg/d)	LSMEANS ± SEM	41.14 ± 1.47 [a]	37.29 ± 0.66 [b]	< 0.01	< 0.05	NS
	Slope DIM ± SEM	0.049 ± 0.018	0.049 ± 0.018			
DMI (kg/d)	LSMEANS ± SEM	18.56 ± 0.77 [a]	20.51 ± 0.35 [b]	< 0.0001	0.37	0.05
	Slope DIM ± SEM	0.05 ± 0.02 [a]	0.10 ± 0.01 [b]			
BW (kg)	LSMEANS ± SEM [1]	672.29 ± 14.12	642.73 ± 6.42	< 0.0001	< 0.01	< 0.001
	Slope DIM ± SEM	-1.69 ± 0.22 [a]	-0.80 ± 0.10 [b]			
EBAL (MCal/d)	LSMEANS ± SEM	-11.14 ± 1.25 [a]	-4.32 ± 0.50 [b]	< 0.0001	< 0.0001	NS
	Slope DIM ± SEM	0.11 ± 0.01	0.11 ± 0.01			

[ab] LSMEANS and slope with different superscript differ ($P < 0.05$).

[1] Trend for a difference between IMBALANCED and OTHER$_{IMBAL}$ ($P = 0.06$).

Table 7. Prediction accuracy within each run and summary for different classification models to predict BALANCED cows on 14 DIM (D14) and 35 DIM (D35), or D14, or D35.

Milk biomarker predictor set	Class	Run1	Run2	Run3	Run4	Run5	Run6	Run7	Run8	Run9	Run10	Mean	Min	Max	CV
Metabolites and enzymes	D14 and D35	70%	74%	63%	70%	64%	67%	58%	60%	70%	64%	66%	58%	74%	8%
	D14	69%	59%	68%	54%	68%	63%	62%	65%	57%	61%	63%	54%	69%	8%
	D35	60%	74%	60%	61%	58%	61%	71%	56%	65%	47%	61%	47%	74%	12%
+ DIM & Parity	D14 and D35	68%	63%	70%	71%	71%	72%	58%	72%	80%	71%	69%	58%	80%	9%
	D14	67%	59%	66%	72%	65%	65%	59%	71%	65%	64%	65%	59%	72%	7%
	D35	69%	64%	71%	63%	59%	60%	67%	59%	60%	60%	63%	59%	71%	7%
+ DIM & Parity + Milk yield features	D14 and D35	63%	70%	61%	72%	67%	73%	61%	65%	51%	70%	65%	51%	73%	10%
	D14	57%	60%	75%	68%	70%	59%	66%	67%	60%	65%	65%	57%	75%	9%
	D35	73%	68%	72%	64%	64%	58%	55%	59%	57%	68%	64%	55%	73%	10%
IgG Glycans	D14 and D35	67%	52%	66%	59%	57%	56%	68%	73%	53%	58%	61%	52%	73%	12%
	D14	43%	43%	47%	50%	52%	43%	45%	39%	42%	50%	45%	39%	52%	9%
	D35	68%	58%	66%	58%	58%	47%	52%	67%	58%	66%	60%	47%	68%	12%
+ DIM & Parity	D14 and D35	63%	53%	73%	58%	63%	61%	61%	64%	55%	51%	60%	51%	73%	11%
	D14	50%	62%	49%	52%	51%	40%	41%	62%	58%	42%	51%	40%	62%	16%
	D35	55%	59%	57%	47%	57%	61%	63%	62%	55%	62%	58%	47%	63%	8%
+ DIM & Parity + Milk yield features	D14 and D35	69%	64%	57%	62%	60%	62%	67%	50%	65%	63%	62%	50%	69%	9%
	D14	45%	55%	37%	56%	42%	52%	45%	58%	56%	55%	50%	37%	58%	15%
	D35	37%	65%	54%	56%	60%	66%	63%	64%	65%	60%	59%	37%	66%	15%
FT-MIR spectra	D14 and D35	67%	76%	72%	70%	77%	73%	78%	78%	68%	75%	73%	67%	78%	5%
	D14	70%	68%	68%	64%	71%	70%	73%	68%	63%	67%	68%	63%	73%	5%
	D35	65%	71%	63%	62%	64%	65%	66%	80%	69%	66%	67%	62%	80%	8%
+ DIM & Parity	D14 and D35	76%	77%	70%	78%	77%	74%	78%	78%	78%	74%	76%	70%	78%	4%
	D14	68%	70%	59%	67%	61%	79%	71%	70%	74%	70%	69%	59%	79%	8%
	D35	68%	67%	71%	69%	47%	63%	63%	65%	70%	64%	65%	47%	71%	10%
+ DIM & Parity + Milk yield features	D14 and D35	76%	76%	69%	65%	79%	76%	69%	73%	72%	74%	73%	65%	79%	6%
	D14	71%	67%	72%	74%	63%	66%	66%	65%	71%	61%	68%	61%	74%	6%
	D35	61%	57%	63%	56%	64%	67%	65%	66%	65%	76%	64%	56%	76%	9%

[1] color code: green indicates better (i.e. lower CV or higher mean, min, max accuracy), red indicates worse (i.e. higher CV, lower mean, min, max accuracy)

Table 8. Prediction accuracy within each run and summary for different classification models to predict IMBALANCED cows on 14 DIM (D14) and 35 DIM (D35), or D14, or D35.

Milk biomarker predictor set	Class	Run1	Run2	Run3	Run4	Run5	Run6	Run7	Run8	Run9	Run10	Mean	Min	Max	CV
Metabolites and enzymes	D14 and D35	81%	82%	84%	77%	71%	79%	80%	81%	83%	79%	80%	71%	84%	5%
	D14	79%	82%	80%	79%	81%	77%	81%	80%	80%	81%	80%	77%	82%	2%
	D35	85%	85%	75%	85%	76%	71%	81%	82%	79%	83%	80%	71%	85%	6%
+ DIM & Parity	D14 and D35	81%	87%	76%	80%	87%	88%	77%	76%	76%	78%	81%	76%	88%	6%
	D14	81%	80%	84%	75%	80%	84%	76%	78%	79%	83%	80%	75%	84%	4%
	D35	83%	83%	76%	76%	77%	76%	81%	80%	81%	79%	79%	76%	83%	4%
+ DIM & Parity + Milk yields features	D14 and D35	87%	78%	82%	79%	79%	82%	79%	79%	80%	80%	80%	78%	87%	3%
	D14	84%	83%	76%	78%	82%	86%	79%	82%	83%	72%	81%	72%	86%	5%
	D35	81%	88%	79%	87%	80%	78%	87%	77%	80%	81%	82%	77%	88%	5%
IgG Glycans	D14 and D35	81%	78%	83%	80%	80%	81%	82%	84%	70%	80%	80%	70%	84%	5%
	D14	73%	71%	84%	75%	77%	71%	81%	74%	80%	78%	76%	71%	84%	6%
	D35	76%	73%	79%	74%	73%	71%	80%	73%	74%	80%	75%	71%	80%	4%
+ DIM & Parity	D14 and D35	81%	82%	78%	79%	82%	79%	74%	77%	81%	80%	79%	74%	82%	3%
	D14	80%	77%	75%	81%	72%	77%	79%	79%	72%	81%	77%	72%	81%	4%
	D35	75%	75%	73%	72%	75%	76%	77%	80%	77%	77%	76%	72%	80%	3%
+ DIM & Parity + Milk yields features	D14 and D35	81%	79%	76%	80%	81%	76%	68%	76%	78%	79%	77%	68%	81%	5%
	D14	76%	82%	79%	78%	77%	81%	82%	76%	82%	79%	79%	76%	82%	3%
	D35	63%	72%	69%	74%	71%	76%	74%	68%	70%	78%	72%	63%	78%	6%
FT-MIR Spectra	D14 and D35	84%	81%	78%	79%	76%	87%	76%	81%	83%	79%	80%	76%	87%	4%
	D14	77%	84%	83%	81%	86%	84%	76%	86%	87%	79%	82%	76%	87%	5%
	D35	69%	72%	79%	76%	73%	75%	76%	85%	72%	78%	76%	69%	85%	6%
+ DIM & Parity	D14 and D35	78%	78%	73%	74%	77%	81%	73%	82%	80%	78%	77%	73%	82%	4%
	D14	85%	87%	86%	80%	82%	84%	86%	80%	87%	85%	84%	80%	87%	3%
	D35	73%	74%	77%	79%	76%	76%	73%	74%	79%	80%	76%	73%	80%	3%
+ DIM & Parity + Milk yields features	D14 and D35	80%	77%	71%	75%	73%	87%	86%	82%	84%	79%	79%	71%	87%	7%
	D14	86%	87%	84%	80%	87%	83%	81%	87%	85%	85%	84%	80%	87%	3%
	D35	69%	70%	73%	76%	78%	77%	73%	80%	73%	76%	74%	69%	80%	5%

[1] color code: green indicates better (i.e. lower CV or higher mean, min, max accuracy), red indicates worse (i.e. higher CV, lower mean, min, max accuracy)
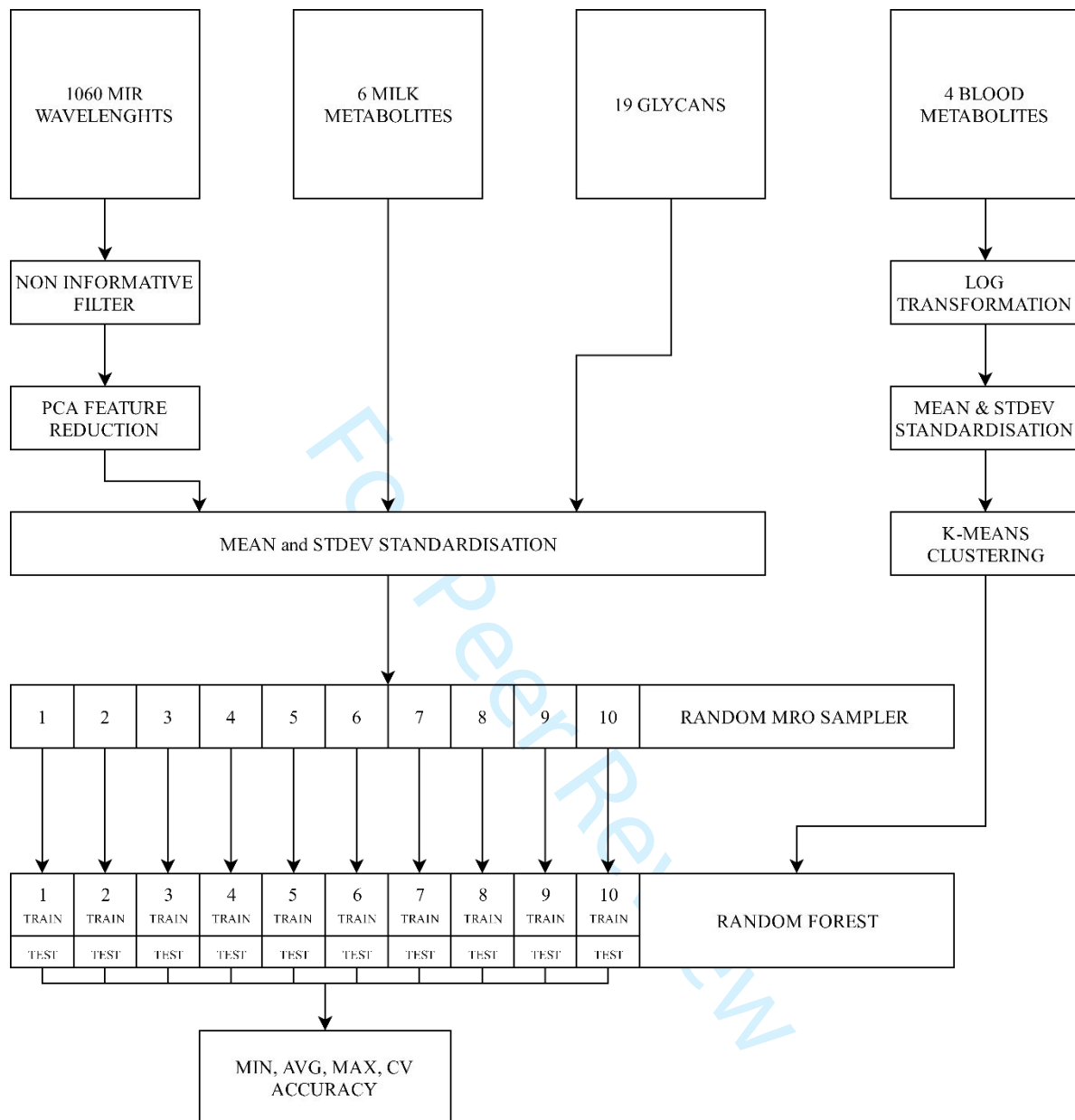
615

616

FIGURES

Figure 1. Statistical prediction methodology for metabolic clustering in dairy cows.
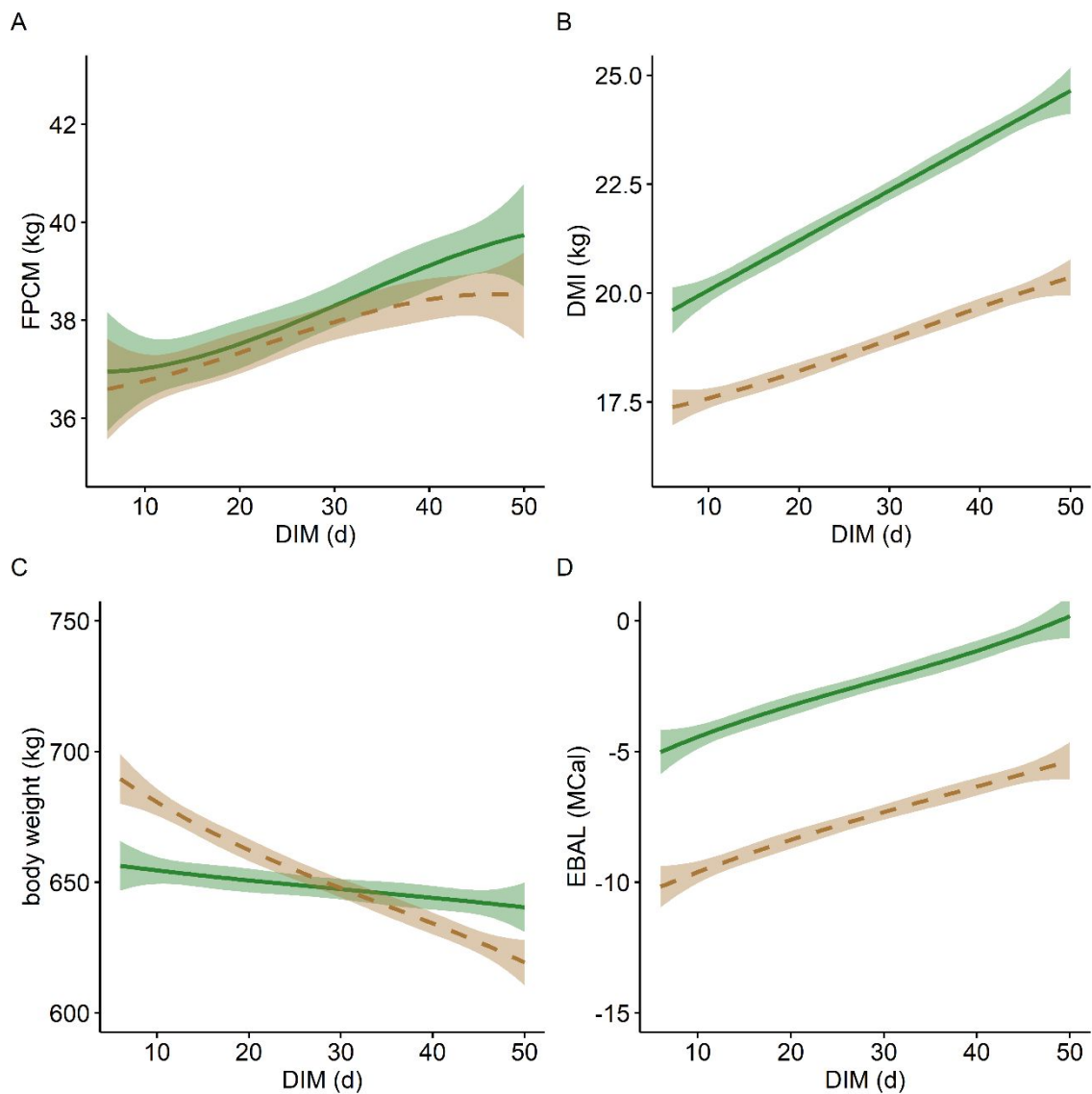
Figure 2. Fat and protein corrected milk yield (FPCM, A), dry matter intake (DMI, B), body weight (BW, C) and energy balance (EBAL, D) in BALANCED (green, continuous line) and OTHER$_{BAL}$ (orange, dashed line) metabolic clusters. Lines represent the LSMEANS of the models and the colored area represent the 95% confidence limits.
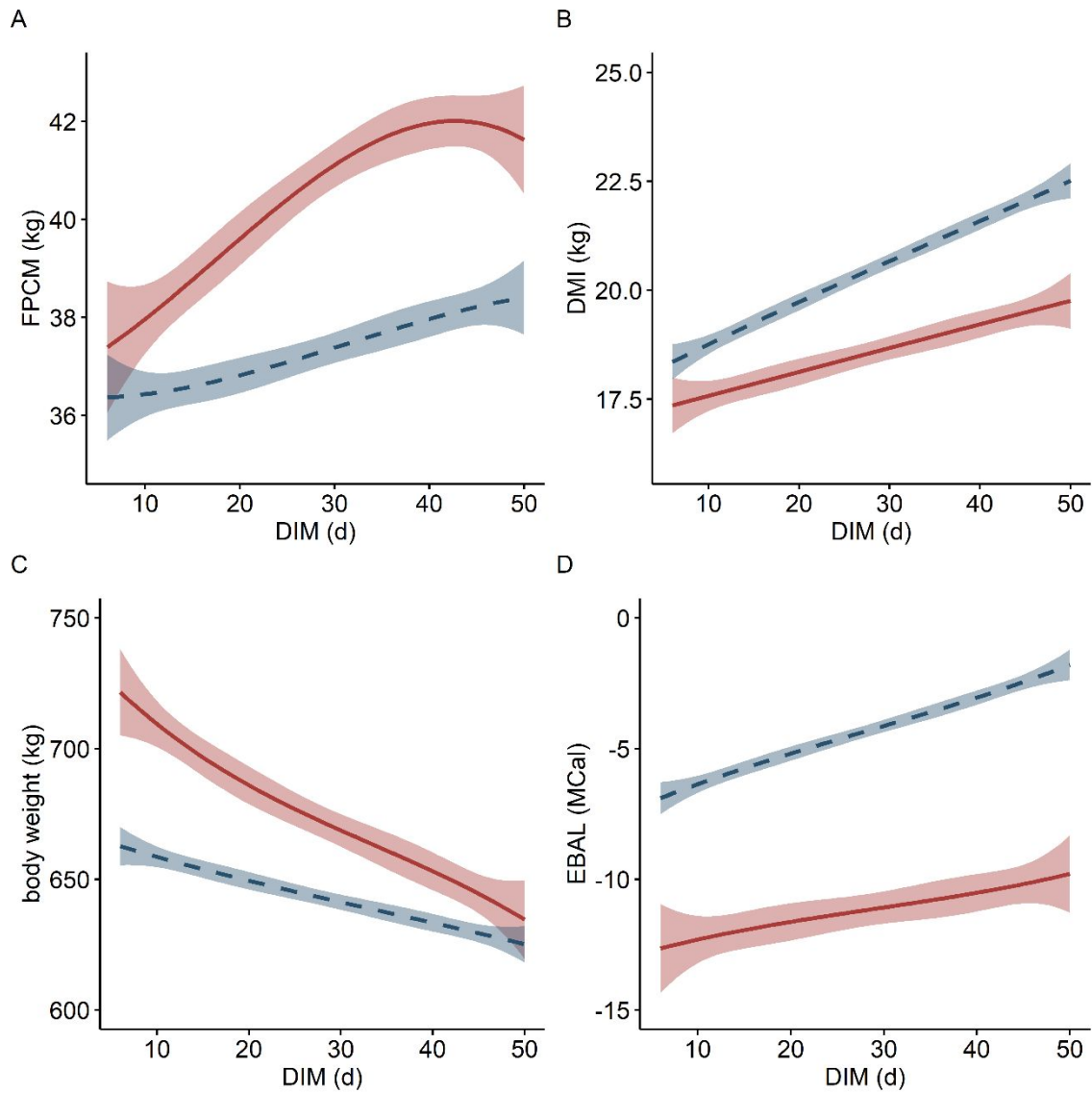
A

B

C

D

Figure 3. Fat and protein corrected milk yield (FPCM, A), dry matter intake (DMI, B), body weight (BW, C) and energy balance (EBAL, D) in IMBALANCED (red, continuous line) and OTHER$_{IMBAL}$ (blue, dashed line) metabolic clusters. Lines represent the LSMEANS of the models and the colored area represent the 95% confidence limits.

619