Milk Urea Nitrogen: Precision, Accuracy, and Individual Animal Variability

K. F. Reed, E. M. Wood, M. Portnoy, D. Barbano
Department of Animal Science
Cornell University

Introduction

Nitrogen loss from dairy farms negatively impacts the environment by contributing to greenhouse gas emissions, soil acidification, ground water contamination and surface water eutrophication (Hristov et al. 2011). Much of the N lost from dairy farms originates from manure N and, in particular, urinary N. Thus, reducing the amount of nitrogen excreted by individual cows is an important step in reducing the detrimental environmental effects of the dairy industry. Nutrition strategies that increase the efficiency with which feed N is converted into milk N reduce the amount of manure N produced per unit of milk. Milk Urea Nitrogen (MUN) is commonly used as an indicator of protein metabolism and nitrogen efficiency in lactating dairy cattle to guide management and diet formulation decisions.

Urea is a byproduct of protein metabolism generated when the liver converts the ammonia produced during amino acid catabolism into non-toxic urea that is primarily excreted in urine. Due to the ability of urea to diffuse and equilibrate across membranes that separate blood, urine, and milk, plasma urea N (PUN), MUN, and urinary urea N (UUN) concentrations are highly correlated such that one measurement can be used to estimate the others (Gustafsson and Palmquist, 1993). Since milk sampling already occurs regularly on farms and is less invasive than blood sampling, MUN is the preferred method for estimating UUN excretion and assessing N efficiency on commercial dairies.

Unfortunately, previous research raised uncertainty about the precision and accuracy of commercial MUN testing, suggesting that results from a the same set of bulk tank samples sent to multiple labs had a wider range of reported results than the recommended target range of 8-12 mg/dl (Weeks and Hristov, 2017). Thus, one objective of this study is to reevaluate the precision and accuracy of mid-infrared spectroscopy (MIR) for MUN analysis. Additionally, accurate interpretation of MUN values requires an understanding of natural variation in MUN over time. While dairy farms often measure MUN during routine bulk tank sampling, the movement towards precision management will benefit from individual cow and pen-level metrics like MUN. Therefore, the secondary objective of this study is to quantify the expected variation in MUN over the course of lactation in individual cows. Combination of testing precision and expected variation will enable more effective interpretation of MUN results.

Methods

Precision and Accuracy of MIR MUN Analysis

Bulk tank samples were collected for 7 consecutive days and sent to 3 commercial labs (Labs A, B, and C) and the Barbano Lab (Lab D) in duplicate. Bulk tank samples were collected daily between 10:00 and 13:00. Samples were immediately placed on ice, stored overnight at 40 ^{o}F, and either delivered the following morning to Labs A and D or shipped to the additional commercial labs (Labs B and C).

Additional sample sets from the Federal Milk Market Administrator (FMMA) quality assurance program were prepared by Lab D and sent out to Labs A, B, and C for MUN analysis. The FFMA quality assurance program prepares 10 milk samples every 2 weeks from around the country that are composite samples from multiple bulk tanks in that region. The sets are used to ensure the accuracy of milk testing labs meet the USDA's standards for milk payments which are based on milk fat and protein content. MUN is not included in this quality assurance program so labs are not required to report their MUN results from these sample sets. Lab D prepared three sets of the FFMA samples in duplicate (i.e. 20 samples each) on three separate weeks for shipment to Labs A, B, and C.

Mid-infrared spectroscopy was used to evaluate MUN content of all samples at each lab. In addition, Lab D performed the Megazyme Urea/Ammonia Assay Procedure (Barbano and Coon, 2017) on all samples and this enzymatic assay was used as the reference chemistry for subsequent data analysis and comparison.

Individual Cow Variation

Milk samples were collected from 16 multiparous Holstein cows at each of 3X daily milkings during 3, 7-day periods in early, middle, and late-lactation. Samples were collected in triplicate in period 1 (P1) and in duplicate for periods 2 and 3 (P2 and P3). One set was sent to commercial Lab A for MIR analysis of MUN and milk components. Lab D analyzed the second set with MIR analysis for milk components and MUN and performed the enzymatic assay for MUN on the third sample set collected for P1.

Animal Care and Sample Collection

Animals were housed at the Cornell University Ruminant Center and all procedures were reviewed and approved by the university's IACUC. During P1, cows were housed in tie stalls. In P2 and P3 cows were moved to pens and were required to be in the pen for a minimum of one-week prior to each sampling period. Two of the 16 cows were culled during the study period leaving 14 cows that made it through the each of the 3 periods.

Cows were fed a standard high-cow TMR with an average CP content of 15.4% that ranged from 14.7-16.2%.

Milk samples were collected from individual cows using DeLaval in-line sampler and production, time and date were recorded for each sample. All samples were transferred to 1 L plastic bottles, inverted to mix, and aliquoted into sub-sample tubes for Labs A and D. Samples were stored at 40 ^{o}F before delivery to Labs A and D. Samples collected for MIR analysis were processed immediately, while samples to be used in the enzymatic MUN assay were frozen and stored for later processing.

For all samples, MUN was measured using MIR technology. Lab A analyzed samples using Milkoscan FT+ and Milksocan FOSS 7 spectrometers while Lab D used the Delta FTA. Both labs reported values for milk fat, protein, lactose, somatic cell count (SCC), and MUN.

Statistical Analysis

All data analysis was performed in R version 3.6.3. Mixed models were fit with the lmer() package and all other functions were performed using base packages.

Precision and Accuracy of MIR MUN

Statistical methods for evaluation of methodological agreement described by Lynch (1998) and used by Kaylegian et al. (2006) were applied to bulk tank MIR MUN analysis results. The Mean Difference (MD) was calculated by subtracting the reference chemistry value from the MIR spectroscopy value and averaging the difference over the sample sets. The standard deviation of the difference (SDD) was calculated as the square root of the summed squared value of the differences divided by the number of samples. The Euclidian distance (ED) was calculated as the distance from the origin of the points when the SDD is plotted against the MD. The coefficient of variation (CV) was calculated as the SDD divided by the mean reference chemistry MUN for each sample set. The repeatability (sr) was estimated for each commercial lab by calculating the square-root of the summed squared differences between duplicate analyses on the same sample divided by the number of samples. The reproducibility (sR) was also estimated for all labs by calculating the square-root of the summed, squared differences between the MIR analysis and the reference chemistry divided by the total number of samples tested at each lab.

Sample resutls were also fit to a linear mixed model:

$$\mathrm{MUNDiff}_{ij} = \beta_0 + \beta_1 \mathrm{MUNRefC}_{ij} + \beta_2 \mathrm{Prot}_{ij} + \beta_3 \mathrm{Fat}_{ij} + \lambda_i + \sigma_{ij} \quad [1]$$

In Eq. [1], $\mathrm{MUNDiff}_{ij}$ is the difference between the MIR analytical value for MUN and the reference chemistry MUN for *j*th sample from the *i*th lab; β_0 is the intercept that represents the mean difference between the reference chemistry and the MIR analysis; β_1 is the slope that represents the change in MUNDiff as the reference chemistry MUN value moves away from the mean of the reported values; $\mathrm{MUNRefC}_{ij}$ is the mean-centered value of the MUN reference chemistry, Prot_{ij} and Fat_{ij} are the MIR values for the true protein and fat composition and β_2 and β_3 are the slopes that represent the change

in MUNDiff as milk protein and fat increase, respectively; λ_i is the random effect of the *i*th lab; and σ_{ij} is the residual random error.

Individual Animal MUN Variation

The individual animal MUN data from all three periods was fit to the following models:

$$\mathrm{MUN}_{ijklm} = \mathrm{Lab}_i + \mathrm{Milking}_j + \mathrm{Period}_k + \alpha_l + \delta_m + \sigma_{ijklm} \quad [2]$$

$$\mathrm{MUN}_{ijklm} = \mathrm{Lab}_i + \mathrm{Milking}_j + \mathrm{Period}_k + \beta_{1i}\mathrm{Fat}_{ijklm} + \beta_{2i}\mathrm{Protein}_{ijklm} + \alpha_l + \delta_m + \sigma_{ijklm} \quad [3]$$

In Eq. [2] MUN_{ijklm} represents the raw MUN value for a sample tested by the *i*th Lab (labs A or D), collected during the *j*th milking of the *k*th period from the *l*th animal on the *m*th date. Lab was included as a fixed effect rather than fitting a model to each lab separately. Neither DIM nor CP level are included as variables due to a high correlation between these two potential variables as a result of the short sampling periods. Instead, period is included as a variable, because the effect of period is related to the stage in lactation and accounts for external factors including weather, pen, etc.

In Eq. [3], the dependent variable is the same but the fixed effects include parameters to estimate the effects of fat and protein content on reported MUN concentration. In Eq. [3], β_{1i} is the effect of milk fat corresponding to the *i*th lab and β_{2i} is the effect of milk protein corresponding to the *i*th lab. These parameters are included because the bulk tank analysis showed that fat and protein content impacted the MIR difference from enzymatic MUN measurements. The parameters act as a correction factor and should therefore remove any effect cause by fat and protein interference with MIR analysis.

In both models, α_l is the random effect of the *l*th animal, δ_m is the random effect of the *m*th day, and σ is the residual random error

Results and Discussion

Precision and Accuracy of MIR MUN Analysis

The plot of the SDD vs. MD is presented in Figure 1 and the ED, which are not significantly different between labs, are presented in Table 1. There is no apparent pattern or grouping in the plot in Figure 1 which suggests that there is no systematic bias in MUN reporting for the labs included in this study. The only potential pattern that emerges is that the points from the first machine in Lab D (D1), all fall in the negative range of the x-axis (to the left of the vertical line at MD=0) which suggests that the MIR results from this machine within this lab tend to under estimate MUN.

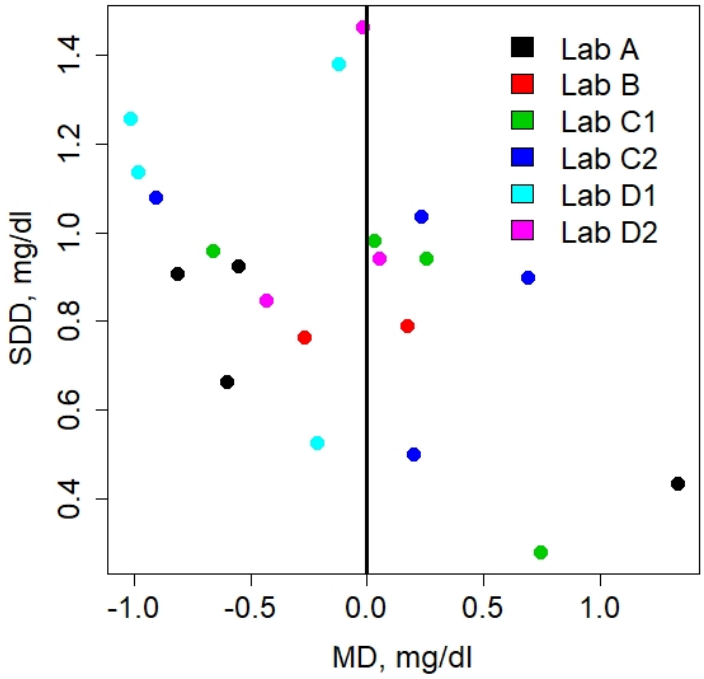

Figure 1. A plot of the standard deviation of the difference (SDD) vs mean difference (MD) for each lab (A-D) with labs C and D reporting results for two different machines.

Table 1. Euclidean Distance of MIR analysis of MUN for Labs A-D. Labs C and D reported results for two different machines which is indicated by the label.

Lab	Euclidean Distance
A	1.15
B	0.810
C2	0.978
C1	1.04
D1	1.27
D2	1.12

Repeatability (s_r) and reproducibility (sR) estimates for commercial labs are shown in Table 2 and ranged from (0.297 - 0.469) and (0.555-0.791) respectively. Repeatability is interpreted as the expected variability of a result reproduced by the same lab on the same sample. Similarly, reproducibility is the expected difference or variation between two labs or methods. In this analysis, reproducibility measures the ability of commercial lab MIR to reproduce the enzymatic assay. Repeatability and reproducibility values are interpreted like standard deviations. Since all sR values are less than 1, each of the labs is expected to predict MUN within 0.8 mg/dL of the reference chemistry value 68% of the time. This means that 95% of the time, each lab is expected to predict MUN within ± 1.6 mg/dL. Looking at the repeatability measures, all labs have an sr value < 0.5 mg/dl. This means that 95% of repeated sampling is expected to be within ±1 mg/dl. These parameters can also be expressed as percentages, similar to a coefficient of variation,

which indicate the percent of the mean MUN value by which repeated and reproduced analyses are expected to vary.

Table 2. Repeatability and reproducibility values and percentages for the commercial labs included in this study.

Lab	sr (mg/dL)	sr (%)	sR (mg/dL)	sR (%)
A	0.367	2.98	0.785	6.38
B	0.362	2.94	0.555	4.51
C1	0.297	2.41	0.701	5.70
C2	0.469	3.81	0.791	6.43

Differences in sr and sR across labs are most likely due to the use of different machines and calibration methods used for different machines.

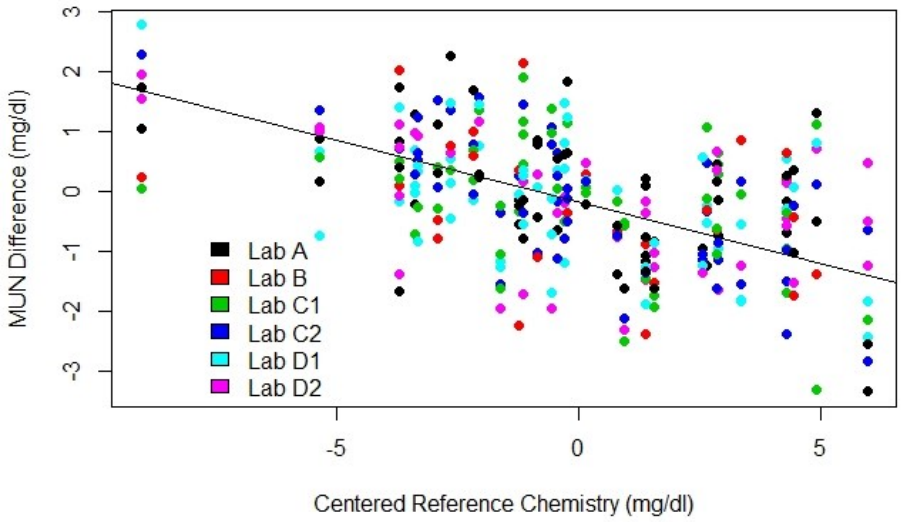

Figure 2. A plot of the differences between the MIR and reference chemistry for MUN analysis vs. the centered reference chemistry value. Line represents the fixed-effect results of the mixed-model regression.

Regression results indicate MIR analysis over-predicts MUN at low MUN concentrations and under predicts MUN at high MUN concentrations. A plot of the MUN differences against the centered reference chemistry values is shown in Figure 2 and the parameter estimates are provided in Table 3.

Table 3. The parameter estimates from a mixed-model analysis described in Eq. [1]

Parameter	Mean	SE
β_0	-2.32	0.436
β_1	-0.206	0.0168
β_2	0.397	0.1629
β_3	0.193	0.0893
σ_{Lab}	NA	0.224
σ_{Res}	NA	0.868

The results of the mixed-model analysis suggest that at the mean milk protein (3.4%), milk fat (4.2%), and MUN (12.8 mg/dl) of this dataset, MIR analysis was not significantly different than the reference chemistry. However, for every 1 mg increase in the reference chemistry value (what is considered to be the true MUN value) above 12.8 mg/dl, MIR analysis underpredicted MUN by an average of 0.206 mg/dl. This means that for a milk sample with 3.4% protein, 4.2% fat, and 15.8 mg/dl MUN, MIR analysis would be expected to underpredict MUN by 0.618 mg/dl. Similarly, as the reference chemistry decreases below the average of 12.8 mg/dl, the MIR analysis is expected to over predict MUN concentration. For a milk sample with 3.4% protein, 4.2% fat, and 7.8 mg/dl MUN, MIR analysis would be expected to over predict MUN by 1.3 mg/dl. In addition, as protein and fat levels deviate from the means in this dataset, MIR MUN analysis is expected to have some systematic over or under prediction depending on the linear combination of the fat and protein levels and their parameter estimate.

The residual standard error estimate (0.87 mg/dl) and random effect of lab (0.22 mg/dl) indicate the amount of uncertainty in MIR analysis of MUN. Combining these variance estimates, we get an overall standard error of 0.90 for MUNDiff which means at the average milk composition values, the 95% CI for differences between the MIR analysis and the reference chemistry to be between -1.8 and +1.8 mg/dl, which is very similar to the results of the reproducibility analysis reported above.

Individual Cow Variability

The mean and SE of the regression parameters corresponding to Eq. [3] and [4] are listed in Tables 4 and 5. A separate intercept representing the average MUN value during Period 1 and Milking 1 was estimated for Labs A and D. For example, for Eq.[2] the Lab_A parameter estimate is the average value for Lab A at milking 1 in Period 1, and in Eq.[3], the Lab_A parameter estimate represents the expected MUN value at milking 1 in Period 1 at a fat and protein content of zero. In both models, the fixed effects estimates indicate the amount by which MUN is expected to increase or decrease based on milking time and period of lactation. For example, in Eq. [2] milk samples collected during milking 1 in Period 2 are expected to have MUN values 0.391 mg/dL less than samples collected during milking 1 in P1. Likewise, samples collected during milking 3 of P3 are expected to have an average net difference in MUN values of 0.472 mg/dL compared to samples taken during milking 1 of P1.

The proportion of variance caused by the random effects for animal and date are similar between the two equations and indicate that approximately 30% of the variation is attributable to individual animals and only about 8% of is explained by variation from day to day. An additional 62% of total variance is attributed to residual random error that cannot be explained by either model but contains the variation associated with lab repeatability. The total variance including the random effect of animal, day, and residual error is 3.667 which equates to a standard error of MUN of 1.91 mg/dL. Thus, the random variation associated with MUN observations over multiple days would be expected to be within ± ~ 4 mg/dl. Removing the random effect of animal, the expected MUN variance of an individual animal across multiple days is 2.55 or a SE of 1.6 mg/dl. From a management perspective, observations varying more than ± 1.91 mg/dL for multiple

animals in a pen or 1.6 mg/dl for a single animal between days indicate that a significant change has taken place. For example, a MUN value of 10 mg/dL one week, followed by a diet change and an MUN of 9 mg/dL the following week, may indicate that the diet change had no significant impact on MUN.

Table 4. Parameter estimates of linear mixed model for the effects of lab, milking time, and period of lactation as described in Eq.[2]

	Estimate	SE	σ	Prop. of variance
Lab_A	7.90	0.349		
Lab_D	7.03	0.349		
$Milking_2$	0.289	0.0877		
$Milking_3$	-0.307	0.0886		
$Period_2$	-0.391	0.314		
$Period_3$	0.779	0.3151		
α		1.05	1.124	30%
δ		0.565	0.327	8.9%
σ		1.54	2.220	61.1%

Table 5. Parameter estimates of linear mixed model for the effects of lab, milking time, period of lactation, milk fat, and milk protein as described in Eq.[3]

	Estimate	SE	σ	Prop. of variance
Lab_A	10.5	0.797		
Lab_D	5.63	0.762		
$Milking_2$	0.351	0.0919		
$Milking_3$	-0.239	0.0934		
$Period_2$	-0.332	0.352		
$Period_3$	0.783	0.366		
β_{1A}	0.140	0.0980		
β_{1B}	-0.599	0.1180		
β_{2A}	-1.08	0.2702		
β_{2B}	2.15	0.196		
α		1.06	1.124	30%
δ		0.572	0.327	8.9%
σ		1.49	2.220	61.1%

Summary

In order to interpret reported MUN values, the precision and accuracy of the metric must be taken into account. The results presented here suggest that MIR analysis of MUN has improved since the 2017 report by Weeks and Hristov as the commercial labs that participated in this study were able to reproduce results of the enzymatic assay within ± 1.6 mg/dl. Further, commercial lab repeatability of MUN was high. However, the systematic bias revealed by the regression analysis indicates that there is still a need for improvement in MIR methods for this important milk component. Further, if MUN is to be used as a metric for management of individual animals, the metric must also be interpreted within the context of that animal's natural variation. Removing the MUN variation between animals, we found that the MUN of an individual cow would be expected to vary ± 1.6 mg/dL from day to day under similar dietary conditions and lactation period.

Take Home Messages

- Commercial lab repeatability for MIR analysis of MUN is ≤ 0.5 mg/dl which is lower than 5% of the average MUN value
- Commercial lab MIR reproducibility of the gold standard method for measuring MUN is ≤ 0.8 mg/dl which means reported MIR values are expected to be within 1.6 mg/dl of the true value
- Current MIR methods for analysis of MUN tend to over predict MUN at values below 12.8 mg/dl and under predict MUN at values above 12.8 mg/dl
- The MUN content of samples from an individual cow is varies between days so repeated samples from the same cow across multiple days within the same stage of lactation would be expected to vary ± 1.6 mg/dL

References

Barbano, D. M. and Coon, C. M. (2017). Determination of Urea and Urea Nitrogen Content of Milk (MUN) by the Enzymatic Method – Megazyme Urea/Ammonia Kit CAT#K-URAMR.

Gustafsson, A. H., and D. L. Palmquist. 1993. Diurnal variation of rumen ammonia, serum urea, and milk urea in dairy cows at high and low yields. J. Dairy Sci. 76:475–484.

Kaylegian, K. E., J.M. Lynch, G.E. Houghton, J.R. Fleming, and D.M. Barbano. 2006. Modified Versus Producer Milk Calibration: Mid-Infrared Analyzer Performance Validation. J. Dairy Sci. 89: 2833-2845

Lynch, J. M. 1998. Use of AOAC International method performance statistics in the laboratory. Journal of AOAC International. 81: 679-684.

Weeks, H. L. and A. N. Hristov. 2017. Short communication: Analytical methods and amount of preservative added to milk samples may alter milk urea nitrogen measurements. J. Dairy Sci. 100:1502-1506.