Improvement in Precision and Accuracy of Mid-Infrared Analysis of Milk Urea Nitrogen and Individual Animal Variation

Kristan F. Reed[1], Emma M. Wood, Mathilda Portnoy, and Dave M. Barbano
Department of Animal Science
Cornell University

Abstract

Milk urea nitrogen (**MUN**) from bulk tank samples is often tracked by nutritionists and farm managers as an indicator the status of the herd's protein nutrition. This study presents an evaluation of the precision and accuracy of the use of infrared spectroscopy to measure MUN and as a preliminary assessment of the expected variation within and between cows. For the first objective, we sent 4 sets of bulk tank samples to commercial labs for spectroscopic analysis and used an enzymatic assay as a gold-standard method for measurement of MUN. Using reproducibility as an indicator of accuracy, we found the commercial lab analyses were expected to be within 1.6 mg/dL of the enzymatic assay, or true value, with very little differences between labs. We also found the commercial lab repeatability to be high with an expected coefficient of variation of repeated samples close to 3% for most labs. To achieve our second objective, we followed 16 cows through 3 periods of lactation (early – average DIM = 40; middle – average DIM = 140; and late – average DIM = 240), collecting milk samples in duplicate from each of the 3X milkings for 7 consecutive days. Milk samples were sent to 2 labs for spectroscopic analysis of MUN and the results were fit to a linear mixed model with a continuous autoregressive correlation covariance structure. We found that both milking time and period of lactation have significant effects on MUN and that milk fat and protein composition influences the spectroscopic estimate of MUN differently depending on the lab.

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 are often employed on farms and result in reduced amounts of manure N produced per unit of milk (Chase, 2018). Milk urea nitrogen is commonly used as an indicator to guide and provide feedback about the effectiveness of nutrition strategies to manage nitrogen efficiency on farms because the statistical relationships between MUN and urinary urea nitrogen have been well-studied and widely reported (Gustafsson and Palmquist, 1993; Spek et al., 2013a, 2013b). Physiological factors beyond nitrogen intake and efficiency also influence MUN and the relationships between MUN, nitrogen intake, and urinary urea nitrogen

[1]Contact at: 272 Morrison Hall, Ithaca, NY 14850, (530) 402-5682 Email: kfr3@cornell.edu

(Spek et al., 2013a, 2013b; Barros et al., 2019), but nitrogen intake and efficiency remain the dominant factors driving MUN levels. This fact, combined with accessibility of this metric on commercial farms, has led to widespread use of MUN as an on-farm metric to assess nitrogen efficiency. In addition to its application as an indicator of protein efficiency, some recent work (Albaaj et al., 2017; Raboisson et al., 2017) suggests that high MUN may also be an indicator of reproductive performance. Further, the multiple connections of MUN to dairy cattle performance have led others to evaluate the genetic variation and heritability of MUN to assess the potential for its incorporation in breeding programs (Beaston et al., 2019; Bobbo et al., 2020).

Unfortunately, previous research has highlighted variation in both precision and accuracy of commercial MUN testing, raising questions about the utility of this metric. Weeks and Hristov (2017) found that average reported MUN from same set of bulk tank samples sent to multiple labs ranged from 6.5 to 14.9 mg/dL, which is wider than the commonly recommended target for MUN of 8 to 12 mg/dL (Chase, 2018). Most commercial labs use indirect, spectrophotometric methods for milk analysis because it is faster and cheaper than traditional wet-chemistry methods. In addition, the most common application of spectrophotometric analysis of milk composition restricts the spectral analysis to the mid-infrared range and we, therefore, refer to this method as mid-infrared (**MIR**) spectroscopy. Precision and accuracy of MIR spectroscopy depend on many factors, including the calibration methods, milk sample composition, and internal instrument factors (Kaylegian et al., 2006).

Thus, the first objective of this study was to reevaluate the precision and accuracy of MIR spectroscopy for MUN analysis. In addition to the knowledge of the accuracy of the lab analysis method, interpretation of MUN values requires an understanding of expected variation in MUN over time. Dairy farms often measure MUN during routine bulk tank sampling, and the extent and sources of variation in bulk-tank MUN has been studied (Hristov et al., 2018; Siachos et al., 2019). However, the movement towards precision management and advance of technologies that enable daily monitoring of individual cow performance has increased interest in use of MUN to inform individual and pen-level decisions. The potential to use routine or daily MUN measurements from individual animals and the interest in use of MUN to inform breeding programs require an understanding of expected variation in MUN between animals and within animals over time. Therefore, the secondary objective of this study was to quantify the expected variation in MUN over the course of lactation in individual cows.

Methods

Bulk tank sample collection and 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°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 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 to meet the USDA's standards for milk payments which are based on milk fat and protein contents. 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 3 sets of the FFMA samples in duplicate (i.e. 20 samples each) on 3 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 MUN variability

Milk samples were collected from 16 multiparous Holstein cows at each of 3X daily milkings during 3, 7-day periods in early (average DIM = 40; P1), middle (average DIM = 140; P2), and late-lactation (average DIM = 240; P3). All samples were collected in duplicate and one set was sent to Labs A and D for MIR analysis of MUN and milk components.

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 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 to 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. Samples were stored at 40°F before delivery to Labs A and D. Lab A used the Milkoscan FT+ and Milkoscan 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 lme() package and all other functions were performed using base packages.

Precision and accuracy of MIR MUN

We used statistical methods for evaluation of methodological agreement described by Lynch (1998) and used by Kaylegian et al. (2006) for analysis of bulk tank MIR MUN. The mean difference was calculated by subtracting the reference chemistry value from the MIR spectroscopy value and averaging the differences over the sample sets. We calculated the standard deviation of the difference as the square root of the summed squared value of the differences divided by the number of samples and the Euclidean distance as the distance from the origin of the points when the standard deviation of the differences was plotted against the mean difference (Figure 1). We calculated a coefficient of variation by dividing the SDD divided by the mean reference chemistry MUN for each sample set. We estimated repeatability (sr) 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. Finally, we estimated reproducibility (sR) 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.

We also fit bulk tank MUN observations to a linear mixed model:

$$MUNDiff_{ij} = \beta_0 + \beta_1 MUNRefC_{ij} + \beta_2 Prot_{ij} + \beta_3 Fat_{ij} + \lambda_i + \sigma_{ij} \quad [1]$$

In Eq. [1], MUNDiffij 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; $MUNRefC_{ij}$ 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

To understand how MUN varies within and between animals over time, we fit the individual animal MUN data from all 3 periods to the following models:

$$MUN_{ijklm} = Lab_i + Milking_j + Period_k + \alpha_l + \delta_m + \varepsilon_{ijklm} \quad [2]$$

$$MUN_{ijklm} = Lab_i + Milking_j + Period_k + \beta_1 Milk._{jklm} + \beta_{2i} Fat_{ijklm} + \beta_{3i} Protein_{ijklm} + \alpha_l + \varepsilon_{ijklm} \quad [3]$$

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 day. Lab was included as a fixed effect to accommodate differences in MIR machines and analysis processes 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 2 factors due to short sampling periods. Instead, period is included as a fixed effect that includes the stage in lactation, changes in diet due to feed variability, and external factors like weather and pen location. Milk production ($Milk\bullet_{jklm}$) is also included as a fixed effect.

In Eq. [3], the dependent variable is the same, but the fixed effects include parameters to estimate the lab-specific effects of fat and protein contents on reported MUN concentration. In Eq. [3], β_{2i} is the effect of milk fat corresponding to the *i*th lab and β_{3i} 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 contents impacted the MIR difference from enzymatic MUN measurements. The parameters act as a correction factor and should therefore remove any effect caused by fat and protein interference with MIR analysis.

In both models, α_l is the random effect of the *l*th animal and ε_{ijklm} is the residual random error modeled with a continuous auto-regressive correlation structure to account for dependency of repeated observations within animal that are unequally spaced in time.

Results and Discussion

Precision and accuracy of MIR MUN analysis

Figure 1 presents the plot of the standard deviation of the differences vs. the mean differences and the Euclidean distances, 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 mean distance = 0), which suggests that the MIR results from this machine within this lab systematically underestimate MUN.

Repeatability (**sr**) and reproducibility (**sR**) estimates for commercial labs are shown in Table 2 and ranged from (0.297 to 0.469) and (0.555 to 0.791), respectively. We can interpret repeatability 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 2 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 a 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 indicates the percent of the mean MUN value by which repeated and reproduced analyses are expected to vary. Differences in sr and sR across labs are most likely due to the use of different machines and calibration methods used for different machines.

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.

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 we consider 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, we would expect the MIR analysis to underpredict MUN by 0.618 mg/dL. Similarly, as the reference chemistry decreases below the average of 12.8 mg/dL, we expect the MIR analysis to over predict MUN concentration. For a milk sample with 3.4% protein, 4.2% fat, and 7.8 mg/dL MUN, we expect the MIR analysis to over predict MUN by 1.3 mg/dL. In addition, as protein and fat levels deviate from the means of this dataset, we expect a systematic over or under prediction of MIR MUN analysis, 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 the difference between MIR and enzymatic analysis of MUN, 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. [2] and [3] are listed in Tables 4 and 5. We estimated a separate intercept for Labs A and D, which represents the average MUN value during P1 and Milking 1 for each lab. For example, for Eq.[2] the Lab_A parameter estimate is the average value for Lab A at Milking 1 in P1, and in Eq.[3], the Lab_A parameter estimate represents the expected MUN value at Milking 1 in P1 at the average fat and protein contents. 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 P2 were 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 were expected to have an average net increase 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 2 equations and indicate that between animal variation accounts for approximately 35%. The remaining 65% 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 estimate for the correlation between observations from an individual animal over time ($\rho^{m,m'}$) is close to 0.15 mg/dL for both models. The continuous autocorrelation structure accounts for the different length of time between observations in this dataset so that expected correlation of observations from consecutive days is 0.15 mg/dL, but the correlation of observations from different periods that were, for example, 100 days apart is equal to ($\rho^{m,m'}$)100 or 4.07e-83 mg/dL; i.e. there is essentially no observed correlation of MUN observations between different periods. In fact because the estimate of $\rho^{m,m'}$ is quite small to begin with, the expected correlation of MUN observations from the same animal decays to zero quickly with an expected correlation of only 0.003 mg/dL when observations are 3 days apart.

The total variance including the random effect of animal, day, and residual error is 3.25, which equates to a standard error of MUN of 1.80 mg/dL.Thus, the random variation associated with MUN observations from a single pen, over multiple days would be expected to be within $\pm \sim 4$ mg/dL. Removing the random effect of animal, the expected MUN residual variance of an individual animal across multiple days is 2.09 or a SE of 1.45 mg/dL. From a management perspective, observations varying more than ± 1.8 mg/dL for multiple animals in a pen or 1.45 mg/dL for a single animal between days are a sign 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, would not indicate that the diet change had a significant impact on MUN.

Conclusion

To interpret reported MUN values, we must take the precision and accuracy of the metric into account. The results presented here suggest that MIR analysis of MUN has improved since the 2017 report by Weeks and Hristov (2017) 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 to remove the effect of milk fat and protein composition that varies by lab. 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 ± 1.45 mg/dL from day to day with an small expected correlation between days under similar dietary conditions.

References

Albaaj, A., G. Foucras, and D. Raboisson. 2017. Changes in milk urea around insemination are negatively associated with conception success in dairy cows. J. Dairy Sci. 100:1-9.

Barbano, D.M., and C.M. Coon. 2017. Determination of urea and urea nitrogen content of milk (MUN) by the enzymatic .ethod – Megazyme Urea/Ammonia Kit CAT#K-URAMR.

Barros, T., K.F. Reed, J.J. Olmos Colmenero, and M.A. Wattiaux. 2019. Short Communication: Milk urea nitrogen as a predictor of urinary nitrogen and urea nitrogen excretions of late-lactation dairy cows fed nitrogen-limiting diets. J. Dairy Sci. 102:1601-1607.

Beaston, P.R., S. Meier., N.G. Cullen, and H. Edning. 2019. Genetic variation in milk urea nitrogen concentration of dairy cattle and its implications for reducing urinary nitrogen excretion. Animal. 13: 2164-2171.

Bobbo, T., M. Penasa, A. Rossoni, and M. Cassandro. Short Communication: Genetic aspects of milk urea nitrogen and new indicators of nitrogen efficiency in dairy cows. J. Dairy Sci. 103: 9207-9212.

Chase, L.E. 2018. Precision Feed Management – What have we learned? Proceedings of the Cornell Nutrition Conference (Syracuse, NY 2018). Accessed on Dec 12, 2020: https://hdl.handle.net/1813/59851

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.

Hristov, A.N., M. Hanigan, A. Cole, R. Todd, T.A. McAllister, P.M. Ndegwa, and A. Rotz. 2011. Review: Ammonia emissions from dairy farms and beef feedlots. Can. J. Anim. Sci. 91:1-35, doi:10.4141/CJAS10034.

Hristov, A.N., M. Harper, J. Oh, F. Giallongo, J.C. Lopes, G. Cudoc, J. Clay, R. Ward, and L.E. Chase. 2018. Short Communication: Variability in milk urea nitrogen and dairy total mixed ration composition in the northeastern United States. J. Dairy Sci. 101: 1579-1584.

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.

Raboisson, D., A. Albaaj, G. Nonne, and G. Foucras. 2017. High urea and pregnancy or conception in dairy cows: A meta-analysis to define the appropriate urea threshold. J. Dairy Sci. 100:7581-7587.

Siachos, N., N. Panousis, and G.E. Valergakis. 2019. Association of bulk tank milk urea nitrogen concentration with elevated individual cow values and investigation of sampling frequency for accurate assessment. Tropic. Anim. Health and Prod. 51:2431-2436.

Spek, J.W., J. Dijkstra, G. Van Duinkerken, and A. Bannink. 2013a. A review of factors influencing urea concentration and its relationship with urinary urea excretion in lactating dairy cattle. J. Ag. Sci. 151: 407-423.

Spek, J.W., J. Dijkstra, G. Van Duinkerken, W.H. Hendriks, and A. Bannink. 2013b. Prediction of urinary nitrogen and urinary urea nitrogen excretion by lactating dairy cattle in northwestern Europe and North America: A meta-analysis. J. Dairy Sci. 96: 4310-4322.

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.

Table 1. Euclidean Distance of mid-infrared spectroscopy 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

Table 2. Repeatability (sr) and reproducibility (sR) 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

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

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

	Estimate	SE	σ^2	Proportion of Variance
Lab_A	9.22	0.435		
Lab_D	8.38	0.4.35		
$Milking_2$	0.202	0.0568		
$Milking_3$	-0.455	0.0646		
$Period_2$	-0.559	0.152		
$Period_3$	0.437	0.167		
β_1	-0.104	0.0196		
α		1.06	1.124	30%
σ		1.48	2.220	61.1%
$\rho^{m,m'}$		0.155		

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

	Estimate	SE	σ^2	Proportion of Variance
Lab_A	13.2	1.01		
Lab_D	8.55	1.01		
$Milking_2$	0.138	0.0643		
$Milking_3$	-0.541	0.0719		
$Period_2$	-0.111	0.212		
$Period_3$	0.980	0.239		
β_1	-0.104	0.0196		
β_{2A}	0.467	0.0834		
β_{2D}	-0.337	0.109		
β_{3A}	-1.92	0.315		
β_{3D}	1.743	0.289		
α		1.08	1.16	35%
σ		1.45	2.10	65%
$\rho^{m,m'}$		0.143		

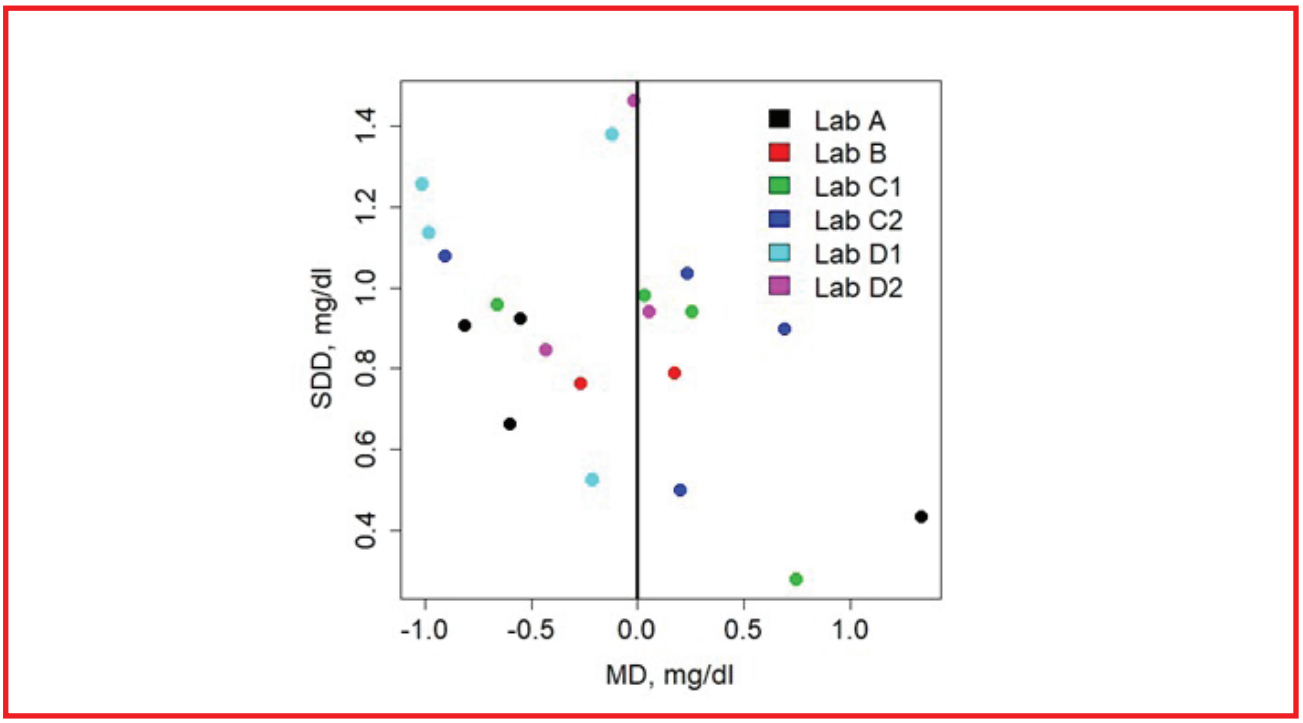

Figure 1. A plot of the standard deviation of the difference (SDD) vs mean difference (MD) for each lab (A-D) with labs C and D reporting results for two different machines.

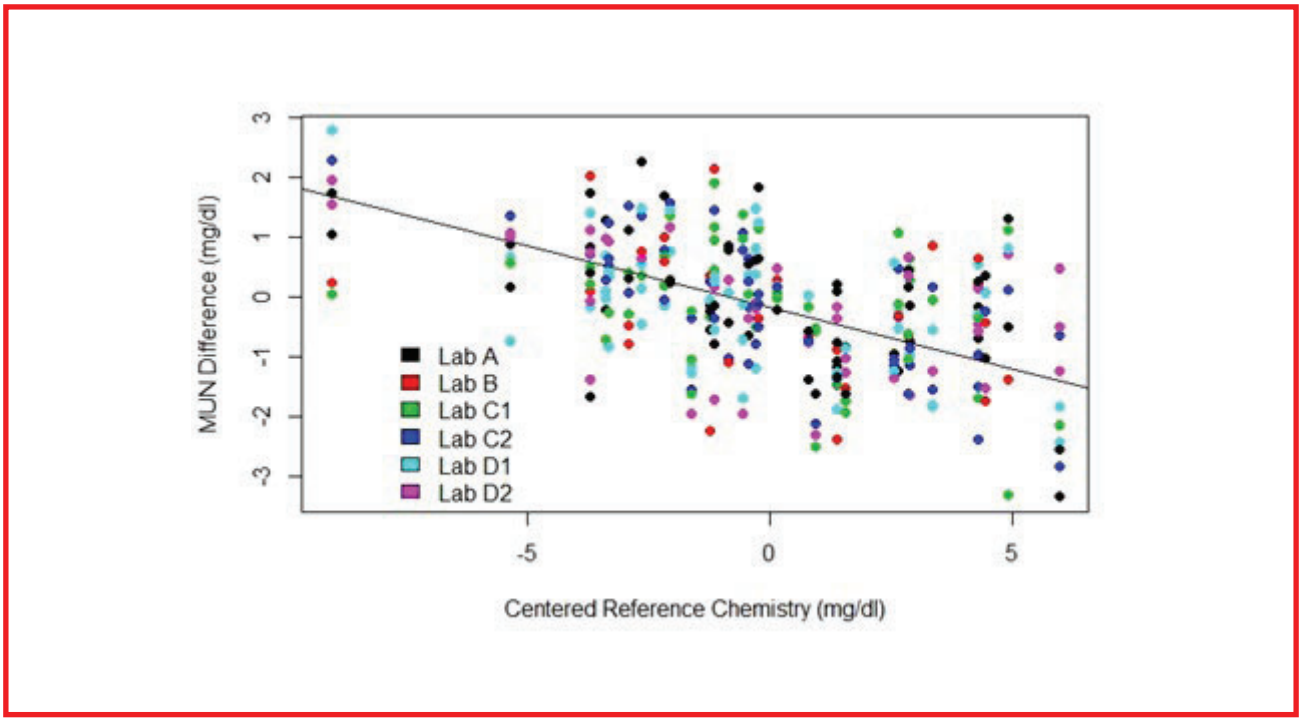

Figure 2. A plot of the differences between the mid-infrared spectroscopy and reference chemistry for MUN analysis vs. the centered reference chemistry value. The line represents the fixed-effect results of the mixed-model regression.