Tso et al., Metabolomics 2015, 5:3
DOI: 10.4172/2153-0769.1000151

Metabolomics: Open Access

Research Article Open Access

Consistency of Metabolite Determination from NMR Spectra over Time and Between Operators

Tso V[1,4]*, Eisner R[1,4], Macleod S[2], Ismond KP[1,4], Foshaug RR[1,4], Wang H[3,4], Joseph R[4], Chang D[4], Taylor N[5] and Fedorak RN[1]

[1]*Department of Medicine, University of Alberta, Edmonton, AB, Canada*
[2]*Department of Oncology, University of Alberta, Edmonton, AB, Canada*
[3]*Department of Surgery, University of Alberta, Edmonton, AB, Canada*
[4]*Metabolomic Technologies Inc, Edmonton, AB, Canada*
[5]*Chenomx Inc, Edmonton, AB, Canada*

Abstract

With the recent increase in publications of metabolomic-based studies, it is apparent that there are considerable differences in findings between laboratories. Even within a single centre, variability introduced by different highly skilled operators contributing to the same experiment can alter the results. These inconsistencies may contribute to unforeseen experimental confounders and, as such, represent a critical barrier to metabolomics based studies. This issue can impact studies throughout the research spectrum, from basic science to clinical research. One of the potential sources of unforeseen variance in many metabolomic-based studies occurs during the identification and quantification of metabolites. In our study, we examined the profiling consistency of NMR spectra as essential to the interpretation and analysis of metabolite concentrations. This manuscript describes how to evaluate the consistency of metabolite identification and quantification from NMR spectral profiling, between operators and over time, and this methodology can be applied to other quantification techniques in the 'omics' realm. We also present the results of our analysis, ranking the urine metabolites' consistency via NMR quantification. Results from this study will allow metabolomics researchers to better assess consistency and thus ensure that metabolite differences truly reflect biological differences rather than experimental variability.

Keywords: Metabolomics; Metabolite quantification; Quality control; Urine; Spectroscopy; Confounders; Operators; Chenomx; Batching; Profiling NMR spectra

Abbreviations: HMDB: Human Metabolome Database; MS: Mass Spectrometry; NMR: Nuclear Magnetic Resonance

Introduction

Profiling metabolites using spectra acquired by either nuclear magnetic resonance (NMR) or mass spectrometry (MS) is complex and requires a highly trained and skilled operator. This process is made more consistent by using software to identify spectra of known metabolites (e.g., Chenomx NMR Suite) [1]. Importantly, employing proper experimental controls ensures the consistent acquisition of NMR spectra regardless of instrument used or laboratory site [2-4].

However, there are many sources of variation in the determination of metabolite concentration from spectral profiling. For example, the stability of metabolites in urine samples varies depending on storage conditions emphasizing the need for consistent sample handling [5-9]. A more recent study has highlighted the importance of parameter settings on solvent peak suppression [1]. Other studies have shown the importance of the normalization method in analysis of NMR spectra, which can depend on the sample being studied, and on the signal to noise ratio of the peaks [10,11]. In long-term studies, differences in sampling due to batch effect have been reported which can be controlled by calibration based on either quality control samples or sample randomization [12]. Another option to control for batch effects relies upon normalization using a known cellular response to an external agent (e.g., a pharmaceutical) [13]. Particular metabolites have also been shown to be difficult to analyze. Several studies have shown that there is more variability seen with lower metabolite concentration [14,15]. Furthermore, NMR spectral analysis has also been shown to be more difficult in highly complex areas, where multiple metabolites overlap [14-16]. It is also well known that pH changes can cause some metabolite peaks to shift.

In analysis of NMR spectra, well trained people are not highly variable (within themselves), but there is a potential for variability when different operators are used for analysis. One study has shown similar analytical results between three analysts examining four samples for nine metabolites [14]. Another study used five people with multiple rounds of analysis of 18 spectra, to demonstrate good agreement between people for the most common metabolites [16]. Furthermore, the experimental group (sampling method) was generally more important than variation between people [16]. Despite the consistencies seen between people, it is still recommended that analysis be performed by the same person or group over a single study [16]. However, for studies with a large number of samples, or long-term experiments, single operator analysis may not be feasible.

In this study, metabolite spectra of urine samples acquired using NMR were assessed for consistency of metabolite quantification either over time (3 years) or between two groups of operators. The aim was to identify the metabolites that are reliably quantified to improve our standard operating protocols and experimental procedures.

We first performed a subjective analysis of 70 metabolites, categorizing them based on their apparent consistency in metabolite concentration over time. We also looked at the difference between the

***Corresponding author:** Tso V, 7-120 Katz Building, 11315 87 Ave, University of Alberta Campus, Edmonton, AB, Canada, Tel: 780-492-2925; Fax: 780-492-9271; E-mail:vtso@ualberta.ca

Received August 17, 2015; **Accepted** September 15, 2015; **Published** September 17, 2015

Citation: Tso V, Eisner R, Macleod S, Ismond KP, Foshaug RR, et al. (2015) Consistency of Metabolite Determination from NMR Spectra over Time and Between Operators. Metabolomics 5: 151. doi:10.4172/2153-0769.1000151

analyses of two skilled operators with the same training, on 60 NMR spectra. Additionally, we re-profiled the same NMR spectra two or three times over the course of three years. This allowed us to identify metabolites that may be difficult to quantify over time, and likely represents a change in policy and/or change in the skill level of the profiler.

Material and Methods

Study participants and urine collection

This study used NMR spectra acquired from 988 urine samples previously obtained as part of a regional colon cancer screening program in Edmonton, Canada (SCOPE®, Stop Colorectal Cancer through Prevention and Education) [17-19]. Study participants of average or increased colorectal carcinoma risk were recruited. On day of entry, participants provided informed consent, a urine sample; and also completed a demographic survey. Participants were excluded if they were under 40 or over 75 years of age or had findings of colonic or ileal disease at the time of colonoscopy. Ethics approval was obtained from the Health Research Ethics Board at the University of Alberta. The www.ClinicalTrials.gov identifier is NCT01486745.

NMR spectra acquisition

Spectra were collected using a 600 MHz NMR spectrometer (Oxford Instruments, Oxfordshire, UK) with a VNMRS two-channel console (Varian Inc., Palo Alto, CA, USA) running VNMRJ software version 2.2C on a RHEL 4 (Red Hat) host computer. The spectrometer was equipped with an HX probe with Z-axis gradients. The first increment of a 2D-^{1}H, ^{1}H-NOESY pulse sequence was utilized for the acquisition of ^{1}H-NMR data and for suppressing the solvent signal. Experiments used a 100 ms mixing time along with a 990 ms pre-saturation (~ 80 Hz gamma B_1). Spectra were collected at 25°C for a total of 32 scans over a period of 3.5 min.

NMR metabolite quantification subsets used for consistency analysis

The metabolite quantification was first performed in 2010 using the targeted profiling techniques of Chenomx NMR Suite v7.7 (Chenomx, Inc., Edmonton, Canada) and is labeled original spectra [18]. The quantification process was completed at Chenomx by a group of operators consisting of 1 or 2 trained staff members and verified by a separate staff member.

Subsequently, using the same NMR spectra, re-quantification was carried out as part of the consistency assessment (Figure 1). Identical Chenomx protocols were used in the re-quantification. First, in March 2013, 100 samples were re-quantified and labeled Subset 1 (Figure 1A). Second, in September 2013 the remaining 888 samples were re-quantified and labeled Subset 2 (Figure 1A). Third, in December 2013, a further 60 samples from Subset 1 were re-quantified and labeled Subset 3 (Figure 1B). Fourthly, in December 2013, 60 samples from Subset 2 were re-quantified independently by two operators and labeled Subset 4 (Figure 1C). Named metabolites are also identified using the HMDB ID from the Human Metabolome Database, version 3.6 (www.hmdb.ca).

Metabolite consistency over time: Subjective comparison

From the same NMR spectra we compared the original 2010 metabolite quantification data to that of the 2013 metabolic quantification data from Subsets 1 and 2 in order to examine the consistency of the metabolite quantification between the two time points. The concentration of each metabolite determined in 2010 was plotted against that determined for the same sample in 2013. Thus X-Y plots of metabolite concentrations were created containing either 100 (Subset 1) or 888 (Subset 2) points. These graphs were examined and subjectively classified into one of four consistency groups (Excellent, Good, Fair, and Poor) by two authors (VT and RE) independently. Differences in categorization were resolved by re-examination and consensus.

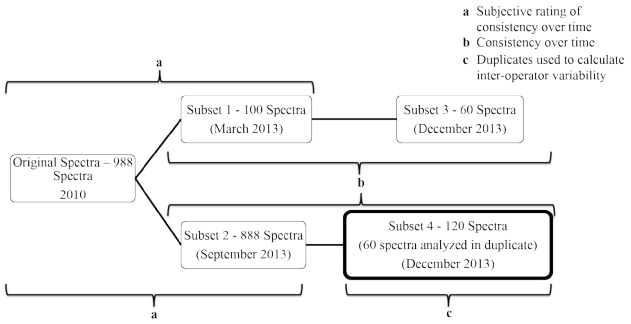

Figure 1: Diagram of spectral analysis timing and relationship to subsets. The original data was profiled in 2010 (988 spectra). Subset 1 consists of 100 spectra originally profiled in 2010, re-profiled in March 2013. Subset 2 consists of 888 spectra originally profiled in 2010, re-profiled in September 2013. Both Subset 3 and Subset 4 contain 60 spectra profiled a third time in December 2013. The Subset 3 spectra were all drawn from within Subset 1, while Subset 4 was drawn from within Subset 2. (a) Both Subset 1 and Subset 2 are used in the subjective rating of consistency over time. (b) The comparisons of Subset 1 to Subset 3 and Subset 2 to Subset 4 represent the quantitative consistency over time analysis. (c) The 60 spectra from Subset 4 were profiled in parallel by two different operators, and these duplicates were used to calculate inter-operator variability.

Metabolite consistency over time: Objective comparison

We further examined consistency over time using Subsets 3 and 4 from December 2013 (Figure 1B) and comparing Subsets 1 vs. 3 and 2 vs. 4. An X-Y plot was created for each metabolite, with the concentration determined in March 2013 or September 2013 plotted against the concentration determined at December 2013. The difference between the concentrations for each was analyzed by determining the p value, average difference between the two values, absolute average difference, and correlation coefficient. The p value was calculated using the paired Mann-Whitney test in R. For average difference, the difference between the two concentrations assigned for a particular metabolite in a single sample was calculated and these values were averaged over all samples for a particular metabolite. The absolute average difference is the average of the absolute value of the differences. Spearman's correlation coefficient (r_s) was chosen since this makes no assumptions about the distribution of values. We considered r_s values below 0.8 as "inconsistent", which is more stringent than the minimum value. Finally, the fraction of zeros assigned to each metabolite was calculated from the number of zeros assigned to the metabolite divided by the total number of samples profiled. We determined the difference between the fraction of zeros for each group of operators. A metabolite was considered "inconsistent" if the difference in the fraction of zeros was greater than 10 percentage points.

Inter-operator metabolite quantification consistency

Using Subset 4 (December 2013, Figure 1C), the 60 NMR spectra were analyzed to determine metabolic concentrations by two operators independently. An X-Y plot was created for each metabolite, with the concentration determined by one operator group plotted against the

concentration determined by the other operator group. We used the same difference measures and fraction of zeros to compare time points as was used to compare the consistency of metabolite quantification in two time points separated by 3 years. Similarly, a metabolite was considered "inconsistent" if the correlation coefficient was below 0.8 or if the difference in fraction of zeros was greater than 10 percentage points.

Results and Discussion

If the correlation is relatively low, there is a possibility that this metabolite is difficult to profile. Metabolites that are not consistently analyzed may adversely affect modeling performed on the metabolite concentrations. For both inter-operator and over time experiments, we quantified the number of times metabolites were assigned the concentration of zero. This can result from either a concentration that is lower than the limit of detection, or because a group of operators cannot separate the peak from the surrounding ones (a highly complex region). Both of these factors have been shown to increase variability in metabolite concentration previously [14-16], but not by a quantification of zeros. Furthermore, our study included larger sample sizes than many previous studies. The most consistent metabolites determined here would likely be the most robust for use in clinical models based on metabolite concentration in urine, and thus be useful for future diagnostic studies. Likewise, the metabolites identified as difficult to quantify in these analyses may not be as useful for the construction of models from this data. For studies that are concerned with the accuracy of the metabolite concentrations (e.g., trying to establish the normal range for a metabolite), the confidence in quantification is crucial for knowing how much these measures can be trusted.

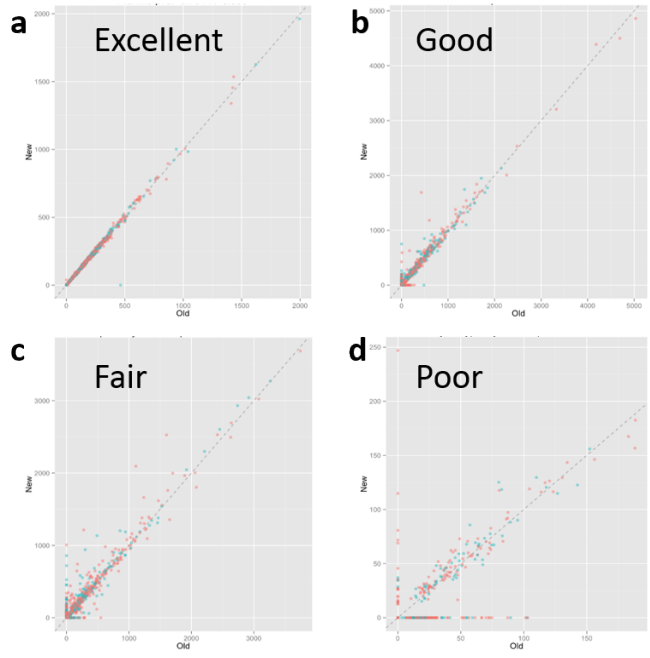

Figure 2: Illustrative plots for each of the four subjective categories. (**a**) Excellent (L-alanine), (**b**) Good (taurine; HMDB00251), (**c**) Fair (p-methylhistidine; HMDB00001), and (**d**) Poor (3-hydroxyphenylacetate; HMDB00440)) plots of metabolite concentrations at two different time points. Concentrations on the y-axis (New) are from 2013, while concentrations on the x-axis (Old) are from 2010.

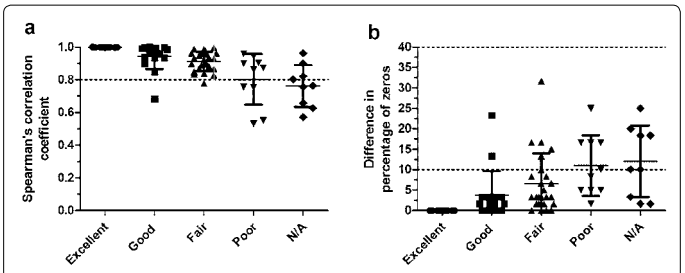

Figure 3: Correlation and difference in percentage zeros seen with the comparison of inter-operator variability in the subjective groups. (**a**) Spearman's correlation coefficient (r_s) was calculated from plots of metabolite concentrations determined from analysis done by two different operators. (**b**) Difference in percentage zeros between the two operators. The grey shaded area in represents metabolites that are identified as potentially variable in all graphs (**a**) r_s<0.8 and (**b**) difference in percentage of zeros >10. The metabolites rated subjectively in Table 1 retain their grouping here. N/A is Not Analyzed. Error bars represent standard deviation of the values.

Subjective Rating	Number of Metabolites	Percent of Total Metabolites Analyzed
Excellent	9	13%
Good	17	24%
Fair	25	36%
Poor	10	14%
Not Analyzed	9	13%
Total	70	100%

Table 1: Number of metabolites and percent of total metabolites analyzed for the subjective rating of consistency over time. The metabolite categories are consistent with the example plots seen in Figure 2.

Consistency over time: Subjective

The metabolite concentrations from the original NMR spectra (2010) were compared to that from Subsets 1 and 2 (2013) to examine the consistency of the metabolite quantifications between the two analyses conducted 3 years apart. Figure 2 shows example plots of each of the four consistency categories (Excellent, Good, Fair, and Poor). This method is most effective for metabolites with large sample numbers, since trends are visually more obvious. As this analysis required larger data sets, the subjective ratings used different data sets than the objective analysis described in the next section (Supplementary Table S1). Metabolites not analyzed at both time points were not examined by this method (n=9), though they were examined in the following objective analyses. Results were not controlled for differences between the 2 groups of operators (Figure 3).

The consistency of 61 metabolites was categorized as: 9 Excellent, 17 Good, 25 Fair, and 10 Poor (Table 1). Only 9 metabolites were not categorized because they were only profiled at one of the time points (Supplementary Figure S1). This categorization was used as a comparison for our later, more objective, analyses. The most consistent metabolites (rated Excellent) are listed in Table 2.

Consistency over time: Objective

We compared the metabolite concentrations determined at two different times using Subsets 1 vs. 3 and 2 vs. 4 (Figure 1B). In this part of the study, the same group of operators was ensured across a one year span. We also examined the use of p values, average difference and average absolute difference to differentiate more consistent metabolites (Supplementary Figure S2). However, these were not used in future calculations.

Correlation coefficient: Spearman's correlation coefficient (r_s) was

used as a measure of consistency in metabolite concentrations over time. There is less consistency in the first comparison, between Subsets 1 and 3, compared to the second, between Subsets 2 and 4 (Figure 4A and 4B). In both comparisons, the r_s values for Excellent metabolites was very close to 1, while the r_s for Good metabolites was greater than 0.8; the exception being the metabolite ethanol in Subsets 2 vs. 4 (Figure 4B, Supplementary Table S2). The r_s results for Fair metabolites were greater than 0.8 in both comparisons albeit there were five metabolites with r_s less than 0.8 in the comparison between Subsets 1 and 3. Most of these metabolites were close to 0.8, except for benzoate. Poor metabolites have r_s results that were below 0.8 in Subsets 1 vs. 3, while only three were below 0.8 in the comparison of Subsets 2 and 4. Of the unrated metabolites, only two were above 0.8 in Subsets 1 vs. 3, while myo-inositol (HMDB00211) and acetoacetate (HMDB00060) had particularly poor correlations. Conversely, metabolites from Subsets 2 and 4, including myo-inositol, were not identified as inconsistent. Most of the metabolites identified as inconsistently quantified in Subsets 2 vs. 4 were similarly identified in Subsets 1 vs. 3, however more metabolites were identified as inconsistent in the Subsets 1 vs. 3. These results illustrate that correlation is a good but not perfect measure of quantification consistency.

Difference in percentage of zeros: We also investigated the difference in percentage of zeros in Subsets 1 vs. 3, and Subsets 2 vs. 4 as another measure of consistency of metabolite quantification. Some metabolites were often classified as zero, but these were not considered to be inconsistent unless the percentage of zeros assigned from a time point differed from the comparison time point. For example, asparagine (Fair; HMDB00168) had about 75% zeros in both Subsets 2 and 4 (no difference), while it was identified as inconsistently quantified in Subsets 1 vs. 3, because Subset 1 and Subset 3 had a large difference in percentage of zeros (45% vs. 12%, respectively).

We saw 37 differences in the percentage of zeros in Subsets 1 vs. 3 (Figure 4C) and only 6 in Subsets 2 vs. 4 (Figure 4D, Supplementary Table S2). Two of the metabolites identified as inconsistent in Subsets 2 vs. 4 were not identified in Subsets 1 vs. 3: mannitol (Fair; HMDB00765) and methanol (Poor; HMDB01875). However, the differences in Subsets 1 and 3 were 8.3 and 6.7 percentage points, respectively. Metabolites rated Excellent had few differences in percentage of zeros, while some metabolites in the other three classification groups did. Most of the metabolites categorized as either Poor or not rated had many differences in zeros. Two of the metabolites with the largest difference in percentage zeros were ethylmalonate (Poor, 60 percentage points; HMDB00622) and myo-inositol (not rated, 63 percentage points).

Metabolite (HMDB ID)	Inter-operator group		Subset 1 vs. 3		Subset 2 vs. 4	
	r_s	zero	r_s	zero	r_s	zero
Creatinine (HMDB00562)	1	0	0.999	0	1	0
Dimethylamine (HMDB00087)	0.999	0	0.999	0	1	0
Hippurate (HMDB00714)	0.999	0	0.999	0	0.999	0
Glycine (HMDB00123)	0.998	0	0.998	0	0.999	0
L-Alanine (HMDB00161)	0.999	0	0.995	0	0.999	0
Trigonelline (HMDB00875)	0.999	0	0.972	0	0.999	0
Trimethylamine N-oxide (HMDB00925)	0.999	0	0.998	0	0.998	0
Formate (HMDB00142)	0.997	0	0.994	1.67	0.997	0
Succinate (HMD00254)	0.996	0	0.987	0	0.997	0

Table 2: A list of the most consistent metabolites determined in the subjective analysis (rated Excellent) and the values for Spearman's correlation coefficient (r_s) and the percentage difference in zero fraction for the inter-operator and over-time experiments (Subset 1 vs. 3 and Subset 2 vs. 4).

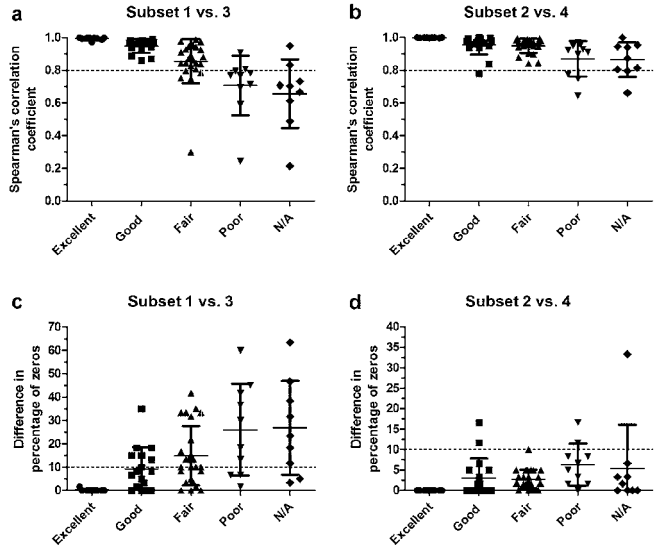

Figure 4: Correlation and difference in percentage zeros seen with comparisons over time in the subjective groups. (**a** and **b**) Spearman's correlation coefficient (r_s) was calculated from plots of metabolite concentrations determined from analysis done at two different time points. (**a**) Subsets 1 vs. 3, (**b**) Subsets 2 vs. 4. (**c** and **d**) Difference in percentage zeros between the two time points. (**c**) Subsets 1 vs. 3, (**d**) Subsets 2 vs. 4. The grey shaded area represents metabolites that are identified as difficult to quantify in all graphs: (**a**) and (**b**) r_s<0.8; (**c**) and (**d**) difference in percentage of zeros >10. The metabolites rated subjectively in Table 1 retain their grouping here. N/A is Not Analyzed. Error bars represent standard deviation of the values.

Comparing subjective and objective metabolite quantification consistency over time

We found 22 metabolites that were inconsistently quantified in either the over time or inter-operator experiments: Excellent=0, Good=1, Fair=6, Poor=8, and not rated=7 (Supplementary Table S3). All of the metabolites considered inconsistent and difficult to quantify in the consistency of quantification over time experiments were also identified in the inter-operator group consistency study, except for xylose (HMDB00098). This metabolite had r_s of 0.805 and 0.897 in the subjective and objective over time experiments, respectively; there was almost no difference in percentage of zeros, unlike in the inter-operator experiment.

Comparing the subjective and objective approaches to assessing the consistency of metabolite quantification highlighted that they yielded the same results. All of the metabolites rated Poor in the subjective analysis, and many of those rated Fair, were also considered inconsistent in the objective analysis and or the inter-operator group consistency analysis (Supplementary Table S3). Metabolites rated Excellent were universally highly correlated, with a small difference in the percentage of zeros. The similarity in results between our subjective and objective analyses indicates that either is useful in selecting reliable metabolites from complex biofluids for future clinical models or quantification in general.

As the first to study the consistency of metabolite quantification over a span of years, it is important to report the finding of notable differences between the data sets. For example, comparing Subsets 1 vs. 3 resulted in many more metabolites being labeled inconsistent than in Subsets 2 vs. 4. It is possible that the 6 month longer time span for Subsets 1 and 3 included a technological, maintenance, or training

event that was not shared with us by Chenomx for any number of reasons. Subtle, undocumented changes have the potential to alter the robustness and reproducibility of metabolite quantification for many metabolites within a complex biosample.

We also found it useful to differentiate between metabolites with a high percentage of zeros as a result of low concentration from those metabolites in regions of complex NMR spectra. Though both of these metabolite types are variable [2,10,16], this variability may affect models in different ways (e.g., one may be more subjective). However, at this point the only way to differentiate them will be to examine the spectra directly.

Many of the metabolites found to be inconsistent to quantify in one study was corroborated by a second investigatory route. As a result, we saw that none of the metabolites labeled as Excellent in our subjective analysis were considered inconsistent in our other investigations. Conversely, all of the metabolites rated as Poor were found to be inconsistent by one or more analyses. Good metabolites were considered inconsistent, primarily according to the differences in the percentage of zeros. Ethanol was the only Good metabolite that had a low correlation, and this occurred in the consistency over time (and intra-operator studies). This is due to the high percentage of zeros, since ethanol was not detected in most urine samples. Some of the Fair metabolites were considered inconsistent, which was expected as these metabolites were difficult to evaluate subjectively. The not rated metabolites were inconsistently quantified over time thereby increasing the likelihood that they may be biasing the data set and its interpretations. Since many of these metabolites may have been added to the reference database as a placeholder to be fully characterized at a later time, investigators should carefully examine recently added metabolites for consistent quantification over time or between the 2 groups of operators prior to widespread use in modeling.

Inter-operator variability

To assess the consistency in metabolite profiling between different operators, 60 NMR spectra from Subset 4 (Figure 1C) were quantified by 2 groups of operators independently to determine 70 metabolite concentrations for each sample. Similar to the subjective analysis, an X-Y plot was created for each metabolite, with the concentration determined by one operator plotted against the concentration determined by the other operator group. The difference between the two concentrations was analyzed by using multiple potential measures. Several measures did not allow us to easily differentiate more consistent metabolites: p-values, average difference, or absolute average difference (Supplementary Figure S1A, S1C and S1D).

Correlation coefficient: We calculated r_s for each metabolite. If r_s <0.8, the metabolite was classified as inconsistent. Using the groupings determined by subjective rating, we graphed the r_s (Figure 3A). All of the excellent metabolites had r_s values approaching 1. Those rated Good had r_s values generally above 0.9, with only one metabolite (ethanol) falling below 0.8. Even those rated as Fair had an r_s of greater than 0.8, except for xylose. Conversely, close to half of the metabolites rated as Poor had r_s below 0.8. Additionally, of the metabolites not subjectively ranked, more than half had r_s below 0.8. Thus, the use of Spearman's correlation coefficient (r_s) to determine the inter-operator variability allowed us to identify 11 inconsistently quantified metabolites (r_s<0.8) out of 70 (Supplementary Table 1).

Difference in percentage of zeros: The difference in percent zeros between the 2 groups of operators was calculated for each metabolite (Figure 3B). If the difference was larger than 10 percentage points the metabolite was considered difficult to quantify reliably. All Excellent metabolites had no difference in the percentage of zeros. Many of the Good metabolites also had low differences in percentage of zeros; however two were larger than 10 percentage points: creatine (HMDB00064) and tartrate (HMDB00956). Fair metabolites contained a mixture of metabolites of both low and high differences in percentage zeros. Half of the metabolites rated Poor (five) and two thirds of the metabolites not subjectively ranked had differences in percentage of zeros above 10 percentage points (Supplementary Table 1).

In summary, though most of the metabolites identified as difficult to quantify by r_s were also identified by differences in percentage of zeros, the opposite was not always true. Many metabolites were identified as difficult to quantify by differences in percentage of zeros that were not identified by r_s.

Most metabolites considered difficult to quantify by the inter-operator group experiment were also considered difficult to quantify by at least one of the overtime experiments. This is not surprising, as there is a potential for the personnel composition of an operator group to change in the overtime comparison, and this was not controlled for in these experiments. However, this does demonstrate that for future experiments it may be possible to examine variability over time alone, and not need to examine inter-operator variability. Additionally, similar to a previous study and for most of the metabolites tested, the 2 groups of operators were quantifying the metabolites consistently, especially those ranked Excellent and Good. The correlation values were all greater than 0.5, and the correlations were generally higher the better the subjective metabolite label. However, a previous study also found that low concentration metabolites correlated with high between-person variability [16]. This may explain some of the differences in the percentage of zeros seen with the not rated metabolites, or those ranked fair or poor.

The final ranking of the 70 metabolites based upon both subjective and objective experimental findings are presented in Table 3. No similarities were found between metabolites in the groups. For example, the super class of the metabolites classified as excellent included organonitrogen compounds, alkaloids and derivatives, organic acids and derivatives, benzenoids, and organoheterocyclic compounds. Similar variety at the super class strata was found for the group of metabolites ranked as good, fair, or poor.

Conclusion

The goal of this paper was to develop methods to identify and characterize potential sources of variability affecting metabolomic data interpretation and quantification. As statistical methods evolve, groups will seek to enhance sample size by combining existing data or conduct longer term experiments. However, results derived from these practices could lead to unforeseen biases in data interpretation or quantification. We have outlined a procedure that could be followed by other labs to examine their own data in a similar manner and determine which metabolites are inconsistently analyzed and thus difficult to quantify.

The protocol outlined here will contribute to the selection of metabolites used to create a robust dataset that will be more consistent. These measures can also aid statistical analysis by prioritizing reliable metabolites and sorting redundant features before statistical analysis, or incorporated into Bayesian networks, where expert knowledge (i.e., confidence of measures) is also modeled. Importantly, our procedure can be readily incorporated into routine metabolomic data analyses, such as those described by Sun and colleagues [20].

Final Metabolite Classification			
Metabolite	**HMDB ID**	**Metabolite**	**HMDB ID**
Excellent			
Alanine	HMDB00161	Hippurate	HMDB00714
Creatinine	HMDB00562	Succinate	HMDB00254
Dimethylamine	HMDB00142	Trigonelline	HMDB00875
Formate	HMDB00142	Trimethylamine N-oxide	HMDB00925
Glycine	HMDB00123		
Good			
1,6-Anhydro-D-glucose	HMDB00640	Glucose	HMDB00122
2-Hydroxyisobutyrate	HMDB00729	Lactate	HMDB00190
3-Hydroxyisovalerate	HMDB00754	Salicylurate	HMDB00840
4-Hydroxyphenylacetate	HMDB00020	Sucrose	HMDB00258
Acetate	HMDB00042	Tartrate	HMDB00956
Ascorbate	HMDB00044	Taurine	HMDB00251
Citrate	HMDB00094	Tyrosine	HMDB00158
Creatine	HMDB00064	Urea	HMDB00294
Ethanol	HMDB00108	Valine	HMDB00883
Fair			
1-Methylnicotinamide	HMDB00699	Histidine	HMDB00177
3-Methylxanthine	HMDB01886	Hypoxanthine	HMDB00157
3-Aminoisobutyrate	HMDB03911	Leucine	HMDB00687
3-Hydroxymandelate	HMDB00750	Mannitol	HMDB00765
3-Indoxylsulfate	HMDB00682	N,N-Dimethylglycine	HMDB00092
Acetone	HMDB01659	O-Acetylcarnitine	HMDB00201
Asparagine	HMDB00168	Pantothenate	HMDB00210
Benzoate	HMDB01870	Propylene glycol	HMDB01881
Betaine	HMDB00043	Pyroglutamate	HMDB00267
Carnitine	HMDB00062	Threonine	HMDB00167
Ethanolamine	HMDB00149	Xylose	HMDB00098
Fumarate	HMDB00134	trans-Aconitate	HMDB00958
Glutamine	HMDB00641	p-Methylhistidine	HMDB00479
Glycolate	HMDB00115	t-Methylhistidine	HMDB00001
Poor			
2-Oxoglutarate	HMDB00208	Methanol	HMDB01875
3-Hydroxyphenylacetate	HMDB00440	myo-Inositol	HMDB00211
Acetoacetate	HMDB00060	Pyruvate	HMDB00243
cis-Aconitate	HMDB00072	Serine	HMDB00187
Ethylmalonate	HMDB00622	Tryptophan	HMDB00929
Isoleucine	HMDB00172	Uracil	HMDB00300
Lysine	HMDB00182		

Table 3: List of 70 metabolites and their final classification based upon the results of all of the experiments.

Our key finding is that metabolites that are difficult to profile are inconsistent and could potentially lead to a misleading result when combined with analysis of data over time (i.e., batch effect). This effect would be magnified for more inconsistent compounds. When using the Chenomx software in conjunction with the information provided here, meaningful metabolite profiles should ideally include excellent and good metabolites while minimizing fair metabolites and avoiding those ranked as poor. This approach to analyzing the robustness and reproducibility of metabolite profiles indicative of disease or a condition provides another layer of validity and potential clinical utility to the unique metabolite profile.

Acknowledgements

The authors thank Metabolomic Technologies Inc. (Edmonton, AB) and the National High Field Nuclear Magnetic Resonance Centre (NANUC) at the University of Alberta (Edmonton, AB) for technical assistance.

References

1. Lacy P, McKay RT, Finkel M, Karnovsky A, Woehler S, et al. (2014) Signal intensities derived from different NMR probes and parameters contribute to variations in quantification of metabolites. PLoS One 9: e85732.
2. Viant MR, Bearden DW, Bundy JG, Burton IW, Collette TW, et al. (2009) International NMR-based environmental metabolomics inter-comparison exercise. Environ Sci Technol 43: 219-225. w
3. Keun HC, Ebbels TMD, Antti H, Bollard ME, Beckonert O, et al. (2002) Analytical reproducibility in ^{1}H NMR-based metabonomic urinalysis. Chem Res Toxicol 15: 1380-1386.
4. Dumas ME, Maibaum EC, Teague C, Ueshima H, Zhou B, et al. (2006) Assessment of analytical reproducibility of ^{1}H NMR spectroscopy based metabonomics for large-scale epidemiological research: The INTERMAP study. Anal Chem 78: 2199-2208.
5. Saude EJ, Sykes BD (2007) Urine stability for metabolomic studies: Effects of preparation and storage. Metabolomics 3: 19-27.
6. Saude EJ, Adamko D, Rowe BH, Marrie T, Sykes BD (2007) Variation of metabolites in normal human urine. Metabolomics 3: 439-451.
7. Townsend MK, Clish CB, Kraft P, Wu C, Souza AL, et al. (2013) Reproducibility of metabolomic profiles among men and women in 2 large cohort studies. Clin Chem 59: 1657-1667.
8. Lauridsen M, Hansen SH, Jaroszewski JW, Cornett C (2007) Human urine as test material in ^{1}H NMR-based metabonomics: Recommendations for sample preparation and storage. Anal Chem 79: 1181-1186.
9. Maher AD, Zirah SFM, Holmes E, Nicholson JK (2007) Experimental and analytical variation in human urine in ^{1}H NMR spectroscopy-based metabolic phenotyping studies. Anal Chem 79: 5204-5211.
10. Wang B, Goodpaster AM, Kennedy MA (2013) Coefficient of variation, signal-to-noise ratio, and effects of normalization in validation of biomarkers from NMR-based metabonomics studies. Chemom Intell Lab Syst 128: 9-16.
11. Kohl SM, Klein MS, Hochrein J, Oefner PJ, Spang R, et al. (2012) State-of-the art data normalization methods improve NMR-based metabolomic analysis. Metabolomics 8: 146-160.
12. Fages A, Pontoizeau C, Jobard E, Levy P, Bartosch B, et al. (2013) Batch profiling calibration for robust NMR metabonomic data analysis. Anal Bioanal Chem 405: 8819-8827.
13. Wang K, Sun J, Zhou S, Wan C, Qin S, et al. (2013) Prediction of drug-target interactions for drug repositioning only based on genomic expression similarity. PLoS Comput Biol 9: e1003315.
14. Slupsky CM, Rankin KN, Wagner J, Fu H, Chang D, et al. (2007) Investigations of the effects of gender, diurnal variation, and age in human urinary metabolomic profiles. Anal Chem 79: 6995-7004.
15. Sokolenko S, Blondeel EJM, Azlah N, George B, Schulze S, et al. (2014) Profiling convoluted single-dimension proton NMR spectra: A plackett-burman approach for assessing quantification error of metabolites in complex mixtures with application to cell culture. Anal Chem. American Chemical Society 86: 3330-3337.
16. Tredwell GD, Behrends V, Geier FM, Liebeke M, Bundy JG (2011) Between-person comparison of metabolite fitting for NMR-based quantitative metabolomics. Anal Chem 83: 8683-8687.
17. Wang H, Tso V, Wong C, Sadowski D, Fedorak RN (2014) Development and validation of a highly sensitive urine-based test to identify patients with colonic adenomatous polyps. Clin Transl Gastroenterol 5: e54.
18. Eisner R, Greiner R, Tso V, Wang H, Fedorak RN (2013) A machine-learned predictor of colonic polyps based on urinary metabolomics. Biomed Res Int 2013: 303982.
19. Wong CKW, Fedorak RN, Prosser CI, Stewart ME, Van Zanten SV, et al. (2012) The sensitivity and specificity of guaiac and immunochemical fecal occult blood tests for the detection of advanced colonic adenomas and cancer. Int J Colorectal Dis 27: 1657-1664.
20. Sun L, Li J, Zhou K, Zhang M, Yang J, et al. (2013) Metabolomic analysis reveals metabolic disturbance in the cortex and hippocampus of sub-chronic MK-801 treated rats. PLoS One 8: e60598.