Research Article   **Open Access**

Optimizing Urine Processing Protocols for Protein and Metabolite Detection

Nazema Y Siddiqui[1]*, Laura G DuBois[2], Lisa St John-Williams[2], Will Thompson J[2], Carole Grenier[3], Emily Burke[4], Matthew O Fraser[5], Cindy L Amundsen[1] and Susan K Murphy[3]

[1]Division of Urogynecology & Reconstructive Pelvic Surgery; Department of Obstetrics & Gynecology, Duke University, Durham, NC, USA
[2]Center for Genomic & Computational Biology; Duke University, Durham, NC, USA
[3]Division of Gynecologic Oncology; Department of Obstetrics & Gynecology, Duke University, Durham, NC, USA
[4]Department of Biostatistics and Bioinformatics, Duke University, Durham, NC, USA
[5]Division of Urology; Department of Surgery, Duke University, Durham, NC, USA

Abstract

Background: In urine, factors such as timing of voids, and duration at room temperature (RT) may affect the quality of recovered protein and metabolite data. Additives may aid with detection, but can add more complexity in sample collection or analysis. We aimed to identify the optimal urine processing protocol for clinically-obtained urine samples that allows for the highest protein and metabolite yields with minimal degradation.

Methods: Healthy women provided multiple urine samples during the same day. Women collected their first morning (1st AM) void and another "random void". Random voids were aliquotted with: 1) no additive; 2) boric acid (BA); 3) protease inhibitor (PI); or 4) both BA + PI. Of these aliquots, some were immediately stored at 4°C, and some were left at RT for 4 hours. Proteins and individual metabolites were quantified, normalized to creatinine concentrations, and compared across processing conditions. Sample pools corresponding to each processing condition were analyzed using mass spectrometry to assess protein degradation.

Results: Ten Caucasian women between 35-65 years of age provided paired 1st morning and random voided urine samples. Normalized protein concentrations were slightly higher in 1st AM compared to random "spot" voids. The addition of BA did not significantly change proteins, while PI significantly improved normalized protein concentrations, regardless of whether samples were immediately cooled or left at RT for 4 hours. In pooled samples, there were minimal differences in protein degradation under the various conditions we tested. In metabolite analyses, there were significant differences in individual amino acids based on the timing of the void.

Conclusions: For comparative translational research using urine, information about void timing should be collected and standardized. For urine samples processed in the same day, BA does not appear to be necessary while the addition of PI enhances protein yields, regardless of 4°C or RT storage temperature.

Keywords: Urine; Proteomics; Metabolomics; Boric acid; Protease inhibitors; First morning void; Random void; Urine protein degradation

Introduction

There is significant interest in studying urine proteins and metabolites as potential biomarkers for clinical diseases. Urine serves as an easily accessible biologic fluid that can be accessed using non-invasive methods. Urine is proximate to the bladder wall, and also contains renally-cleared systemic compounds and metabolites. Thus urinary biomarkers may be helpful in distinguishing pathologic versus normal biologic processes for renal, genitourinary, and other medical conditions.

In clinically obtained urine samples, multiple factors may introduce variability and affect the predictive value of urine protein and metabolite data. In general, normal (non-proteinuric) urine has low quantities of protein. Some would argue that 1st morning voids, containing the highest protein concentrations, are helpful for proteomic studies. However, logistically there is an obligate time delay when study participants collect their 1st morning void, and factors such as time at room temperature, ongoing protease activity, or bacterial contamination from urethral microbes may affect data quality. Thus prior studies have suggested collecting the 2nd morning or other random "spot" urine [1]. However, it remains unclear if the addition of protease inhibitors or bacteriostatic agents may preserve proteins and metabolites in 1st morning samples and facilitate their use. This is relevant since urinary proteomic studies require maximal concentrations of protein from urine with minimal loss [2].

Although it may be appealing to simply collect random or "spot" urine samples in a clinical setting, clinic-based staff may need additional training with aliquoting and sample processing. These staff may potentially introduce more variability than when specimens are processed with laboratory personnel, and it is hard to know if the resources expended to process samples in clinic are justified. Furthermore, urine samples produced in a clinical environment may stay at room temperature for hours prior to final processing in a laboratory. Even if urine is immediately collected and processed, the presence of antibacterial agents or protease inhibitors may affect the quality and reproducibility of protein and metabolite yield [1].

To optimize data quality, investigators have proposed adding bacteriostatic agents such as boric acid or sodium azide to urine [3], though it is still hard to know if this should be prioritized. Furthermore,

the toxicity of sodium azide makes its use difficult in the clinic. Protease inhibitors (PI) are thought to be less important in urine [1], which has generally low levels of endogenous proteases, but it is unclear if a PI may actually be helpful when urine remains at room temperature prior to transport from the clinic to another laboratory. Furthermore, the use of additives such as PI and bacteriostatic compounds has not been tested in combination on clinically-obtained urine samples. There is a need to confirm that findings obtained in laboratory settings remain relevant in the clinical environment where more variability exists. There is also a need to standardize collection and storage protocols for translational research, which is particularly relevant in light of recent National Institutes of Health concerns highlighting how different environments or protocols could affect reproducibility of preclinical data [4]. Therefore, our objective was to compare and quantify protein and metabolite yields from urine obtained in a clinical setting and subjected to different processing conditions.

Materials and Methods

Urine collection

After Duke University Institutional Review Board approval was obtained, we recruited healthy Caucasian women, ages 35-65 to provide multiple same-day urine samples. Participants first provided informed consent, and were then screened for inclusion. Inclusion criteria were normal urinalyses, a negative pregnancy test (for women of childbearing age), and the absence of urinary symptoms such as frequency or nocturia, based on the validated Urinary Distress Inventory short form [5].

Participants were provided a home urine collection kit with detailed instructions to collect their first morning (1st AM) void, separate the urine into multiple aliquots, and store at 4°C. Women were instructed to bring refrigerated samples to the clinic within 4 hours. In the clinic, women provided another "random void" or "spot" sample, which was also divided into multiple aliquots. Random void samples were either immediately stored at 4°C, or left at room temperature (RT) for 4 hrs prior to cooling to 4°C. For both 1st AM and random samples, aliquots contained 10 ml of urine and either: 1) no additive; 2) boric acid (BA) [10 millimoles (mM) powdered BA]; 3) a protease inhibitor (PI) tablet [cOmplete™ ULTRA Tablets, Mini, *EASYpack*; Sigma-Aldrich Corp., St. Louis, Missouri]; or 4) both 10 mM BA+PI. For aliquots with boric acid, we used 10 mM of BA based on prior recommendations calling for 2-20 mM [1].

In the home urine collection kit, vials with additives were color-coded; participants received separate color-coded disposable pipettes and detailed instructions regarding pipetting urine in such a way as to avoid cross-contamination. In the clinic, a research coordinator who was trained in basic laboratory procedures performed all of the pipetting. After all samples were cooled to 4°C, they were transferred to the laboratory where they were spun at 1800g for 5 min at 4°C, and the supernatant was stored at -80°C.

Protein analyses

For analyses of proteins, urine samples were thawed and protein concentrations were determined using 50 µl of sample and a mini-Bradford Assay (Biorad, Inc). To assess for potential degradation in samples left at room temperature or without protease inhibitors or bacteriostatic agents, sample pools were created for the processing conditions described above. For each pool, an equal amount of protein from each participant per processing condition was added to create the pooled sample. Pooled samples were concentrated and buffer exchanged

into 50 mM ammonium bicarbonate using 10 kDa MWCO Amicon 4 spin filters (Millipore). Concentrations were determined after buffer exchange and samples were normalized to 9 micrograms protein in 50 mM Ammonium Bicarbonate with 0.1% w/v Rapigest SF (Waters), reduced with 10 mM DTT at 80°C for 15 minutes, alkylated with 25 mM iodoactamide in the dark at room temperature for 30 minutes, then digested overnight with 200 ng trypsin (Promega) at 37°C.

Approximately 200 ng of digested samples were analyzed using a nanoAcquity UPLC (Waters Corporation, Milford, MA) coupled to a Q-Exactive Plus tandem mass spectrometer (Thermo) via a nanoelectrospray ionization source. Peptides were separated using a 25 cm × 75 um HSS T3 C18 column, with a gradient from 5 to 40% MeCN (0.1% formic acid) over 90 minutes, at a flow rate of 0.4 µL/min and column temperature of 55°C. Mass spectrometry analysis was performed with MS1 resolution 70,000 (@ 200m/z) and data-dependent MS/MS sequencing at MS2 resolution 17,500 for the top 10 most abundant precursor ions. Normalized collision energy of 30V and dynamic exclusion of 30 seconds was employed. Data were converted within Proteome Discoverer (Thermo) to .mgf searchable files, and submitted to Mascot v2.4 (Matrix Sciences, Inc). The data were searched against the Swiss Prot database with human taxonomy selected, 5 ppm precursor and 0.02 Da product ion mass accuracy, allowing for variable modifications of deamidation of asparagine [N] and glutamine [Q] and oxidation of methionine [M]. For identifications from each pool, we assessed the percent of all peptides with a semitryptic end (only 1 end with amino acids arginine [R] or lysine [K] versus fully tryptic with both termini being either R or K). Percentages of peptides with semi-tryptic cleavage sites were used as a proxy for protein degradation. Database search results and comparison of semitryptic cleavage between samples was performed in Scaffold v4.3 software (Proteome Software, Seattle WA). Peptide identifications were accepted if they could be established at greater than 99.0% probability to achieve a false discovery rate (FDR) less than 1.0% by the Peptide Prophet algorithm [6] with Scaffold delta-mass correction. Protein identifications were accepted if they could be established at greater than 97.0% probability to achieve an FDR less than 1.0% and contained at least 1 identified peptide. The Scaffold file has been made available for download using the following link: (https://discovery.genome.duke.edu/express/resources/4021/4021_032715_Condensed.sf3).

Metabolite analyses

For analyses of metabolites, thawed supernatant samples were prepared using the Absolute*IDQ*® p180 kit (Biocrates Innsbruck, Austria) according to the manufacturer instructions. After the addition of 10 µL internal standard to the 96 or 56-well extraction plate, 15 µL of each urine supernatant sample were added to the appropriate wells. The plate was dried under a gentle stream of nitrogen. An additional 15 µL of each urine sample was added to the appropriate wells and dried under a gentle stream of nitrogen. The samples were derivatized with phenyl isothiocyanate then eluted with 5mM ammonium acetate in methanol. Samples were diluted with either 40% methanol in water for the UPLC analysis (15:1) or running solvent (a proprietary mixture provided by Biocrates) for flow injection analysis (20: 1). A pool of equal volumes of all samples (Metabolite Study Pool) was created, prepared, and analyzed using the same techniques as individual study samples to assess potential batch effects.

LC separation of amino acids and biogenic amines was performed using a Waters (Milford, MA) Acquity UPLC using a Waters Acquity 2.1 mm x 50 mm 1.7 µm BEH C18 column fitted with a Waters Acquity BEH

C18 1.7 µm Vanguard guard column. Acylcarnitines, sphingolipids, and glycerophospholipids, were analyzed by flow injection analysis (FIA). Using electrospray ionization, samples for both UPLC and FIA were introduced directly into a Xevo TQ-S triple quadrupole mass spectrometer (Waters) operating in the Multiple Reaction Monitoring (MRM) mode. MRM transitions (compound-specific precursor to product ion transitions) for each analyte and internal standard were collected over the appropriate retention time. The UPLC-MS/MS data were imported into Waters application TargetLynx™ for peak integration, calibration, and concentration calculations. The UPLC-MS/MS data from TargetLynx™ and FIA-MS/MS data were analyzed using Biocrates MetIDQ™ software.

Normalization and statistical analyses

Creatinine concentrations were calculated in each sample using a colorimetric Urinary Creatinine Assay Kit (Cell BioLabs, Inc, San Diego, CA). Creatinine concentrations were used to create normalized protein and metabolite concentrations. Normalized sample concentrations from various conditions were compared as outlined in the study schema in Table 1. We compared normalized protein concentrations from the Bradford Assay based on the timing of the void (1st AM vs. random), presence of additives, and time at RT, using non-parametric Kruskal Wallis tests. Normalized metabolite concentrations were similarly compared across conditions. To assess for protein degradation per processing condition, proportions of semi-tryptic cleavages were directly compared. All statistical analyses were performed using SPSS Version 20.0 (Chicago, Illinois) with $p < 0.05$ considered statistically significant.

Results

Ten women provided same-day urine samples for our study. Since we purposely recruited a homogeneous population to minimize biologic variability, all were healthy non-pregnant white women between 35-65 years of age. For each participant, we obtained 1st morning and random voided samples, with median urine creatinine concentrations of 92.4 mg/dL (IQR 70.3, 125.1). We processed samples with additives (BA, PI, both BA + PI), and tested samples that were immediately cooled to 4°C versus those kept at room temperature (RT) for 4 hours.

Concentrations of urinary protein based on time of void and time at RT

Median protein concentrations from the 1st morning void and random voids were normalized to urine creatinine and are listed in Table 2 and displayed in Figure 1. Regarding timing of voids, we found that 1st AM voids ± BA had higher normalized protein concentrations compared to random voids ± BA [median 0.025 µg/µl (IQR 0.02, 0.04) vs. 0.017 (IQR 0.01, 0.03), p=0.03]. In random voids without additives, when samples were left at RT for 4 hours there were no differences in normalized protein compared to those immediately cooled to 4°C [Immediately cooled median 0.02 µg/µl (IQR 0.01, 0.03) vs. RT × 4hr median 0.02 (IQR 0.01, 0.05), p=0.72].

	1st AM void		Random clinic void			
	No add	+BA	No add	+BA	+PI	+BA+PI
Immediately Cooled (4° C)	0.03 [0.02, 0.04]	0.02 [0.02, 0.03]	0.02 [0.01, 0.03]	0.01 [0.01, 0.02]	0.05 [0.03, 0.10]	0.03 [0.03, 0.06]
RT x 4 hr			0.02 [0.01, 0.05]	0.02 [0.01, 0.02]	0.04 [0.03, 0.07]	0.04 [0.03, 0.06]

Table 2: Normalized protein concentrations (µg/µL by Bradford Assay)

1st AM void: First morning void; No add: no additive; RT: room temperature; BA: boric acid; PI: protease inhibitor. Protein concentrations normalized to urine creatinine.Data displayed as median [interquartile range]; All comparisons tested using Kruskal-Wallis.Comparison of time of void - 1st AM ±BA vs. random ±BA: **p=0.03**Comparison of boric acid - 1st AM no add vs. 1st AM +BA: p=0.65Comparison of time at RT - Random no add 4° C vs. random no add RT x 4hr: p=0.72Comparison of PI - Random voids no add ±BA vs. random PI ±BA: **p<0.001**Combination - random voids compare: 1) no add; 2) +BA; 3) +PI; 4) +BA+PI: **p <0.001** (Multiple pairwise testing reveals groups with PI to be significantly different than non-PI)
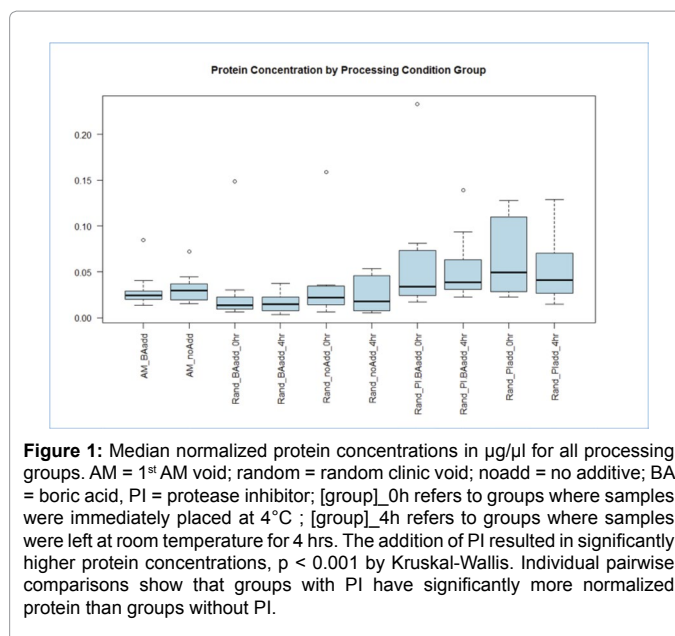

Figure 1: Median normalized protein concentrations in µg/µl for all processing groups. AM = 1st AM void; random = random clinic void; noadd = no additive; BA = boric acid, PI = protease inhibitor; [group]_0h refers to groups where samples were immediately placed at 4°C ; [group]_4h refers to groups where samples were left at room temperature for 4 hrs. The addition of PI resulted in significantly higher protein concentrations, p < 0.001 by Kruskal-Wallis. Individual pairwise comparisons show that groups with PI have significantly more normalized protein than groups without PI.

Concentrations of urinary protein based on additives - BA and PI

All samples were normalized to urine creatinine and additives were tested alone and in combination as displayed in Figure 1. Regarding the use of a bacteriostatic agent, we tested boric acid (BA) since the alternative, sodium azide, has a toxicity profile that generally precludes its use in a clinical setting. When comparing 1st AM or random void samples ± BA, the presence of BA did not result in a significant difference in normalized protein concentrations (p=0.65).

Since BA theoretically reduces bacterial action while BA and PI reduce protein degradation, we chose to make a qualitative assessment of protein degradation using mass spectrometry analysis. We utilized pooled samples from ten processing conditions of void samples (Table 1) to see if timing of void, lower temperatures, addition of BA or protease inhibitors decreased the observation of nonspecific cleavage events. We digested pooled samples with trypsin, analyzed peptides using mass spectrometry, and compared the proportion of individual peptides per sample pool with at least one non-tryptic cleavage site (semi-tryptic

	1st AM void		Random clinic void			
Immediately Cooled (4° C)	(1) no additive	(2) +BA	(3) no additive	(4) + BA	(5) + PI	(6) + BA + PI
RT x 4 hr			(7) no additive	(8) + BA	(9) + PI	(10) + BA + PI

Table 1: Study Schema (comparison groups; n=10 samples) 1st AM void: First morning void; RT: room temperature; BA: boric acid; PI: protease inhibitor
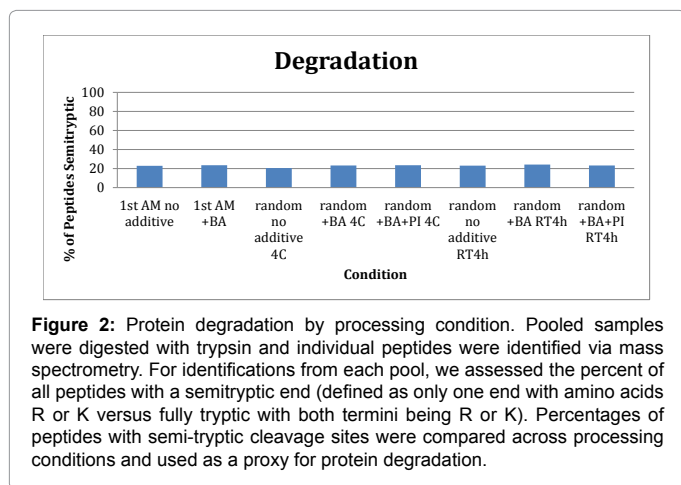
Figure 2: Protein degradation by processing condition. Pooled samples were digested with trypsin and individual peptides were identified via mass spectrometry. For identifications from each pool, we assessed the percent of all peptides with a semitryptic end (defined as only one end with amino acids R or K versus fully tryptic with both termini being R or K). Percentages of peptides with semi-tryptic cleavage sites were compared across processing conditions and used as a proxy for protein degradation.

cleavages, Figure 2). Since we purposely introduced trypsin into samples, we inferred that any non-tryptic cleavage sites occurred due to other unintended cleavage activity. Thus the presence of non-tryptic cleavage sites was used as a proxy for protein degradation. Mass spectrometry revealed only modest differences in semi-tryptic cleavages between the conditions. The values ranged from 20.5% semi-tryptic cleavages in random voids that were immediately cooled, to 24.3% in random voids + BA at RT for 4 hr. Under the various conditions we tested, native proteolytic cleavage measured by mass spectrometry does not appear to be significantly altered.

Regarding protease inhibitors, in all combinations, whether samples were immediately cooled, left at RT for 4 hours, or also had BA, the addition of PI significantly improved median protein concentrations ($p < 0.001$, see Figure 1). We further explored this relationship by comparing random voids ± BA to random voids with PI ± BA, and confirmed higher normalized protein concentrations in samples with PI [median 0.02 µg/µl (0.01, 0.03) no PI vs. 0.04 (0.03, 0.08) with PI, $p < 0.001$, Figure 3]. However, when identifying proteins using mass spectrometry the presence of PI may add a layer of complexity. As depicted in Table 3, the number of identified proteins was highest in samples that were immediately cooled but had BA, while for the same samples; the number of identified proteins was lower when PI was added. Thus in our study the addition of PI resulted in higher overall protein concentrations but less ability to identify proteins with mass spectrometry.

Concentrations of urinary metabolites

Among urinary metabolites, we first compared amino acids (AA) under different processing conditions (Table 4). The majority of AA were identified in our samples, with the exception of aspartic acid (Asp), citrulline (Cit), and glutamic acid (Glu), which were lower than the limit of detection in almost every sample. Using the same testing paradigm that we applied to proteins, we compared normalized AA concentrations across various processing conditions. The only processing condition that revealed significant differences in recovery of metabolites was time of void (1st AM vs. random), where 1st AM voids resulted in different normalized concentrations for the following AA: Arginine, Asparagine, Histidine, Isoleucine, Serine, and Valine ($p < 0.05$ for all).

Regarding biogenic amines, 9 amines had values in over half of the samples and were normalized and analyzed further. These included: ADMA, alpha.AA, carnosine, dopamine, histamine, kynurenine,
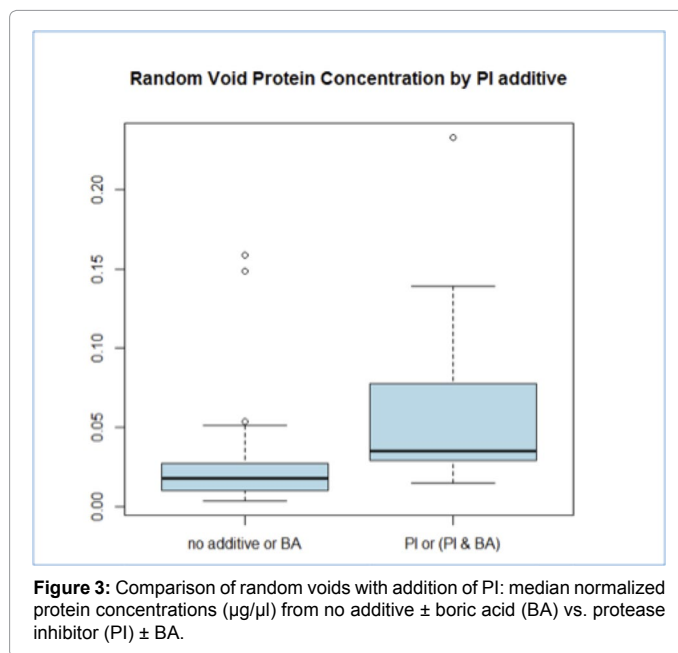

Figure 3: Comparison of random voids with addition of PI: median normalized protein concentrations (µg/µl) from no additive ± boric acid (BA) vs. protease inhibitor (PI) ± BA.

Sample Pool	Number of Identified Spectra	Number of identified proteins	Total Number of Spectra	%ID (identification rate)
1st AM 4C no add	6794	544	45770	15%
1st AM 4C +BA	9528	589	46591	20%
Random 4C no add	7076	499	46210	15%
Random 4C +BA	8797	516	46533	19%
Random 4C +BA +PI	6074	439	47431	13%
Random RT no add	8581	559	49220	17%
Random RT + BA	6832	457	47670	14%
Random RT + BA +PI	5463	472	48280	11%

Table 3: Proteins and Mass Spectra Identified per Processing Condition

serotonin, t4.OR.Pro, and taruine. Similar to AA metabolites, we generally did not identify significant differences in metabolite concentrations based on processing condition with the exception of AMDA and serotonin, where significantly higher values were found in 1st AM compared to random voids ($p < 0.01$ for both). Most acylcarnitines, sphingolipids, and glycerophospholipids were less than the lower limit of detection such that we were not able to analyze these by condition.

Discussion

We tested 10 urine storage and processing conditions to ascertain the importance of the timing of the void (1st AM vs. random void), time at room temperature (immediately cooled vs. 4 hr at RT), and presence of additives (BA, PI, or combination of both BA and PI). We identified higher normalized protein concentrations in 1st AM voids compared to random voids. Normalized proteins concentrations were highest in samples where PI was added, regardless of the time of void or duration at room temperature, though the ability to identify proteins using mass spectrometry was lower with the presence of PI. Protein concentrations did not differ between samples that were immediately cooled and those left at room temperature for 4 hours. The presence of boric acid, an antimicrobial agent, did not significantly alter protein concentrations

THREONINE					
1st AM void		Random clinic void			
No add	+BA	No add	+BA	+PI	+BA+PI
Immediately Cooled (4° C) 37.4 [25.4, 156.0]	40.7 [22.2, 152.0]	53.6 [14.7, 130.8]	32.7 [11.2, 116.3]	51.2 [11.5, 80.9]	26.1 [7.3, 81.7]
RT x 4 hr		21.6 [10.5, 92.9]	51.4 [14.5, 80.0]	33.9 [13.1, 88.5]	26.6 [11.0, 55.5]

TYROSINE					
1st AM void		Random clinic void			
No add	+BA	No add	+BA	+PI	+BA+PI
Immediately Cooled (4° C) 29.5 [14.8, 59.8]	36.8 [11.8, 66.1]	33.6 [10.2, 50.2]	23.2 [8.2, 61.1]	33.1 [12.2, 54.2]	15.6 [5.5, 32.1]
RT x 4 hr		53.5 [5.4, 52.6]	57.7 [10.7, 69.9]	45.3 [5.7, 56.7]	44.1 [4.5, 40.7]

PROLINE					
1st AM void		Random clinic void			
No add	+BA	No add	+BA	+PI	+BA+PI
Immediately Cooled (4° C) 2.9 [1.9, 6.1]	4.6 [1.7, 5.8]	(n=8) 4.0 [1.9, 5.9]	(n=8) 3.3 [1.3, 4.9]	(n=8) 4.4 [1.8, 5.9]	(n=6) 3.1 [1.7, 5.9]
RT x 4 hr		(n=6) 4.5 [2.0, 6.4]	(n=7) 4.9 [2.4, 8.2]	(n=7) 4.2 [2.3, 5.5]	(n=6) 3.2 [2.2, 6.4]

GLUTAMATE					
1st AM void		Random clinic void			
No add	+BA	No add	+BA	+PI	+BA+PI
Immediately Cooled (4° C) (n=3) 5.6 [4.8-8.6]	(n=2) 13.9 [3.8-24.1]	(n=4) 7.3 [2.9-23.8]	(n=4) 7.1 [2.9-25.1]	(n=4) 7.7 [2.3-21.6]	(n=3) 7.2 [2.8-24.3]
RT x 4 hr		(n=3) 2.8 [2.4-17]	(n=3) 8.2 [2.1-12.3]	(n=3) 6.0 [3.1-15.3]	(n=2) 4.0 [2.3-5.8]

Table 4: Amino Acid concentrations (μM by AbsoluteIDQ® p180 kit) n=10 unless otherwise noted; 1st AM void: First morning void; No add: no additive; RT: room temperature; BA: boric acid; PI: protease inhibitor Data displayed as median [interquartile range] or [range] if n<5No significant differences identified in Thr (p=0.14), Tyr (p=0.05), Pro (p= 0.32), or Glu (p=0.62) per processing condition

in our study, and the amount of degradation (as estimated by the proportion of peptides with semi-tryptic cleavage sites) did not vary with the presence of BA. With regards to urinary amino acids and biogenic amine metabolites, certain metabolites were found in significantly higher concentrations in the 1st AM void. Otherwise metabolites did not significantly vary based on the processing conditions we tested.

Proteomic and metabolomic studies with human urine samples are continually hampered by the lack of standard guidelines for urine collection and processing, and many investigators have identified the need for standardization [7-9]. The standardization of biobanking protocols would allow for comparisons of results between studies and would also allow for collaborative approaches between biobanks. Multiple strategies for urine collection and processing currently exist

[9]; these protocols variably advocate for different timing of urine samples (spot or 2nd AM), while some investigators advocate for collection of first morning voids or 24-hour urine samples to maximize protein yields [10]. Furthermore, within these collection strategies, different groups advocate for the presence or absence of additives such as protease inhibitors or bacteriostatic agents [9] despite the fact that it is not clear if these additives are helpful in non-proteinuric samples [11].

The Human Kidney & Urine Proteome Project (HKUPP), which is associated with the Human Proteome Organization (HUPO) has proposed a standardization protocol after numerous consensus meetings [12]. One strength of our study is that we tested many of the conditions addressed by this protocol to provide empirical data to guide scientists and systems biologists regarding which processing conditions most affect protein and metabolite data. Our data corroborate the major suggestions provided by the HKUPP guidelines, showing that random or 2nd morning voided samples are reasonable to use for biobanking. Our study shows that urine protein concentrations are not significantly compromised when using random voided samples, but it still may be important to collect urine samples from patients at similar times of the day to minimize variability based on diurnal patterns. Based on our data, this is particularly relevant for metabolite studies.

Our data further corroborate existing evidence suggesting that the urinary proteome does not change significantly when stored for up to 6 hours at room temperature, or up to 3 days at 4°C [13-15]. We showed minimal differences in proteins and metabolites from urine samples that remain at room temperature for up to 4 hours compared to those that are immediately cooled.

Regarding the use of additives, the KHUPP guidelines call for the addition of a bacteriostatic agent such as boric acid or sodium azide. Due to potential toxicity, sodium azide is not useful in a clinical setting. Boric acid is certainly more feasible though in our study the presence of BA did not significantly alter normalized protein concentrations. However it is notable that in our proteomic analyses, we were able to identify more proteins in samples that had boric acid, suggesting that some breakdown or ongoing proteolytic activity was occurring in samples without boric acid. Our samples were all processed and ultimately frozen within 6 hours (4 hours at 4°C or RT in the clinic followed by 1-2 hours for transport, centrifugation, aliquoting, and freezing at -80°C in the lab). Bacteriostatic agents may be more helpful when there are longer transit times. Our metabolomic results did not vary based on the presence or absence of a bacteriostatic agent suggesting that the addition of boric acid does not impede useful metabolomics data; others have found that filtration might be preferable to bacteriostatic additives in optimizing metabolomic results [16]. Thus, for translational studies where urine is processed and frozen within the same day, it may be helpful to prioritize other components of collection rather than the addition of bacteriostatic agents.

Finally, regarding protease inhibitors, our findings showed slightly higher normalized protein concentrations when PI were present, though these differences were minimal. These findings are similar to others who have found minimal variability in urine proteins when PI is used [17]. Zhou et al. [18] assessed sample processing conditions, emphasizing the presence of protease inhibitors for recovery of exosomes in urine. They found that protease inhibitors were necessary for preservation of exosome-associated proteins, and thus there may be special situations where PI should be prioritized. Regarding PI, Havanapan and Thongboonkerd found that PI remain helpful in renal proteomics, but they did not recommend the use of PI for urinary

proteomics [11]. Our data were equivocal, showing that PI may help to maximize protein yield, but protein identifications may be limited in mass spectrometry based proteomics.

In addition to adding to existing literature supporting processing standards, our study is strengthened by its applicability to true clinical practice. We collected urine samples in a clinical setting where subjects themselves or clinical research coordinators had to perform initial processing, and samples obtained in the clinic were then transported to a research laboratory. Although this process is not as rigorous as immediately processing samples post-voiding, our study reflects the actual experiences of many translational researchers and increases the generalizability of our results. Strength of our design is that of comparing samples from the same subjects. This allowed us to minimize biologic variability and focus on the variability introduced by different processing conditions to further elucidate which processes should be prioritized for high quality results.

Despite these strengths, our study also has several weaknesses. We assessed for degradation by mass spectrometry using sample pools rather than individual samples, because of the relatively low throughput and high cost of performing mass spectrometry on all samples. In this instance, we felt that it was reasonable to pool samples since we were testing processing conditions overall and not the presence of any particular protein. Since we were looking for large differences in semi-tryptic cleavages, we felt that these were unlikely to be masked by one dominant sample in a pool, and found that pooling our samples was a cost-efficient way to assess for degradation. However, our pooled technique does not allow us to assess for statistical differences in the level of degradation in samples. Another limitation is that our metabolite analyses are mainly comparisons of the amino acids and biogenic amines that were detected in all of our samples. We did not specifically compare acylcarnitines, sphingolipids, and glycerophospholipids due to limited data within the range of detection, and we are unable to comment on whether processing techniques might affect results for these other metabolites.

In summary, for studies performed in a clinical setting, our findings may allow investigators to prioritize which processing conditions are most important for maximizing data from clinically-obtained urine samples. For research focusing on urine proteins, 1st AM voids may provide higher protein yields, but random or "spot" urine samples, which are logistically more feasible, appear to be a reasonable alternative. The addition of protease inhibitors at any void time significantly improves the amount of protein recovered, but may limit detection of peptides in mass spectrometry based proteomics. For research focusing on urine metabolites, certain metabolites exist in higher concentrations in 1st AM voids, and thus standardization of timing of voids may be more important. Bacteriostatic agents do not appear to offer significant benefit for samples processed and frozen on the same day, and samples that remain at room temperature for up to 4 hours are not significantly different than those that are immediately cooled.

References

1. Thongboonkerd V (2007) Practical points in urinary proteomics. J Proteome Res 6: 3881-3890.

2. Adachi J, Kumar C, Zhang Y, Olsen JV, Mann M (2006) The human urinary proteome contains more than 1500 proteins, including a large proportion of membrane proteins. Genome Biol 7: R80.

3. Thongboonkerd V, Chutipongtanate S, Kanlaya R (2006) Systematic evaluation of sample preparation methods for gel-based human urinary proteomics: quantity, quality, and variability. J Proteome Res 5: 183-191.

4. Collins FS, Tabak LA (2014) Policy: NIH plans to enhance reproducibility.

Nature 505: 612-613.

5. Barber MD, Walters MD, Bump RC (2005) Short forms of two condition-specific quality-of-life questionnaires for women with pelvic floor disorders (PFDI-20 and PFIQ-7). Am J Obstet Gynecol 193: 103-113.

6. Keller A, Nesvizhskii AI, Kolker E, Aebersold R (2002) Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. Anal Chem 74: 5383-5392.

7. Emwas AH, Luchinat C, Turano P, Tenori L, Roy R, et al. (2015) Standardizing the experimental conditions for using urine in nmr-based metabolomic studies with a particular focus on diagnostic studies: A review. Metabolomics 11: 872-894.

8. Decramer S, Gonzalez de Peredo A, Breuil B, Mischak H, Monsarrat B, et al. (2008) Urine in clinical proteomics. Mol Cell Proteomics 7: 1850-1862.

9. Thomas CE, Sexton W, Benson K, Sutphen R, Koomen J (2010) Urine collection and processing for protein biomarker discovery and quantification. Cancer Epidemiol Biomarkers Prev 19: 953-959.

10. Bottini PV, Ribeiro Alves MA, Garlipp CR (2002) Electrophoretic pattern of concentrated urine: comparison between 24-hour collection and random samples. Am J Kidney Dis 39: E2.

11. Havanapan PO, Thongboonkerd V (2009) Are protease inhibitors required for gel-based proteomics of kidney and urine? J Proteome Res 8: 3109-3117.

12. Standard protocol for urine collection and storage normal urine.

13. Theodorescu D, Wittke S, Ross MM, Walden M, Conaway M, et al. (2006) Discovery and validation of new protein biomarkers for urothelial cancer: a prospective analysis. Lancet Oncol 7: 230-240.

14. Schaub S, Wilkins J, Weiler T, Sangster K, Rush D, et al. (2004) Urine protein profiling with surface-enhanced laser-desorption/ionization time-of-flight mass spectrometry. Kidney Int 65: 323-332.

15. Lee RS, Monigatti F, Briscoe AC, Waldon Z, Freeman MR, et al. (2008) Optimizing sample handling for urinary proteomics. J Proteome Res 7: 4022-4030.

16. Bernini P, Bertini I, Luchinat C, Nincheri P, Staderini S, et al. (2011) Standard operating procedures for pre-analytical handling of blood and urine for metabolomic studies and biobanks. J Biomol NMR 49: 231-243.

17. Hepburn S,, Cairns DA, Jackson D, Craven RA, et al. (2015) An analysis of the impact of pre-analytical factors on the urine proteome: Sample processing time, temperature, and proteolysis. Proteomics Clin Appl 9: 507-521.

18. Zhou H, Yuen PS, Pisitkun T, Gonzales PA, Yasuda H, et al. (2006) Collection, storage, preservation, and normalization of human urinary exosomes for biomarker discovery. Kidney Int 69: 1471-1476.