TO: CDER Biomarker Qualification Program

FROM: Kellie B. Kelm, PhD, Deputy Director, Division of Chemistry and Toxicology Devices, FDA/CDRH/OIR/DCTD

THROUGH: Courtney H. Lias, PhD, Director, Division of Chemistry and Toxicology Devices, FDA/CDRH/OIR

RE: CDRH review of the analytical studies submitted for the FNIH Biomarkers Consortium Kidney Safety Biomarker Project Team and the Critical Path Institutes Predictive Safety Testing Consortium Nephrotoxicity Working Group qualification of a safety biomarker panel composite of six novel urinary kidney proteins

DATE: June 19, 2018

BACKGROUND: Individual biomarkers that make up the safety panel composite measure:

1. Urinary clusterin (CLU) – the secreted isoform (vs. nuclear) is measured for kidney injury. It's stated that CLU is expressed "in response to kidney injury in the proximal and distal tubules, glomerulus, and collecting duct".
2. Urinary cystatin-C (CysC) – a potential protein biomarker for glomerular filtration and tubular dysfunction.
3. Urinary Kidney Injury Molecule 1 (KIM-1) – a glycoprotein whose mRNA levels are elevated after initiation of kidney injury and the protein has been evaluated in many studies looking at kidney injury biomarkers.
4. Urinary N-Acetyl-beta-D-glucosaminidase (NAG) – a large enzyme where increased levels have been seen to associate with AKI and treatment with nephrotoxic compounds.
5. Urinary neutrophil gelatinase-associated lipocalin (NGAL) – NGAL is induced in epithelial cells with inflammation or other types of injury. With kidney injury, increased plasma and urine levels of NGAL have been seen.
6. Urinary osteopontin (OPN) – increased mRNA and protein levels of OPN have been seen after kidney injury

Laboratory: All urine biomarker assay validation data was generated at a single site laboratory, Pacific Biomarkers Incorporated (PBI). The PBI laboratory is a CLIA certified and College of American Pathologists (CAP) accredited laboratory.

Pre-analytical processes: Briefly, urine samples were collected in preservative free urine collection cups and were centrifuged at room temperature at 2000 x g for 10 minutes, aliquoted into cryotubes and frozen at -70°C within 3 to 4 hours. Samples were shipped on dry ice to a central storage facility and were maintained at -70°C. Frozen samples were

shipped on dry ice to the PBI laboratory. All samples analyzed for clusterin in the analytical studies for clusterin were desalted, except for precision.

CLU, CysC, KIM-1, NAG, NGAL, and OPN were measured with assays labeled "For Research Use Only" (RUO). Urinary creatinine (uCr) was measured using FDA-cleared assays according to manufacturer instructions when necessary as part of the clinical and analytical validation protocols. This uCr assay is standardized against ID-MS and the primary reference material used as a standard for the assay is the U.S. NIST SRM 914.

Urine Biomarker Assay Method and Manufacturer:

Assay	CLU	CysC	KIM-1	NGAL	NAG	OPN
Method	ELISA	ELISA	ELISA	ELISA	Colorimetric	ELISA
Manu-facturer	R&D Systems	R&D Systems	R&D Systems	BioPorto Diagnostics	Roche Diagnostics	R&D Systems
Range^ (calibrators)	3-200 ng/mL	3-100 ng/mL	30-2000 U/L	0.0-1.0 ng/mL	0.31 – 55.25 U/L*	0-20 ng/mL
Procedural dilution	4			100		440

^ Ranges inferred according to the concentrations of the calibrators as listed in the package inserts. Typically the analytical measuring range is defined in large part by the calibrators (however it can be larger if accuracy of dilution above the calibrators is validated).

* Assay has single calibrator. Range inferred using the lower limit of quantitation (LLOQ) and the upper end of the linearity studies.

Each laboratory using this biomarker will have to create their own quality control material. The quality control material should be urine-based controls and 2-3 levels are recommended to cover the measuring range of each analyte.

Only the assays stated above were used in the assessment of the analytical validity of these individual biomarkers; performance of other assays that measure these analytes is unknown. Other assays should be validated to confirm that they are appropriate for use as part of the safety panel composite measure.

CDRH REVIEW OF THE ANALYTICAL EVALUATIONS:

A note on analytical studies: In the clinical study, the biomarkers of interest and creatinine were measured in all clinical samples. The results for each biomarker in the clinical study was creatinine corrected. The analytical studies, with the exception of the reference range and the bridging study, were performed without creatinine correction.

I. Precision: In general, within-run precision was evaluated using low, medium and high in-house control urine samples assayed 16 to 20 times in one analytical run at PBI. The samples were preserved (i.e., treated with protease inhibitors and/or other treatments). Between-run

precision was evaluated using the same control samples used for within-run precision in three separate analytical runs.

1. CLU: Precision was performed using neat, centrifuged urine samples that were not desalted. (All other studies were performed using centrifuged and desalted urine samples.) Typically, additional preanalytical processing steps like desalting adds imprecision to tests.

Within-run precision results:

	Mean (ng/mL)	SD (ng/mL)	%CV	N
Low	16	1.7	10.5	20
Med	105	6.6	6.3	17
High	238	13.8	5.8	16

Between-run precision (CV) results:

	Mean (ng/mL)	SD (ng/mL)	%CV	N
Low	14	2.1	14.5	3
Med	101	4.4	4.3	3
High	226	10.7	4.7	3

2. CysC

Within-run precision:

	Mean (ng/mL)	SD (ng/mL)	%CV	N
Low	14.2	0.55	3.9	20
Med	27.2	0.81	3.0	20
High	79.4	3.44	4.3	20

Between-run precision:

	Mean (ng/mL)	SD (ng/mL)	%CV	N
Low	12.5	1.41	11.3	16
Med	23.8	1.59	6.7	16
High	70.9	4.09	5.8	16

3. KIM-1

Within-run precision:

	Mean (U/L)	SD (U/L)	%CV	N
Low	169	14	8.3	20
Med	579	6.1	1.1	20
High	1161	61.2	5.3	20

Between run precision:

	Mean (U/L)	SD (U/L)	%CV	N
Low	169	14	8.3	20
Med	579	6.1	1.1	20
High	1161	61.2	5.3	20

4. NAG

Within-run precision:

	Mean (pg/mL)	SD (pg/mL)	%CV	N
Low	2.41	0.061	2.5	20
High	9.90	0.187	1.9	20
Low dilx3	0.86	0.037	4.3	20

Low dilx3 is the low control diluted 3 fold.

Between run precision:

	Mean (pg/mL)	SD (pg/mL)	%CV	N
Low	2.45	0.137	5.6	24
High	9.96	0.245	2.5	24
In-house	2.57	0.115	4.5	24

5. OPN

Within-run precision

	Mean (ng/mL)	SD (ng/mL)	%CV	N
Low	807	25.7	3.2	20
Med	4524	176.6	3.9	19
High	1775	53.5	3.0	18

Between run precision

	Mean (ng/mL)	SD (ng/mL)	%CV	N
Low	797	99.6	12.5	7
Med	4413	441.8	10.0	7
High	1796	138.0	7.7	7

6. NGAL:

Within-run precision

	Mean (ng/mL)	SD (ng/mL)	%CV	N
Low	4.2	2.29	5.2	20
Med	20.0	1.85	9.3	20
High	44.5	2.29	5.2	20

Between run precision:

	Mean (ng/mL)	SD (ng/mL)	%CV	N
Low	3.9	0.26	6.7	10
Med	18.0	0.81	4.5	10
High	46.0	2.14	4.7	10

II. Dilutional Linearity:

Quantitative assays should be linear over their entire measuring range. The Clinical and Laboratory Standards Institute EP6-A: *Evaluation of the Linearity of Quantitative Measurement Procedures: A Statistical Approach, 1st Edition* states:

> "the mathematical relationship between the response and analyte concentrations should be sufficiently well defined to allow the selection of a suitable transformation of the dose-response curve to linear form. Furthermore, for an analytical method, one cannot interpolate between points unless one knows the results are linear.
>
> The reason for ensuring linearity of analytical methods is important on clinical grounds. Clinicians know that the relationship between an analyte and the pathophysiologic process is usually nonlinear, but they expect that the results reported to them by the laboratory include a linear relationship between the result and the true concentration, count, or activity recovered. For example, if the true amount of an analyte in a sample were to double, clinicians expect that would be reflected in a doubling of the measured value."

1. CLU:

Two linearity series were evaluated. Two desalted samples were each independently diluted with the calibrator diluent for a total of 7 concentrations for each series. Each sample was tested with a single replicate. The first series (linearity B) was prepared by serially diluting a single donor urine pool with clusterin value of approx. 470 ng/mL. The second series (linearity D) was prepared by serially diluting a single donor urine pool collected at PBI stabilized with protease inhibitor cocktail at approximately 255 ng/mL clusterin. The submitter looked at the linear regression and the % recovery at each dilution. The linear regression analysis for each dilution series is:

Linearity B:
Slope 1.005 ± 0.036
Intercept 11.5 ± 7.3

Linearity D:
Slope 0.993 ± 0.040
Intercept -7.0 ± 4.6

The B series has an intercept greater than the LLOQ. In addition, the % recovery analysis shows that in the B series a % recovery of 132.8% was obtained (outside +/- 20%) and the D series had a result at 72.7%.

2. CysC:

Two linearity series were evaluated. One sample was diluted with the calibrator diluent for a total of 7 concentrations in the series. Each sample was tested with a single replicate. The first series (linearity B) used a single donor urine pool with cystatin C value of approx. 90 ng/mL. This study gave results that met pre-validation specifications. However, a pattern of over recovery at higher cystatin C (close to 120% recovery at 45 ng/mL) suggested that dilutional linearity was less than might be expected. Therefore, a second linearity experiment was also performed to evaluate the recovery. Thus, it was demonstrated that if urine was first desalted (by gel filtration) and diluted with diluent containing carrier protein (4% BSA), the back calculated cystatin C results were much closer to the target results than for those obtained using untreated urine.

The linear regression analysis for each dilution series is:

Original:
Slope 1.020 ± 0.043
Intercept 1.50 ± 1.71

Desalted:
Slope 1.003 ± 0.004
Intercept -0.08 ± 0.14
Range = 1.4 – 91.1 ng/mL.

3. KIM-1:

One linearity series was evaluated. One single urine sample was diluted with reagent diluent for a total of 6 concentrations in the series. All obtained recovery within 20% of the expected over the range of 64-2051 pg/mL. Regression analysis results:

slope 1.001 ± 0.021
intercept -34.2 ± 20.0

The intercept (-34.2 pg/mL) is much larger than the absolute value of the LLOQ (11.6 pg/mL) and the recoveries ranged from 82 - 99% indicating systemic under-recovery for this assay.

4. NAG:

One linearity series was performed. A 63.0 U/L standard was diluted with saline to a total of 10 concentrations; each was tested by 3 replicates. 0.65 – 55.25 U/L was covered.

Slope 1.001 ± 0.004
Intercept 0.219 ± 0.114

This study gave results that met pre-validation specifications.

5. OPN:

Two linearity series were performed. Two urine samples were diluted with the calibrator diluent to 6 concentrations. Each sample was tested with one replicate.

Series 1

Slope 0.993 ± 0.011
Intercept -0.2 ± 19.5
Range 130-3864 ng/mL

Series 2

Slope 0.992 ± 0.0321
Intercept -60.6 ± 89.6
Range 284-9077 ng/mL

This study gave results that met pre-validation specifications (i.e., percent recovery was within 80-120% at all concentrations) however it was noted that the intercept from Series 2 (-60.6 ng/mL) is very large indicating potential concerns with the systematic negative bias across the range.

6. NGAL:

Two linearity series were performed. Two urine samples (one stabilized with protease inhibitor cocktail) were diluted with the sample diluent to 5 or 6 concentrations. Each sample was tested with two replicates.

Series 1:
Slope 1.005 ± 0.008
Intercept 0.30 ± 0.31

Range: 4.1 – 79.60 ng/mL

Series 2:
Slope 0.826 ± 0.007
Intercept 0.77 ± 0.34

Range 1.2 – 79.40 ng/mL.

This study gave results that met pre-validation specifications.

III. Estimates of the lower limit of quantitation (LLOQ):

The submitter used a single patient pool or low in-house urine-based control material and performed serial dilutions with calibrator diluent or saline to several samples with lower analyte levels. Each sample was analyzed multiple times over several analytical runs. The

submitter performed power fits of the data plotted against the CV for that level; the LLOQ at a CV of 20% was estimated using the equation for the curve.

1. CLU: LLOQ was calculated to be 2.5 ng/mL. Since samples are subjected to a 1:4 procedural dilution, this corresponds to a reportable limit of 10 ng/mL.

2. CysC: LLOQ was calculated to be 1.31 ng/mL.

3. KIM-1: LLOQ was calculated to be 11.6 pg/mL.

4. NAG: For NAG, after the predicted LLOQ was calculated, a two sample precision confirmation of the LLOQ was performed. Samples with means of 0.308 U/L and 0.222 U/L were assessed (10 replicates) and shown to have CVs of 15.3% and 22.3%, respectively, showing that the LLOQ is between these two samples.

Sample ID	**Mean (U/L)**	**St. Dev.**	**%CV**	**N**
QCUNAGL x5	0.31	0.047	15.3	10
QCUNAGL x6	0.22	0.052	23.3	10

LLOQ was determined to be 0.31 U/L.

5. OPN: The LLOQ at a CV of 20% was estimated to be 0.10 ng/mL. Since the samples are diluted 440-fold before analysis, this corresponds to a reportable level of 44 ng/mL.

6. NGAL: The LLOQ at a CV of 20% was estimated to be 0.45 ng/mL = 0.0045 ng/mL when taking into account the 1:100 dilution.

IV. Upper limit of quantification (ULOQ): The upper limit of quantification (ULOQ) is defined as the highest level of analyte with a CV of 20% or less. PBI used the highest reference standard for each assay that met these criteria as the ULOQ. That number was then multiplied by the maximum dilution where assay linearity was still maintained to obtain the upper reportable limit (URL).

VI. Recovery (accuracy):

Several spiked urine samples were used for each biomarker and diluted to 3-6 concentrations and assessed for recovery. Percent recovery for each dilution, the mean of the dilutions across the independent series, the overall mean recovery and the linear regression was evaluated for each.

1. CLU:

Series	N	Slope	Intercept	Error	Range (ng/mL)
B	6	1.007 ± 0.406	-4.7 ± 85.5	12.9	182-219
E	6	0.992 ± 0.051	2.1 ± 15.3	14.9	196-533
F	6	0.984 ± 0.014	10.0 ± 3.9	4.6	162-560

Adequate accuracy was established for clusterin over 162 – 560 ng/mL.

2. CysC:

Series	N	Slope	Intercept	Error	Range (ng/mL)
1	4	1.070 ± 0.087	-1.13 ± 1.22	0.40	10.7-17.3
2	4	0.971 ± 0.033	-0.07 ± 0.31	0.18	5.9-12.8
3	4	0.973 ± 0.101	1.26 ± 2.14	0.32	22.7-26.7
4	3	0.893 ± 0.037	-0.76±0.87	0.34	14.5-26.2

Recovery 4 (a 100 ng/mL standard was used) resulted in % recoveries below 90% (one was at 83.8%).

Adequate accuracy was established for clusterin over 162 – 560 ng/mL.

3. KIM-1

Series	N	Slope	Intercept	Error	Range (pg/mL)
A	3	1.094 ± 0.048	-41.6 ± 20.5	9.4	283-576
B	3	1.059 ± 0.024	14.1 ± 7.1	5.1	164-474
C	3	1.221 ± 0.053	-37.8 ± 19.4	10.8	225-571

Adequate accuracy was established for KIM-1 over 164 – 576 pg/mL.

4. NAG: Concentrated standard was spiked into an in-house control and low control to levels of 3%, 2%, and 1%. Each level was measured three times in one assay run.

Series	N	Slope	Intercept	Error	Range (U/L)
In-house	5	1.000 ± 0.002	0.026 ± 0.040	0.074	2.70 – 4.33, 56.98
Low	5	0.999 ± 0.001	0.086 ± 0.035	0.066	2.5-4.13, 56.98

Recovery series had large gap between the undiluted standard and the other dilutions at 3% and below. This study demonstrated the assay had adequate accuracy over the ranges evaluated.

5. OPN:

Single donor urine samples were protected at time of collection with protease inhibitors and ovalbumin. To determine recovery, three standards were spiked into calibrator diluent and urine samples to targets of 10.0, 5.0, and 2.5 ng/mL, respectively. Each sample was analyzed once in a single analytical run. All samples recovered within 80-120% of expected.

6. NGAL:

Diluted recombinant material was spiked into the urine samples (samples pre-diluted 1:100) at NGAL concentrations of 20%, 10%, 5%, and 2.5% of that of the high level sample.

The NGAL levels were determined in three different analytical runs (A – C below).

Analytical run	**N**	**Slope**	**Intercept**	**Error**
A	6	1.003 ± 0.004	-0.25 ± 0.15	0.28
B	6	0.998 ± 0.004	0.24 ± 0.14	0.27
C	6	0.987 ± 0.015	1.23 ± 0.51	0.85

Adequate accuracy was established for NGAL over 1.2 – 79.4 ng/mL.

V. Interference: PBI conducted a series of studies to evaluate the potential interference of high concentrations of albumin, hemoglobin, and blood contamination on the measurement of urinary CLU, CysC, KIM-1, NAG, NGAL, OPN and uCr. Albumin (5 mg/mL), hemoglobin (~30 mg/dL), and blood (0.2%) did not interfere with the determination of urinary biomarker concentrations. Cisplatin and tobramycin were also evaluated and did not interfere with the biomarkers.

Albumin interferes with test results in samples with low/normal concentrations of biomarkers as follows: 21.9% with clusterin, 21.2% with osteopontin, 28.2% with CysC, and 64.1% with KIM-1 as well as test results for samples with high concentrations of KIM-1 that exhibited 26.9% interference. Since the clinical samples were also evaluated for albumin levels. Samples with high albumin results will be flagged to note that these samples may have inaccurate results for clusterin, clusterin, osteopontin, CysC and KIM-1.

Hemoglobin at a concentration of 0.2% (~ 30 mg/dL, which is visually detectable) causes interference in samples with low/normal concentrations of biomarkers as follows: 22.7% with clusterin, 24.8% with CysC, 62.1% with KIM-1, 31.4% with NAG and 51.8% with total protein. In addition, hematuria (blood) at levels of 0.2% causes interference in samples with low concentration of biomarkers as follows: 218.8% with clusterin, 25.7% with CysC, 51.4% with

KIM-1, 636.3% with albumin and 299.3% with total protein. The submitter stated that samples with visible blood contamination in the clinical study will be documented. Furthermore, the data derived from these samples will not be included in the analysis. Likewise, data collected from samples with visible blood contamination will not be sent for adjudication.

pH and Specific gravity: PBI states that three of the biomarker assays (CysC, CLU,. KIM-1) incorporate a sample preparation step that corrects the pH and the specific gravity of the samples, therefore these were evaluated as potential interferents. Two of the other assays (NGAL, OPN) involve dilutions of small amounts of patient samples into large volumes of buffer, therefore pH and specific gravity aren't considered potential interferents. For NAG, pH is controlled by the buffered reagent substrate solution however specific gravity wasn't evaluated with no explanation provided.

VII. Hook effect: PBI stated that hook effects are not an issue in heterogeneous immunoassays; therefore the potential hook effect was not evaluated. It does appear that these immunoassays incorporate several washes and therefore the susceptibility to hook effect is minimized.

VIII. Reagent Stability: Not evaluated by PBI.

IX. Analyte Stability: To assess the stability of the biomarkers under different storage conditions, aliquots of urine were frozen at -70◦C and then used to examine analyte stability (both 2-8◦C storage and freeze/thaw cycle). The protocols do not mention whether centrifugation of the samples occurred. The results noted for stored samples were compared to those obtained for freshly thawed (baseline) samples. The actual difference and the percent difference from baseline were calculated at each time point. For each condition, the mean percent difference was also determined using the mean baseline and the mean difference. A mean difference of ≤ 10% was considered stable, with no individual changes of more than ± 20%.

1. CLU

Other stability data (i.e. for storage conditions) not provided for clusterin. RUO assay labeling states: Urine - Aseptically collect the first urine of the day (mid-stream), voided directly into a sterile container. Centrifuge to remove particulate matter, assay immediately or aliquot and store at -80°. Avoid repeated freeze-thaw cycles.

2. CysC:

Storage stability: Urine samples not stabilized with protease inhibitor cocktail stored at 2 – 8 °C are stable for at least 1 d (mean difference = -7.6%, 2/3 samples better than +/- 10% change). Samples were not stable through additional freeze/thaw cycles or at room temperature. Submitter claims urine samples stabilized with protease inhibitor cocktail stored at 2 – 8 °C are stable for at least 1 d (mean difference = -9.9%) however 2/3 were worse than +/- 10% change (37.7% and 18.4%). Samples were not stable at room temperature or through additional freeze/thaw cycles.

Urine samples stabilized with protease inhibitor cocktail, which were added to 5% BSA were stable for 3 h at room temperature (only 1 out of 2 samples better than +/- 10% change) and 1

d at 2 – 8 °C (only 1 out of 2 samples better than +/- 10% change, the submitter averaged the results). The samples were not stable for additional freeze/thaw cycles.

Conclusion by the submitter: It appears that cystatin C in urine samples treated with a protease inhibitor cocktail alone are stable for 1d when stored at 2 – 8 °C. However, when 5% BSA is added to the tubes the urine is incubated in before the incubation begins, protease stabilized urine samples were shown to be stable for 3 h at room temperature and 1 d at 2 – 8 °C. Therefore, it is recommended that samples be transferred to tubes containing some carrier protein immediately after collection.

3. KIM-1

Stability: Urine samples stabilized with protease inhibitor cocktail are stable when stored at 2 – 8 °C for 6 d (mean difference = 3.8%), 1 d at room temperature (mean difference = 7.7%), and up to three additional freeze/thaw cycles (mean difference = 0.8%). Out of the 3 samples tested, the lowest sample had significant stability issues under all conditions. The sample was at 118 pg/mL, which is above the 11 pg/mL LLOQ. Urine samples not stabilized with protease inhibitor cocktail are stable when stored at 2 – 8 °C for 3 d (mean difference = -2.1%), 1 h at room temperature (mean difference = 1.4%), and up to two additional freeze/thaw cycles (mean difference = -7.2%).

4. NAG: Five untreated urine samples were evaluated for analyte stability when stored at 2-8°C or when undergoing freeze-thaw cycles. Urine samples stored at 2-8°C were found to be stable for 5 days (change from baseline = -2.5%). The five samples were stable through two additional freeze/thaw cycles (change from baseline = -5.3%).

5. OPN: 3 samples were evaluated. Urine OPN stored at 2 – 8 °C was stable for up to 6 d (mean difference = -1%), and for up to 2 h at room temperature (mean difference = -1% but one sample > 20% change after 24 hr). Samples were also stable through three additional freeze/thaws (mean difference = 4%).

6. NGAL: 3 samples (stabilized) were evaluated. Samples stored at 2 – 8 °C was stable for up to 6 d (mean difference = 9%, one sample at 6 days had a 11% change), and for up to 1 day at room temperature (mean difference = 3.2%, no samples beyond 8.6% difference). Samples were also stable through three additional freeze/thaws (mean difference = -1.8%). Non-stabilized urine samples stored at 2 – 8 °C was stable for up to 3 d (mean difference = -0.6%, 2/3 samples at 3 days had a >11% change), and for up to 3 hr at room temperature (mean difference = 4.3%). The submitter claims samples were also stable through three additional freeze/thaws (mean difference = 8.1% but one of the samples had > 10% change).

A long term stability study for NGAL, clusterin, osteopontin, cystatin C and KIM-1 in centrifuged urine samples stored at -70C was provided. Samples were assessed at 12 and 18 months of storage; 24 month data was available for NGAL and clusterin. Stability at a time

point was acceptable if >67% of the samples met the acceptance criteria and an individual sample was considered unstable if it did not meet the acceptance criteria for two consecutive time points. Stability acceptance criteria was 25% difference from baseline for NGAL and osteopontin and 30% difference from baseline for clusterin, cystatin C and KIM-1. However, different lots, operators, etc. were used at each time point, therefore if the % change in results of the quality controls in the testing exceeded the acceptable between-run precision of the assay, then a correction factor was applied based on the linear regression analysis of the QC results at the time point in question vs. baseline. This process was used for the assessment of cystatin C stability.

The submitter states that osteopontin is stable for 12 months, clusterin was not stable at 12 months, NGAL, KIM-1 and cystatin C are stable for 18 months. The submitter did not evaluate NAG but stated that publications have support long term stability of NAG.

It's unclear whether instability of the analyte in the quality control would be a significant contributing factor to the % change of the results at each time point. It's also unclear how much of a % change in cystatin C levels due to instability was eliminated when the correction factor was applied. Consistently, all biomarkers had multiple samples at the 12 month assessment with % changes > 20% from baseline. Information demonstrating whether the broad acceptance criteria used in this study could impact the clinical use of these biomarker (i.e. the composite measure) for the proposed context of use was not provided. The Agency cautions users of this biomarker that long term storage could impact results; use of tighter acceptance criteria (like those used in the short term stability studies) would lead to shorter claims for long term stability.

X. Bridging study for sample handling and treatment: Due to a difference in the urine matrix used in the assay validation and those collected in the clinical studies to support qualification, PBI conducted a series of bridging studies to evaluate potential differences in the methods for measuring urinary CLU, CysC, KIM-1, NAG, NGAL, OPN and uCr. From these studies it was determined that the difference in urine matrices did not significantly affect the quantitation of the biomarkers and it was concluded that including a centrifugation step in urine sample processing, is preferred to reduce the potential contribution of biomarker activity from cellular contamination. Reported results from these assays were normalized to uCr as measured by the Roche Diagnostics enzymatic uCr assay.

Twenty single-donor urine samples from subjects without documented renal injuries, but including some that may have relatively higher kidney injury biomarker levels will be collected at PBI.

Each individual urine sample will be divided into three equal aliquots. These aliquots will be treated using three different protocols:

1. "N&C": Maintained neat and centrifuged in a swinging bucket centrifuge at ~ 2000 x g for 10 minutes at RT. This type of sample collection and processing was used in the clinical study.
2. "PIC": Treated with a buffered solution of protease-inhibitor cocktail, but not centrifuged.
3. "OA+PIC": Treated with a buffered solution of the protease-inhibitor cocktail and 2.5% ovalbumin, but not centrifuged.

These samples will be processed and placed at -70 to -80°C within 4 hours of collection for all three protocols.

	N&C vs. PIC			Crea corrected			Volume corrected		
	Slope	Intercept	r	Slope	Intercept	R	Slope	Intercept	r
CysC	1.04	1.49	0.9946	0.92	6.36	0.9694	1.04	1.49	0.9946
KIM-1	1.13	16.58	0.9947	1.14	0.0156	0.9899	1.13	16.58	0.9947
NGAL	1.14	-0.32	0.9976	1.10	0.21	0.997	1.14	-0.32	0.9976
OPN**	0.97	122.09	0.9924	1.16	22.45	0.8664	0.97	122.09	0.9924

** OPN is also treated with ovalbumin

PBI concludes that since all slopes are within 0.8 – 1.2, this shows that the analytical validation can be bridged to the clinical samples. FDA notes that the CysC, OPN, KIM1 intercept is above the LLOQ. Slopes beyond 0.90 – 1.10 indicate that the comparisons are not sufficient to state the sample types are equivalent. NAG was not assessed in this study. The comparison for clusterin did not support bridging, therefore all analytical validation studies were performed in the same sample matrix as the clinical validation studies.

Data provided by the submitter showed that some of the biomarkers were impacted by time of collection (first morning vs. midday urine collection) and whether the samples were stabilized (e.g. use of protease inhibitors) or not. Therefore, the submitter recommends that all samples be collected at the same time of day and be pre-analytically treated the same way in any study using this biomarker.

XI. Reference range verification:

PBI performed a reference range study for each biomarker using between 19-25 laboratory volunteers that were apparently healthy. Demographic statistics of the participants were not provided (e.g., gender, age, etc.) although one comment was that the group was "health conscious" and had a young median age < 30 indicates the participants may not be representative of the typical US healthy population. A different number of samples was tested for each biomarker. This small number of samples (19-25) is too small to determine a reference interval or normal range.

1. CLU: Nineteen single donor urine samples were analyzed. The results were averaged and the standard deviation was determined. The results were also normalized with urine creatinine as determined using the Roche Modular P automated analyzer. The reference interval was determined from the group mean ± 2 SD. The absolute range was also determined. The average urine sample clusterin level obtained was 231 ng/mL and the SD for the set of data was 160 ng/mL. Based on these results, the PBI-derived reference interval was < 551 ng/mL. The absolute range of clusterin concentrations observed was 51 – 650 ng/mL. When the urine results were normalized to urine creatinine the mean was 209 ng/mg and the SD was 87 ng/mg. The resulting reference interval was 35 – 383 ng/mg with an absolute range of 55 – 386 ng/mg.

2. CysC: Twenty-five single donor urine samples were analyzed. The results were averaged and the standard deviation was determined. The results were also normalized with urine creatinine as determined using the Roche Modular P automated analyzer. The rough reference interval was determined from the group mean ± 2 SD. The absolute range was also determined.

The average urine sample cystatin C level obtained was 43.4 ng/mL and the SD for the set of data was 29.56 ng/mL. Based on these results, the reference interval estimated at PBI was < 102.5 ng/mL. The absolute range of cystatin C concentrations observed was 10.2 – 117.5 ng/mL. These ranges compares reasonably well with the range levels found in normal urine reported by the assay manufacturer (12.6 – 188 ng/mL; mean = 62.9 ng/mL; standard deviation = 43.8 ng/mL), which was based on the results from 12 individuals. However, the uncorrected value is well below that reported in the most recent version of Tietz Clinical Guide to Laboratory Tests (< 280 ng/mL; see reference 6). However, the submitter suggests that assay standardization and the hydration status of the two groups, as well as the age of the subjects may have differed significantly and therefore, these values are not considered to be inconsistent.

3. KIM-1: Twenty-five single donor urine samples were analyzed. The results were averaged and the standard deviation was determined. The results were also normalized with urine creatinine as determined by the Jaffe method performed on the Roche Modular P automated analyzer. The PBI reference interval estimate was determined from the group mean ± 2 SD. The absolute range was also examined.

The average urine sample KIM-1 level obtained was 590 pg/mL and the SD for the set of data was 512.3 pg/mL. Based on these results, the reference interval determined at PBI was < 1615 pg/mL. The absolute range of KIM-1 concentrations observed was 133 – 1999 pg/mL. When the urine samples were normalized by ratioing with urine creatinine, the mean was 0.514 ng KIM- 1/mg creatine and the SD was 0.3386 ng KIM-1/mg creatinine. The reference interval determined at PBI was < 1.191 ng KIM-1/mg creatinine with an absolute range of 0.117 – 1.351 ng KIM-1/mg creatinine.

4. NAG: Twenty one urine samples were analyzed. The results were averaged and the standard deviation was determined. The PBI reference interval was determined from the group mean ± 2 SD. The absolute range was also determined.

The average uNAG measured was 2.16 U/L and the SD for the set of data was 1.239 U/L. Based on these results, the reference interval determined at PBI was <4.64 U/L. The absolute range observed was 0.59 – 5.73 U/L. When the results were corrected for urine creatinine, the average normalized value was 0.18 U/mmol of Creatinine and the standard deviation was 0.059. Thus, the normalized reference interval determined at PBI was 0.06 – 0.30 U/mmol of creatinine. Larger studies have been done on normal subjects of various age groups and both genders 5, 6. From these studies, a reference interval for normal subjects of <0.78 U/mmol of creatinine was derived. Submitter's data may be impacted by the health and age of the study participants (e.g. health conscious and young median age < 30).

5. OPN: Twenty single donor urine samples protected at time of collection with protease inhibitors and ovalbumin (10% dilution factor), were analyzed. The osteopontin concentrations were averaged and the standard deviation was determined. The PBI reference interval was determined from the group mean ± 2 SD. The absolute osteopontin concentration range was also examined.

The average urine sample osteopontin level observed was 1529 ng/mL and the SD for the set of data was 866.1 ng/mL (after normalization with urine creatinine the mean was 1262 ng/mg and the SD was 383.6 ng/mg). Based on these results, the reference interval determined at PBI was 495 - 2029 ng/mg normalized to urine creatinine (≤3261 ng/mL for non-normalized urine). The absolute range of osteopontin concentrations observed was 329 - 3957 ng/mL. When normalized to urine creatinine the range was 724 - 2293 ng/mg. The range of osteopontin levels for 21 normal urine samples reported by the assay manufacturer was in approximately the same broad range, 122 - 8796 ng/mL with a mean of 2044 ng/mL. These broad ranges are compressed by normalization to urine creatinine.

6. NGAL: Twenty-five single donor urine samples, collected at PBI and Solomon Park, were analyzed. The samples consisted of six single donor urine samples collected at PBI stabilized with protease inhibitor cocktail and Nineteen single donor urine samples collected at Solomon Park Research Laboratories (no details provided on whether they were treated).

The samples were from individuals that were apparently healthy. The results were averaged and the standard deviation was determined. The PBI reference interval was determined from the group mean ± 2 SD. The absolute urine NGAL range was also examined.

Two of the twenty-five reference range results were excluded from the calculations because the results were well segregated from the rest of normal urine results. The average urine sample NGAL level observed was 11.3 ng/mL and the SD for the set of data was 9.03 ng/mL (for samples normalized to urine creatinine the mean was 13.4 ng/mg and the SD was 14.20 ng/mg). Based on these results, the reference interval determined at PBI was <29.4 ng/mL (<41.8 ng/mg normalized to urine creatinine). The absolute range of NGAL concentrations observed was 1.0 – 288.0 ng/mL (1.0 – 43.5 without outliers). When normalized to urine creatinine the range was 0.9 – 191.2 ng/mg (0.9 – 52.7 without outliers). The range of NGAL levels for normal urine reported by the assay manufacturer are 0.7 – 9.8 ng/mL with a mean of 5.3 ng/mL, which was based on the results from seven individuals.

The submitter concludes the range of NGAL levels in the normal urines tested in this study included urines with higher NGAL than those observed at BioPorto. This is probably the results of the limited number of samples tested by BioPorto. Agreement in BioPorto and PBI assay performance is strongly indicated by the same lower NGAL levels.

Note that differences could be caused by different definitions of apparently healthy, difference in the expression of NGAL in gender, race or ethnicity, etc.

XII. Method comparison studies:

The submitter performed three studies called method comparison studies (summarized in 1-3 below) where samples were evaluated for several of the biomarkers of interest and the relationship(s) to each other was analyzed. In these studies, there is no measure of "truth"

which is needed for a method comparison study. A method comparison study is evaluation of a new/modified method to a known/reference/FDA cleared method. Given that there is no "truth" for these samples, this analysis is not useful.

1. Evaluate correlation between clusterin and other renal biomarkers. Nineteen single donor urine samples were analyzed for clusterin, cystatin-C, NGAL, KIM-1 and creatinine. Results were compared using Deming regression analysis.

X-method	Correlation Coefficient (R)	Slope	Intercept (ng/mL)
CYS-C	0.7645	7.684	-94.5
NGAL	-0.2098	-0.013	237.6
KIM-1	0.7049	0.239	96.3
CREAT	0.6055	3.974	-215.4

A conclusion was made that the results correlated significantly but "there is significant variation indicating that these biomarkers reflect different aspects of renal integrity". No acceptance criteria or definition of "significant correlation" was provided.

2. Correlation of cystatin C to creatinine in 22 samples was evaluated. Creatinine ranged from 37.0 to 266.6 mg/dL. CysC ranged from 4.3 to 117.5 ng/mL. Correlation coefficient is 0.9063 (slope 0.436). The conclusion was that "the results indicate a correlation between urine cystatin C and urine creatinine".

3. Renal Biomarker comparison: Twenty-five single donor urine samples collected at the laboratory were analyzed. The samples were from individuals that were apparently healthy. Each sample was assayed for urine creatinine, cystatin C, NGAL, and KIM-1, and results were compared.

The conclusions stated that urine creatinine levels modestly correlate with KIM-1 (R = 0.6350), but quite well with cystatin C (R = 0.9075). Interestingly, creatinine correlated poorly with NGAL (R = 0.3894 and 0.0593 with and without two outlier NGAL points, respectively). Cystatin C and KIM-1 were also moderately correlated (R = 0.6202). Results were not normalized.

4. Urine NAG comparison on Roche modular P analyzer and Mira Analyzer: thirty-five urine specimens frozen at -70ºC were thawed and used for the method comparison. The 35 samples analyzed for uNAG on the two analyzers produced a correlation coefficient of 0.9661, with a slope of 0.856 and an intercept of 0.09 U/L. The average bias seen was -0.72 U/L.

Samples 19 and 35 were excluded from the method comparison because the results from the Cobas Mira were above the linear limit for the instrument. A slope of 0.856 indicates lack of correlation and it's observed that the Mira appears to have positive bias at the higher end of the measuring range.

XIII. Analytical validation of the composite measure: There was no analytical validation of the safety panel composite measure provided for FDA review.