Review of Qualification Data for Biomarkers of Nephrotoxicity Submitted by the Predictive Safety Testing Consortium

Biomarker Qualification Review Team

Melanie Blank
Albert De Felice
Federico Goodsaid
Patricia Harlow
Elizabeth Hausner
David Jacobson-Kram
William Taylor
Aliza Thompson
Douglas Throckmorton
Shen Xiao

Center for Drug Evaluation and Research
U.S. Food and Drug Administration

January 16, 2009

Table of Contents

1. Executive Summary **4**

a. Background 4
b. Sources of Data and Major Findings 4
c. BQRT Conclusions 5
d. BQRT Recommendations 5

2. Background **6**

a. Overview of the Problem 6
b. The Biomarker Qualification Pilot Process 6
i. Definition of Qualification and of Qualification in Context 6
ii. Overview of Pilot Qualification Process 7
c. Commonly Used Laboratory Animal Species for Nonclinical Safety Assessments of Nephrotoxicity 8
d. Detection of Drug-Induced Nephrotoxicity in Animal Models 8
e. Biomarkers of Drug-Induced Nephrotoxicity Proposed by PSTC 9
f. Context Claims Submitted by PSTC for the Qualification of Proposed Biomarkers of Drug-Induced Nephrotoxicity 12

3. Summary of the Supporting Data Submitted by PSTC for the Qualification of Proposed Biomarkers of Drug-Induced Nephrotoxicity **13**

a. Overall Summary of Results 13
b. Summary of Studies Conducted and Biomarkers Measured 16
c. Histopathology Lexicon and Scoring 18
d. Summary Tables of ROC Curves 19
e. Statistical Analysis 20
f. Individual Animal Data 22
g. Recovery Studies 26
h. ROC Analysis of Different Histopathology Lesions 26

4. Reviewer Discussion of Qualification Data **27**

a. The Qualification Process 27
b. Analytical Validation 29
c. Correlation between Histopathology and Biomarker Data 35
i. Exclusion of Animals from ROC Analysis 35
ii. Blinding of Histopathology Analysis 35
iii. ROC Curves Generated for Different Histopathology Grades 43
d. Performance of Proposed Biomarkers Compared with Accessible Biomarkers in Current Use 43
i. Collection of Samples 43
ii. Blinding of Biomarker Assays 44

iii. Urinary creatinine Normalization 44
iv. Background (Control Range) Biomarker Levels 46
v. Threshold Definition 46
vi. Specificity and Secondary Organ Effects 47
vii. Unusual Findings 47
e. Format Issues 48
i. Data for Individual Animals 48
ii. Data from Independent Histopathologist 48

5. Qualification Conclusions 49

a. BQRT Conclusions 49
b. Recommended Application context for voluntary use of these urinary biomarkers:
i. Kim-1 49
ii. Albumin 49
iii. Total Protein 49
iv. β2 Microglobulin 49
v. Cystatin C 50
vi Clusterin 50
vii TFF-3 50
c. Strengths and Limitations of Submitted Data 50
d. Regulatory Recommendations 53

6. Appendices 55
a. Background Information on Proposed Biomarkers 55
i. Kim-1 55
ii. Albumin 56
iii. Total Protein 57
iv. β2 Microglobulin 58
v. Cystatin C 59
vi Clusterin 60
vii TFF-3 61
b. Additional Data Supporting Qualification of Proposed Biomarkers 62
i. PSTC Standardized Kidney Histopathology Lexicon 62
ii. Summary of Individual Studies 64
iii. Summary Tables of ROC Analysis using an Exclusion Model 66
iv. Statistical Analysis using an Exclusion Model 67
v. Statistical Analysis to test whether a Biomarker Adds Value to sCr and BUN 68
vi. PSTC Flow Chart of Sample Handling 69
vii. PSTC Discussion of "Prodromal" Relative to Exclusion and Inclusion Analysis 70
viii. Novartis ROC Analysis of Different Histopathology Lesions 72
ix. Additional PSTC discussion of histopathology procedures 73

7. References 74

Biomarker Qualification Review

1. Executive Summary

This is a review by the Biomarker Qualification Review Team (BQRT) of a submission by the Predictive Safety Testing Consortium (PSTC) for the preclinical qualification of seven urinary biomarkers of nephrotoxicity.

a. Background

The goal of the Pilot Process for Biomarker Qualification at the FDA is to test a process for biomarker qualification through proposals submitted by scientists from multiple organizations. Successful application of this process requires cooperation and communication between sponsors and regulatory scientists.

The FDA believes that biomarker qualification should be an incremental process. Initially, biomarkers will be qualified in a restricted context for use; as additional data becomes available, this context for use can expand. This document discusses what was learned about the qualification process itself and the recommendations for future submissions.

The limited sensitivity and specificity of accessible biomarkers of nephrotoxicity in current use [particularly blood urea nitrogen (BUN) and serum creatinine (sCr)] prevent early detection of drug-induced kidney toxicity. This is an area where the development and qualification of better biomarkers is urgently needed.

b. Sources of Data and Major Findings

Over a period of several months the PSTC submitted a data package supporting the nonclinical use of seven urinary biomarkers of drug-induced kidney toxicity in rats. The performance of these biomarkers was compared with that of current accessible biomarkers, sCr and BUN, against the gold standard of histopathology using Receiver Operating Characteristic (ROC) curves.

In studies conducted at three independent sites, these data were obtained for five to sixteen model nephrotoxicants and two to nine control non-nephrotoxicants for each novel biomarker. Only two nephrotoxicants and no control compounds were studied across all three sites. Despite this limited database, ROC curves for some of these biomarkers show improved specificity and sensitivity at detecting tubular and/or glomerular renal injury compared to BUN and sCr. Since the data used to construct the ROC curves came from a limited number of compounds, animals and independent studies, the utility of these biomarkers for use in a broader context is uncertain. During the review, concerns were raised about the

need for a minimum level of analytical validation of biomarker assays, and the advisability, *a priori*, of more extensive tissue sampling, and of unbiased, blinded histopathology evaluations in biomarker qualification.

c. BQRT Conclusions

Despite the above concerns, the urinary kidney biomarkers (KIM-1, albumin, total protein, β2-microglobulin, cystatin C, clusterin and trefoil factor-3) are considered acceptable biomarkers for the detection of acute drug-induced nephrotoxicity in rats and can be included along with traditional clinical chemistry markers and histopathology in toxicology studies. These biomarkers may be used voluntarily as additional evidence of nephrotoxicity in nonclinical safety assessment studies to complement the standard data (BUN and sCr). In ROC analyses, some of these biomarkers showed better sensitivity and specificity than BUN and sCr relative to histopathological alterations considered to be the gold standard when tested with a limited number of nephrotoxicant and control compounds. While further studies are needed to support a broader application claim, the data submitted thus far appear to be sufficient to support the voluntary testing in specific drug development programs on a case-by-case basis as proposed by the PSTC.

d. BQRT Recommendations

The submitted data support the voluntary use of urinary biomarkers KIM-1, albumin, total protein, β2-microglobulin, cystatin C, clusterin and trefoil factor-3 in preclinical research alongside histopathology to identify drug-induced acute kidney injury in the rat. Further studies are needed to improve the characterization of these markers in different animal models with different drugs and how to interpret different biomarker levels.

In order to gain useful information about biomarker performance in different contexts, including the clinical setting, we recommend the following:

1. A standardized format for submitting preclinical and clinical data is needed for an efficient and accurate review.

2. Consistency in approach, analysis, and presentation is a goal for biomarker qualification submissions. The achievement of this goal will facilitate comparisons between biomarkers of renal toxicity. For example, this will allow comparison of data for different biomarkers submitted by one or more investigators. Moreover, it will help with the creation of databanks that may allow expansion of the qualified context of use of these biomarkers.

3. In concurrence with the PSTC proposal, preclinical studies should demonstrate early detection of drug-induced renal injury and reversibility after drug cessation prior to proceeding to clinical studies.

4. We recommend that the regulatory Division review the protocol and study report for the preclinical reversibility study prior to its conduct.

5. Data are needed to address the correlation between biomarker levels and evolution of lesions with secondary confirmation using appropriate techniques, such as immunohistochemistry and/or in-situ hybridization, when appropriate relative to the biology of the biomarker and any claims concerning localization of injury.

6. While novel renal biomarkers should be tested in humans, they are not currently qualified to be used as primary renal injury monitoring tests or dose-stopping criteria. For the time being, the sponsor and regulatory division will decide on a case by case basis how best to implement these biomarkers in a clinical development program.

2. Background

a. Overview of the Problem

Biomarkers are measurable characteristics used as indicators of physiologic, pathologic and pharmacologic processes. Many commonly used biomarkers lack sensitivity and specificity for early drug-induced organ damage. In particular, drug development has been hampered by a lack of accessible markers of renal injury. Although sCr, BUN, and creatinine clearance have traditionally been used to monitor for drug-induced renal toxicity, these biomarkers are poor predictors of drug-induced renal damage because they lack sensitivity and specificity for renal injury and provide little information on the region of the kidney affected by the drug and/or the mechanism(s) by which this injury occurs. As a result, much research has focused on the development of novel biomarkers of renal toxicity.

In order to improve drug development, the Critical Path Opportunities Report calls for the identification of new safety biomarkers to (1) better identify early toxicity in animal studies, (2) aide in initial dose selection in clinical studies, and (3) improve safety monitoring in phase 1 and 2 clinical trials. Under the FDA Critical Path Initiative, biomarkers will be qualified on the basis of data that support their proposed use in a specified context. The FDA seeks to facilitate the development of biomarkers of renal toxicity by establishing a clear and rigorous process for biomarker qualification.

b. The Biomarker Qualification Pilot Process

i. Definition of Qualification and of Qualification in Context

Biomarkers may be accepted or qualified by different processes. A passive approach uses the peer-reviewed scientific literature and scientific consensus to establish that a biomarker is acceptable for a particular purpose. Passive qualification may be a time-

consuming process with a long lead time from the discovery of a new biomarker to its qualification. The current unstructured process of biomarker acceptance is not efficient nor a reliable process for identifying biomarkers for regulatory decision making.

An active, context-dependent biomarker qualification process provides clearly defined, explicit metrics for incremental success, and hence facilitates the development of novel biomarkers. Thus a uniform, consistent and explicit interpretation of a biomarker measurement in a specific context must be an integral part of biomarker qualification. Under the FDA Critical Path Initiative, biomarkers will be qualified on the basis of the data that support their proposed use in a specific context.

The following criteria for qualification were emphasized by the BQRT:

1) *Only those biomarkers that show potential for a positive impact on drug development will be evaluated for qualification.*

2) *A new biomarker must demonstrate that it provides better than or different information from biomarkers in current use.*

3) *Comparison of the new biomarker against currently accepted biomarkers relative to a widely accepted standard such as histopathology is necessary for a robust qualification process.*

4) *Biomarker qualification for clinical use requires the definition of a range of normal versus abnormal values.*

5) *The technological framework and assays supporting biomarker measurements must be analytically validated.*

ii. Overview of Pilot Qualification Process

The FDA is testing a pilot process for biomarker qualification through the use of Voluntary eXploratory Data Submissions (Goodsaid F., Frueh F. 2007b; Goodsaid F., Frueh F. 2008; Orr, M.S., et al 2007; Goodsaid F., Frueh F. 2007a). It is anticipated that this process will be tested and refined by multiple qualification proposals. The current process follows a number of steps:

1) Submission to the Interdisciplinary Pharmacogenomic Review Group (IPRG) of a preliminary proposal defining the biomarker, proposed context of use, and supporting data. Proposals may be received from industry, academia, government or other individuals within the scientific or medical community.

2) Evaluation of the proposal by the Biomarker Qualification Review Team (BQRT).

3) Decision by the BQRT about whether or not to proceed to a full qualification process.

4) Full submission of qualification data for review by the BQRT.

5) Voluntary eXploratory Data Submission (VXDS) meeting to go over the qualification data and to identify potential information gaps before a full review of the qualification package is completed

6) Full review by BQRT.

7) Internal review and regulatory decision at FDA.

8) Communication of decision to sponsor.

c. Commonly Used Laboratory Animal Species for Nonclinical Safety Assessments of Nephrotoxicity

Species differences in susceptibility to drug-induced toxicity can be attributed to important differences in the rate and extent of absorption of a drug, its rate and type of metabolic conversions, its detoxification mechanisms and excretion. Toxicology studies have been conducted in many species, including the mouse, rat, guinea pig, hamster, rabbit, dog, mini-pig, rhesus monkey, and cynomolgus monkeys. Regulatory guidance (ICH M3, 1997) calls for toxicity testing of pharmaceuticals in two mammalian species – one rodent and one non-rodent, prior to first-in-man clinical studies. Sponsors have generally evaluated renal function in the general toxicology studies using species, such as the rat, dog and monkey, having kidneys with intermediate-sized renal medullas similar to the human kidney (Berndt 1976).

d. Detection of Drug-Induced Nephrotoxicity in Animal Models

Currently, a number of parameters are used to assess renal function and damage in nonclinical safety assessment studies. These include the following:

- *Physical Exam and Behavior:* changes in weight, water consumption and/or drinking behavior, or findings of dehydration (skin tenting).
- *Hematology:* changes in packed cell volume or hematocrit (as indicators of changes in hydration or as a secondary effect on RBC count and hemoglobin). For longer duration exposure, progressive, non-regenerative anemia is usually noted, as well as changes in shape and size of RBCs.
- *Serum chemistries:* changes in Cr, BUN, potassium, phosphorus, calcium and their ratios, and acid base balance, hypoalbuminemia and hypercholesterolemia (as an indicator of glomerular injury leading to a nephrotic state), and hyperamylasemia and hyperlipasemia (as a result of impaired renal excretion).
- *Urinalysis:* the presence and appearance of cells, casts, and crystals, the quantity and type of protein excreted and changes in

urine color, volume, specific gravity and electrolyte concentration/excretion.

- *Histopathology:* is the gold standard for structural changes.

Tests of renal function are not part of the core battery of safety pharmacology studies (ICH S7a). Renal safety pharmacology is generally performed only if there is cause for concern. However, clinical chemistry and urinalysis assays are recommended in general toxicology studies. Abnormalities appearing in these studies could trigger more extensive safety testing.

e. Biomarkers of Drug-Induced Nephrotoxicity Proposed by PSTC

Previously published data on genomic biomarkers of nephrotoxicity (Han et al 2002, Silkensen et al 1997, Verstrepen et al 2001, Amin et al 2004, Thompson et al 2004) support the investigation of a number of accessible protein biomarkers in the rat (Han et al 2002) and monkey (Davis et al 2004). The C-Path Predictive Safety Testing Consortium (PSTC) considered a total of twenty-two (22) biomarkers. However, according to their submission, insufficient data exist for many of these biomarkers to support a claim for qualification.

The PSTC made an original and eight supplementary submissions to the FDA and the EMEA between June and November 2007 to support the preclinical qualification of seven pre-clinical biomarkers of drug-induced acute kidney toxicity. In addition, the PSTC provided a review of the scientific literature on the clinical experience with these biomarkers (see Appendix 6a). The following table provides an overview of key characteristics of these seven biomarkers: kidney injury molecule-1 (KIM-1), albumin, total protein, clusterin, cystatin C, β2-microglobulin, and trefoil factor-3 (TFF-3).

Table 1. Characteristics of Exploratory Biomarkers of Nephrotoxicity

Urinary marker	General attributes	Proposed mechanism by which increased urinary levels seen during kidney injury	Background Data Sited by Sponsor
KIM1	Transmembrane glycoprotein In early studies, shown to be minimally expressed in normal adult rat kidney and markedly expressed after ischemic injury to proximal tubule	Expressed in response to injury Proteolytically processed domain of KIM-1 detected in urine after injury	Up-regulation seen: (1) Models (drugs or environment toxicants) with predominately tubular injury: cisplatin, cadmium, chromium, folic acid, mercuric chloride, ochratoxin, sevoflurane, S-(1,1,2,2-tetrafluroethyl-1-cysteine, (2) Other models: ischemia reperfusion injury, endotoxin, brain dead donor, kidney fibrosis, polycystic kidney disease (animal model), uninephrectomy + BSA administration (proteinuria induced tubular injury), renal cell carcinoma, chronic cyclosporine toxicity (arteriolopathy, tubulointerstitial injury), adriamycin nephrosis (glomerular injury associated with heavy proteinuria)
Albumin	Most abundant plasma protein Widely used in clinical practice as marker of kidney damage in diabetics and end organ damage (kidney) in hypertensives	Classical view: Increased filtration in setting high intraglomerular pressures and/or injury to the filtration barrier. Recent controversy: While traditionally thought not to be significantly filtered, some studies now suggest may be freely filtered and extensively recycled by proximal tubule. If true, increased levels could also be seen with tubular injury.	Increased urinary levels seen: (1)Models (drugs or environment toxicants) with predominately tubular injury: gentamicin, ifosfamide, cis-diamminedichloroplatinum, carboplatin, carbapenem, cadmium ±arsenic, (2) Other models: diabetes and other nephrotic states (glomerular injury), hypertension (vascular injury), hyperfiltration, aminonucleoside administration (glomerular injury associated with heavy proteinuria)
Total Protein	Widely used in clinical practice as a marker of kidney damage, to monitor disease progression and response to therapy	Increased filtration in setting injury to filtration barrier Tubules typically reabsorb protein, hence increase can also be seen with tubular damage	Sponsor emphasizes its use as a prognostic marker of progressive loss of renal function/adverse renal outcomes

Clusterin*	Widely distributed heterodimeric glycoprotein Highly expressed during early states of kidney development, and following glomerular, tubular and papillary injury in animals	Expressed in response to injury	Up-regulation of gene/increased urinary levels seen: (1) Models (drugs or environment toxicants) with predominately tubular injury: sevoflurane (2) Other models: ischemia/ reperfusion injury, puromycin nephrotoxicity (glomerular injury associated with heavy proteinuria), nefiracetam (renal papillary necrosis), nephrectomy, unilateral ureteral obstruction, autosomal-dominant polycystic kidney disease (rat model), glomerulonephritis, acute and chronic transplant rejection, renal dysplasia
Cystatin C	Non-glycosylated low molecular weight protein synthesized continuously produced by nucleated cells Freely filtered and reabsorbed and metabolized by tubules Research also focused on use of serum concentration as measure of GFR	Impaired recycling by proximal tubule as a result of damage to the tubule and/or increased competition with other proteins for tubular transport (e.g. following glomerular damage that has led to a significant leakage of protein into the filtrate)	Sponsor cites studies showing increased urinary excretion in polycystic kidney disease, pyelonephritis and nephrotic syndrome in humans
B2 microglobulin	Single polypeptide chain that is a constituent of the class 1 MHC complex present on membrane of all nucleated cells Freely filtered and reabsorbed and metabolized by tubules	Impaired recycling by proximal tubule as a result of damage to the tubule and/or increased competition with other proteins for tubular transport (e.g. following glomerular damage that has led to a significant leakage of protein into the filtrate)	Increased urinary levels seen: (1) Models (drugs or environment toxicants) with predominately tubular injury: tenofovir disoproxil fumarate, gentamicin, cisplatin, heavy metal exposure (2) Other models: proteinuric states (including membranous nephropathy)
Trefoil factor-3	Peptide present in collecting ducts in normal kidneys, and in other epithelia	Decreased in response to injury	According to sponsor, TTF3 has been studied extensively at Merck as a potential biomarker of proximal tubular injury. Little additional background information is provided.

*For the purposes of this review, clusterin refers to the secreted isoform of clusterin and not the nuclear isoform.

f. Context Claims Submitted by PSTC for the Qualification of Proposed Biomarkers of Drug-Induced Nephrotoxicity

The PSTC makes the following claims for the biomarkers submitted for qualification:

- *The proposed markers 'add information' to sCr and BUN, while six of the seven were also shown to outperform one or both of these clinical chemistry markers.*
- *These kidney biomarkers correlate to either tubular histomorphologic alterations or to glomerulopathy with functional tubular involvement.*
- *Biomarker claims that apply more accurately to acute drug-induced kidney histomorphologic change which are supported by data submitted rather than more traditional chronic kidney injury.*
- *Voluntary use of these biomarkers is claimed by sponsors in preclinical GLP studies.*
- *Voluntary use is proposed of several of these urinary biomarkers (KIM-1, albumin, total protein, ß2-microglobulin, and cystatin C) as bridging markers for early clinical studies on a case-by-case basis when nephrotoxicity is seen in GLP animal toxicology studies. These data would be submitted together with other clinical data to support their use as sensitive biomarkers of kidney injury in humans.*

The PSTC summarized their claims in Table 2, which is based on analyses that included all of the data from the different studies.

Table 2: PSTC Claims: Accessible and Qualified Biomarkers for Regulatory Decision Making that Enable Drug Development

Urinary Biomarker	Qualified Preclinical	Adds inform. SCr and BUN	Outperforms SCr and/or BUN	Analytically Validated Assay	Widely Available Assay	Qualified Clinical
KIM-1	Yes	Yes*	Yes*	Yes	Pending	Yes
Albumin	Yes	Yes*	Yes*	Yes	Yes	Yes
Clusterin	Yes	Yes*	Yes*	Yes	Yes	Pending
TFF3	Yes	Yes*	No	Yes	Pending	Pending
Total Protein	Yes	Yes**	Yes**(SCr)	Yes	Yes	Yes
Cystatin C	Yes	Yes**	Yes**	Yes	Yes	Yes
β2 μ glob.	Yes	Yes**	Yes**	Yes	Yes	Yes

* =Acute tubular alterations
** =Acute glomerular injury with acute tubular re-absorption impairment
Biomarker outperformed (SCr)

3. Summary of the Supporting Data Submitted by PSTC for the Qualification of Proposed Biomarkers of Drug-Induced Nephrotoxicity

a. **Overall Summary of Results**

The data from short term rat GLP toxicology studies conducted at Merck and Novartis were evaluated through the joint FDA/EMEA pilot qualification process in an iterative manner. The results from studies supporting this qualification are summarized using Receiver Operating Characteristic (ROC) curves, which are plots of true positives (sensitivity) against false positives (1-specificity). This is the method of choice to characterize the performance of diagnostics (Metz CE, 1978). In such analyses, the "area under the curve" (AUC) for an ideal biomarker has a value of 1, while the AUC for a biomarker yielding random values is 0.5.

This review focuses on the results from ROC analyses that included all of the data from the different studies (inclusion analysis). Comparison of the performance of each new biomarker to the accepted biomarker standards of BUN and sCr was performed by comparison of the area under the curve (AUC) of the ROC analysis for each new biomarker with similar data for BUN and sCr. Histopathology was used as the gold standard that defined injury. Statistical comparison of the ROC curves is described further below. ROC curves were generated both for data merged from all positive histopathology scores for all studies by study site as well as for data from subset ranges of these scores. The ROC curves for the complete KIM-1 and albumin data from Merck are shown in Figure 1, while the ROC curves for the complete KIM-1, clusterin, total protein, β-2 microglobulin, and cystatin data from Novartis are shown in Figure 2. As shown in Figure 1, according to the Merck inclusion analysis, the AUC value for KIM-1 and albumin was greater than the AUC values for sCr and BUN in detecting kidney injury as determined by histopathology. As shown in Figure 2, according to the Novartis inclusion analysis, the AUC values for KIM-1, clusterin, total protein, β-2 microglobulin, and cystatin were also greater than the AUC values for sCr and BUN. The PSTC also provided exclusion ROC analyses, which excluded data from nephrotoxicant treated animals without positive histopathology, and analyses to test whether a marker adds value to sCr and BUN (see Appendices 6.b.iii, 6.b.iv and 6.b.v).

Figure 1: *Inclusion* model - All Histopathology Grades – All Merck data

AUC (area under the curve) SEN (sensitivity at 95% specificity) are shown. Note that compound treated animals with grade 0 histopathology were *included* for this analysis. ROC curves were generated using all histopathology grades.

KIM-1 **Albumin**

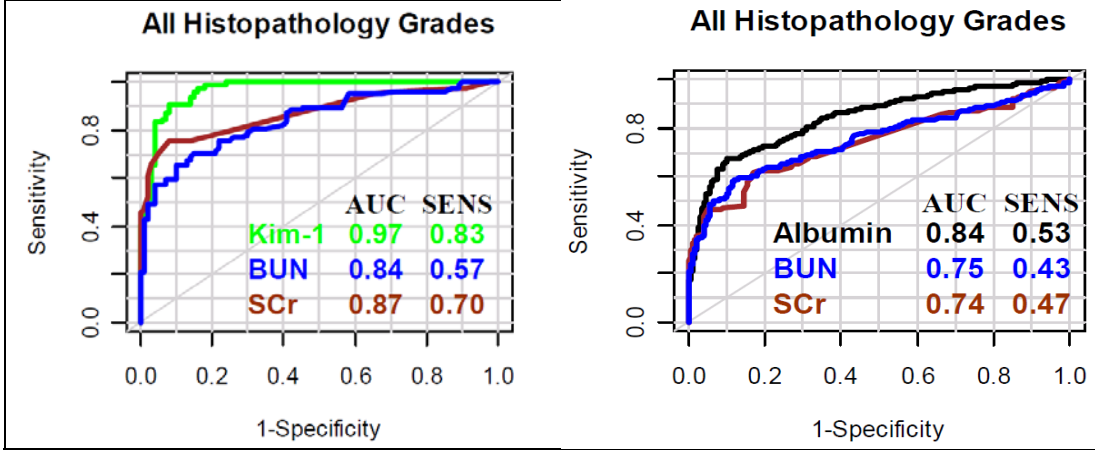
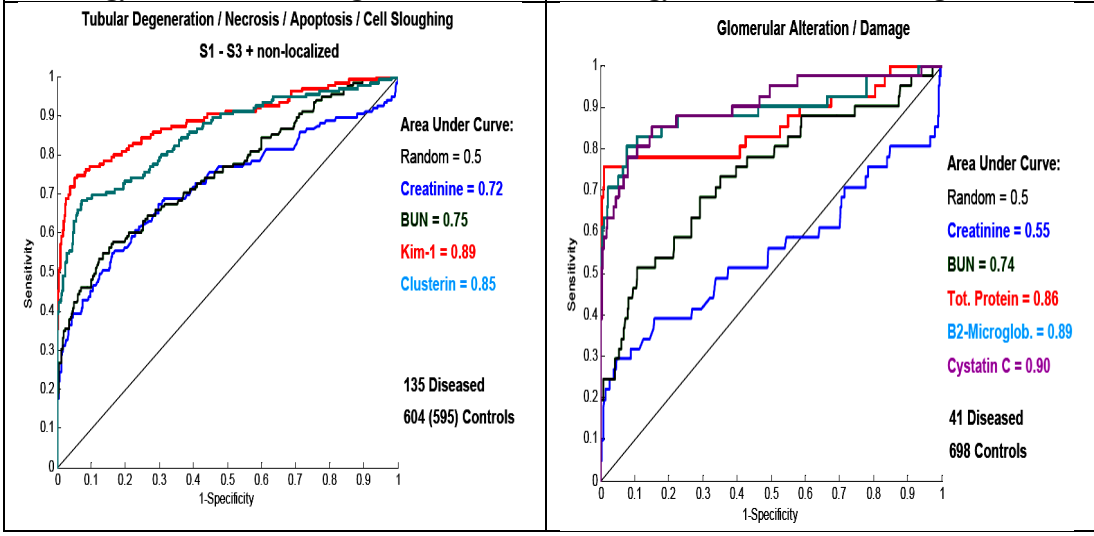


Figure 2. *Inclusion* model - All Histopathology Grades – All Novartis data (VXDS02)

ROC inclusion curves of tubular and glomerular markers. The analysis using animals with all histopathology grades (0-5) included compound treated animals with grade 0 histopathology. The AUC (area under the curve) values and the number of animals (KIM-1 in parentheses) are listed.

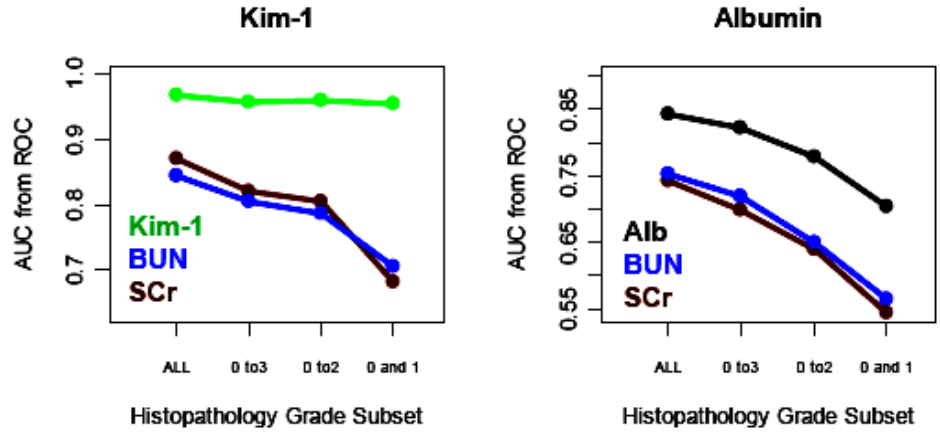
Pathology of "Tubular Damage" **Pathology of "Glomerular Damage"**



ROC data generated for different histopathology ranges of the complete Merck and Novartis datasets are summarized in Figure 3 below, which show the AUC value plotted as a function of the histopathology grade. As shown, for each histopathology grade subset, the AUC value was greater for KIM-1, albumin, clusterin, total protein, β-2 microglobulin, and cystatin than it was for sCr and BUN.

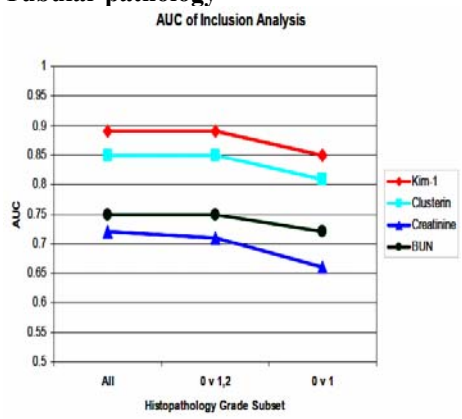
Figure 3: AUC versus. Histopathology Grade Maximum Composite *Inclusion* Model
Each plot shows the AUC from ROC analysis versus. the subset of samples used based on histopathology grade. The plots show the AUC for the biomarker, BUN, and SCr. Note that compound treated animals with grade 0 histopathology were *included* for this analysis.

Merck data

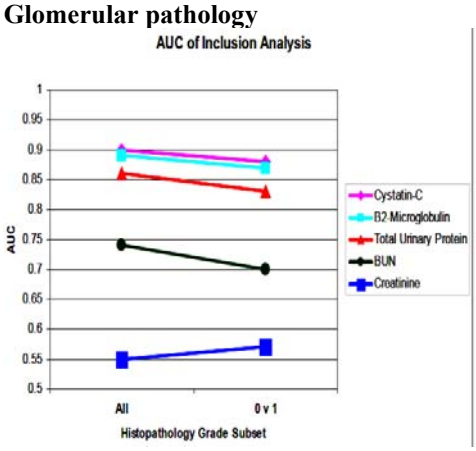


Novartis data

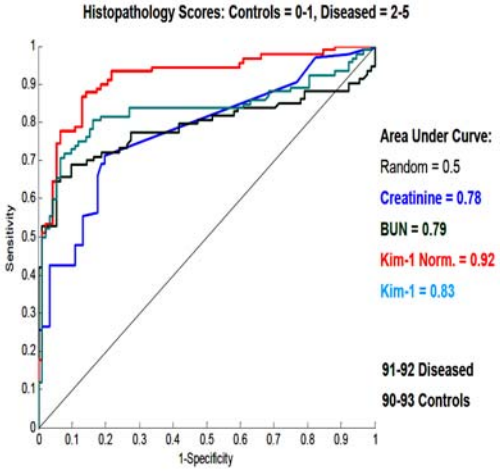
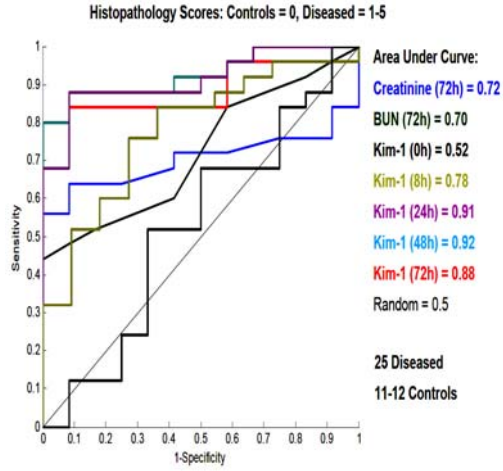
Tubular pathology



Glomerular pathology



The ROC analysis for the FDA data on KIM-1 is shown in Figure 4 below. The FDA time-course Cisplatin study was evaluated separately, because KIM-1 was measured in additional non-terminal urine samples (0, 8, 24, 48, and 72hr). The urinary KIM-1 concentrations at each collection time point were treated as separate parameters and ROC analyses were performed. As shown, the AUC value for KIM-1 was greater than the AUC values for sCr and BUN at the 8, 24, 48 and 72 hour timepoints.

Figure 4. *Inclusion* model – FDA data	
All data	**Cisplatin Time-Course Study**
ROC curves for SCr, BUN, KIM-1 normalized to urinary Creatinine and non-normalized. The control group contains animals with a histopathology score of 0 or 1. The diseased group contains animals with a histopathology score of 2 or higher. For each marker, AUCs and the number of diseased and control animals used for the analyses are shown	ROC curves for SCr, BUN, and KIM-1 (normalized to urinary Creatinine) for different time-points of urine collection. The control group contains vehicle-dosed animals with a histopathology score of 0. The diseased group contains animals with a histopathology score of 1 or higher. AUCs for each marker and the number of diseased and control animals used for the analyses are shown.
	

b. **Summary of Studies Conducted and Biomarkers Measured**

The studies conducted at each site are summarized in Table 3 below. Additional details for each study are provided in the PSTC summary tables found in Appendix 6.b.ii. Although most of the studies were dose-response studies conducted for 7 to 21 days, treatment duration ranged from 3 – 22 days. Time course studies were conducted at Merck (gentamicin and carbapenem A) and FDA (gentamicin, mercuric chloride, chromium and cisplatin). Most of the non-nephrotoxicant studies were conducted at Merck. Only the control biomarkers of BUN and sCr and the novel biomarker KIM-1 were measured at all three sites. The PSTC's tables below (Tables 4 and 5) show the biomarkers measured using the different nephrotoxicants and non-nephrotoxicants.

Table 3: Reviewer's summary of studies conducted

	Novartis	Merck	FDA
Rat strain	Han Wistar	Sprague Dawley, except for two studies	Sprague Dawley
Sex	Male	Male (Only one study (carbapenem-TC) with males and females)	Male
Animal number/group	6	4-6	3-6
Number of nephrotoxicants	8	11	4
Common nephrotoxicants	cisplatin gentamicin	cisplatin gentamicin	cisplatin gentamicin
Number of non-nephrotoxicants	2	9	0
Biomarkers used	BUN, sCr, KIM-1, clusterin, total protein, cystatin, β2-microglobulin	BUN, sCr, KIM-1, albumin, TFF-3	BUN, sCr, KIM-1

Table 4: PSTC Summary tables - Biomarkers measured for each compound studied [(N) Novartis, (M) Merck, (F) FDA] :

	BUN	S Cr	Kim-1	Clusterin	Albumin	TFF3	Tot Prot	Cyst C	Beta2Mic
23 Nephrotox Studies									
Gentamicin (N)	X	X	X	X			X	X	X
Gentamicin (M)	X	X	X		X	X			
Gentamicin (F)	X	X	X						
Cisplatin (N)	X	X	X	X			X	X	X
Cisplatin (M)	X	X	X		X	X			
Cisplatin (F)	X	X	X						
Doxorubicin (N)	X	X	X	X			X	X	X
Doxorubicin (M)	X	X			X				
Vancomycin (N)	X	X	X	X			X	X	X
Furosemide (N)	X	X	X	X			X	X	X
LiCO3 (N)	X	X	X	X			X	X	X
Puromycin (N)	X	X	X	X			X	X	X
Tacrolimus (N)	X	X	X	X			X	X	X
Carbapanem Antibiot. (M)	X	X			X	X			
Cyclosporin (M)	X	X	X		X	X			
Thioacetamide (M)	X	X	X		X	X			
Hexachlorobutadiene (M)	X	X			X				
Allopurinol (M)	X	X			X				
NPAA (M)	X	X			X				
D-Serine (M)	X	X			X				
Propylenimine (M)	X	X			X				
Mercuric Cl (F)	X	X	X		X				
Potassium Chromate (F)	X	X	X		X				

Note: No data were submitted for albumin from the FDA

11 Studies with Non-Nephrotoxics									
ANIT (N)	X	X	X	X			X	X	X
Methapyrilene (N)	X	X	X	X			X	X	X
Isoproterenol (M)	X	X			X	X			
Furan (M)	X	X			X				
Genipin (M)	X	X			X	X			
Cerivastatin (M)	X	X			X	X			
CCl4 (M)	X	X			X				
BrCCl3 (M)	X	X			X				
Water Diuresis (M)	X	X			X	X			
2% NaCl Diuresis (M)	X	X			X	X			
4% Sucrose Diur. (M)	X	X			X	X			

Table 5: PSTC summary of mode of toxicity of nephrotoxicity

Compound	Tubul.	Glom.	Coll. D.	Papillary	Mode of Toxicity
Allopurinol	x				Pyrimidine metabolism abnormality
Carbapanem Antibiotic	x				Reduced pyruvate-stimulated gluconeogenesis, PAH accumulation
Cisplatin	x	(x)	(x)		Direct DNA alkylation of DNA, Ox. stress
Cyclosporin	Tubular, Thick asc.limb				Signal transduction inhibitor, Complex (vasoconstrict., calcification...)
D-Serine	S3				Reactive oxygen from peroxisomal D-amino acid oxidase reaction
Doxorubicin/ Adriamycin	x (2nd)	x			Oxidative stress to glomerular filtration membrane
Furosemide	x				Mineralization
Gentamycin	x				Lysosomal phospholipidosis
Hexachlorobutadiene	S3				Metabolic activation leading to reactive thioketene, Impairment of calcium homeostasis ,
Lithium carbonate	x	(x)	(x)		Influences formation of intracellular cyclic adenosine monophosphate
Mecuric Chloride	x				Induces apoptosis via cytochrome c release from the mitochondria.
NPAA, N-phenylanthranylic acid				x	Acid mucopolysaccharides accumulation
Potassium Chromate	S1-S2				Vacuolization, diffuse necrosis
Propylenimine				x	Alkylating agent
Puromycin	x (2nd)	x			Damage to podocytes
Tacrolimus	x	(x)			Signal transduction inhibitor, Complex (vasoconstrict., calcification...)
Thioacetamide	x				Reduced urinary excretion of tricarboxylic acid cycle intermediates
Vancomycin	x	x			Oxidative stress (free radicals)

c. Histopathology Lexicon and Scoring

The PSTC agreed upon standardized vocabulary of terminology and grading for evaluating renal injury by histopathology. The lexicon used the primary histopathology processes in Table 6 below. The full lexicon in Appendix 6.b.i also lists secondary histopathology lesions and structural elements. Some disagreement among pathologists existed concerning the scoring of background lesions. The severity of the lesions was evaluated in a five-grade system at Merck, Novartis and the FDA, with the exception of the FDA studies with gentamicin, Hg, and Cr, which were not re-scored in time for the VXDS submission deadline. For those studies, scoring was provided for renal tubular injury only, which were the primary treatment-induced lesions found in the histopathological evaluations.

Table 6: Kidney histopathology lexicon and grading	
PSTC (Merck and Novartis) lexicon	**FDA**
Primary histopathology process **Tubular Cell Degeneration/Necrosis/Apoptosis** **Tubular Cell Regeneration** **Tubular Cell Alterations** **Tubular Dilatation** **Necrosis/Infarction** **Glomerular Alteration** **Pelvis Dilatation** **Intratubular Casts** **Inflammation** **Fibrosis** **Edema** **Vascular alteration** **Mineralisation-parenchymal** **Urothelial hypertrophy-hyperplasia**	Histopathology of proximal convoluted tubules showing necrosis, degeneration, regeneration, tubular dilatation, protein casts, and interstitial lymphocytic infiltration
Grading PSTC (scale 0 to 5) 0 = no abnormality noted 1 = minimal/very few/very small 2 = slight / few / small 3 = moderate / moderate number / moderate size 4 = marked / many / large 5 = massive / extensive number / extensive size	**Grading FDA (scale of 0 to 5)** 0 = normal histology, 1 = degeneration only without necrosis, 2 = <25% 3 = > 25% , but <50% 4 = > 50% , but <75% 5 = >75%

d. Summary Tables of ROC Curves

PSTC summary tables for the ROC curves for all histopathology grades of the Merck, Novartis and FDA data (shown in Section 3a) are provided in Tables 7, 8 and 9 below. These tables provide additional information on the ROC analyses using the inclusion model. Additional summary tables for the exclusion model are provided in Appendix 6.b.iii.

Table 7: PSTC Summary Table of ROC Curves - Inclusion model – Merck VXDS02

AUC: area under the curve, se: standard error, fold.cutoff: fold change, FPR: specificity(~0.05 error), TPR: sensitivity, npos: number positive samples by histopathology, nneg: number negative samples by histopathology. Maximum composite score was used for histopathology

All Histopathology Grades					1-SPEC	SENS		
Marker	AUC	SE	log2.cutoff	fold.cutoff	FPR	TPR	npos	nneg
Kim-1.ucr	0.97	0.015	3.07	8.38	0.050	0.83	77	101
SCr	0.87	0.029	0.25	1.19	0.050	0.70	77	101
BUN	0.84	0.031	0.62	1.53	0.040	0.57	77	101
All Histopathology Grades					**1-SPEC**	**SENS**		
Marker	AUC	SE	log2.cutoff	fold.cutoff	FPR	TPR	npos	nneg
Alb.Combined.ucr	0.84	0.017	2.94	7.66	0.049	0.53	251	449
SCr	0.74	0.020	0.32	1.25	0.049	0.47	251	449
BUN	0.75	0.020	0.72	1.65	0.049	0.43	251	449
All Histopathology Grades					**1-SPEC**	**SENS**		
Marker	AUC	SE	log2.cutoff	fold.cutoff	FPR	TPR	npos	nneg
nTFF3.ucr	0.83	0.024	3.01	8.03	0.046	0.30	135	175
SCr	0.88	0.020	0.26	1.20	0.046	0.70	135	175
BUN	0.89	0.020	0.54	1.45	0.046	0.66	135	175

Table 8: PSTC summary table of ROC curves – Inclusion model – Novartis VXDS02

Tubular pathology – All histopathology grades

	Pathology	AUC	Thr	Spec %	Sens %	Direct	N° Ctrls	N° Dis
Kim-1	Necr Apopt Cell Sloughing S1-S3 + non-loc.	0.89 ± 0.02	2.50	95	72	+	595	135
Clusterin	Necr Apopt Cell Sloughing S1-S3 + non-loc.	0.85 ± 0.02	2.30	95	63	+	604	135
Creatinine	Necr Apopt Cell Sloughing S1-S3 + non-loc.	0.72 ± 0.03	1.29	95	39	+	604	135
BUN	Necr Apopt Cell Sloughing S1-S3 + non-loc.	0.75 ± 0.03	1.17	95	42	+	604	135

The areas under curve (AUC) with standard errors, the thresholds and sensitivities for 95% specificity, the number of control animals and the number of injured animals are listed.

Glomerular pathology – All histopathology grades

Target	Biomarker	AUC	Thr	Spec %	Sens %	Direct	N° Ctrls	N° Dis
Specificity	Cystatin-C	0.90 ± 0.03	5.81	99	56	+	698	41
	B2-Microglobulin	0.89 ± 0.03	6.04	99	61	+	698	41
	Total Urinary Protein	0.86 ± 0.04	3.15	99	71	+	698	41
	BUN	0.74 ± 0.05	2.03	99	24	+	698	41
	Creatinine	0.55 ± 0.05	0.85	99	17	-	698	41
Sensitivity	Cystatin-C	0.90 ± 0.03	1.60	85	85	+	698	41
	B2-Microglobulin	0.89 ± 0.03	1.44	82	85	+	698	41
	Total Urinary Protein	0.86 ± 0.04	0.87	47	85	+	698	41
	BUN	0.74 ± 0.05	0.92	41	85	+	698	41
	Creatinine	0.55 ± 0.05	1.40	2	85	-	698	41

The areas under curve (AUC) with standard errors, the thresholds and sensitivities for 99% specificity (top part) and the thresholds and specificities for minimum 85% sensitivity (bottom part), the number of control animals and the number of injured

Table 9: PSTC Summary table of ROC analysis – Inclusion – FDA (VXDS02)

Results of the ROC analyses for the biomarkers listed in Column 2. Column 3: AUCs with standard errors; Column 4: the threshold at a minimum of 95% specificity; Column 5: actual specificity; Column 6: sensitivity. Columns 7: number of controls and column 8 numbers of diseased animals. Animals with a histopathology score of 0 or 1 were assigned to the control group, animals with histopathology scores of 2-5 were assigned to the diseased group.

All FDA data

Histopatohology	Biomarker	AUC	Thr	Spec %	Sens %	N° Ctrls	N° Dis
Controls: 0-1	BUN	0.78 ± 0.03	1.50	96	53	91	93
Diseased: 2-5	Creatinine	0.77 ± 0.03	1.64	97	43	91	94
	Kim-1	0.83 ± 0.03	8.86	96	60	92	92
	Kim-1 Normalized	0.92 ± 0.02	8.93	96	64	92	90
	Urinary Creatinine	0.84 ± 0.03	0.56	96	60	93	90

Cis-platin time course data

Biomarker	Time	AUC	Thr	Spec %	Sens %	N° Ctrls	N° Dis
BUN	72 h	0.70 ± 0.09	1.13	92	64	12	25
Creatinine	72 h	0.72 ± 0.09	1.60	92	48	12	25
Kim-1 Normalized	0 h	0.52 ± 0.10	0.59	92	12	12	25
Kim-1 Normalized	8 h	0.78 ± 0.08	0.73	92	52	11	25
Kim-1 Normalized	24 h	0.91 ± 0.05	1.39	92	88	12	25
Kim-1 Normalized	48 h	0.92 ± 0.04	1.43	92	88	12	25
Kim-1 Normalized	72 h	0.88 ± 0.06	1.32	92	84	12	25

e. Statistical Analysis

Tables 10 and 11 below summarize the statistical analysis performed by Merck and Novartis, respectively, to support the claim that a particular biomarker *outperforms BUN and sCr*. In the Merck studies, only KIM-1

and albumin outperformed BUN and sCr when a maximum composite histopathology score was used. In the Norvartis studies, clusterin and KIM-1 outperformed BUN and sCr for proximal tubular damage while cystatin C and β2-microglobulin outperformed BUN and sCr for glomerular damage. According to the Novartis results, total protein outperformed sCr, but not BUN for glomerular damage. Similar statistical analyses were performed using the exclusion model (see Appendix 6.b.iv). Additional statistical analyses show that all seven biomarkers *added value to BUN and sCr* (see Appendix 6.b.v). Statistical analyses using the FDA data were not performed.

Table 10: PSTC Statistical analysis – Inclusion – Merck (VXDS04)

Comparison: biomarker compared to control (BUN or sCr), npos: number of samples with positive histomorphologic change (Maximum Composite), nneg: number of samples with negative histomóphologic change (Maximum Composite), biomarker AVC: AUC from ROC curve for putative biomarker, CTL AVC: AUC from ROC curve for SCr or BUN, Diff AVC: biomarker AUC - CTL AUC, SE: standard error ofDiff AUC from DeLong analysis, pvalue: p-value from DeLong test, p.Holm: adjusted p-value using Holm procedure for multiplicity in testing TFF3 normalized or ng/mL three ways, q.BH: Benjamini and Hochberg false discovery rate for multiplicity in testing TFF3 normalized or ng/mL three ways. Note that p value fonts at p<.05 are bolded and underlined.

Comparison	npos	nneg	marker AUC	CTL AUC	Diff AUC	SE	pvalue	p.Holm	q.BH
Kim-1.ucr-SCr	77	101	0.967	0.870	0.096	0.028	**0.00068**		
Kim-1.ucr-BUN	77	101	0.967	0.843	0.124	0.031	**0.00008**		
Alb.Combined.ucr-SCr	251	449	0.843	0.742	0.101	0.020	**8.91E-07**		
Alb.Combined.ucr-BUN	251	449	0.843	0.752	0.091	0.021	**0.00001**		
nTFF3.ucr-SCr	135	175	0.828	0.885	-0.056	0.025	**0.02334**	0.07001	0.07001
nTFF3.ucr-BUN	135	175	0.828	0.885	-0.057	0.031	0.06296	0.18887	0.18887
nTFF3.ex-SCr	111	176	0.849	0.887	-0.038	0.023	0.10663	0.21327	0.15995
nTFF3.ex-BUN	111	176	0.849	0.872	-0.023	0.033	0.47583	0.95166	0.68919
nTFF3.conc-SCr	136	187	0.887	0.888	-0.001	0.020	0.96336	0.96336	0.96336
nTFF3.conc-BUN	136	187	0.887	0.876	0.011	0.027	0.68919	0.95166	0.68919

Table 11: PSTC Statistical analysis – Inclusion – Novartis (VXDS02)

Results for statistically comparing the significance of differences of AUCs of the ROC analyses between markers and BUN or sCr for the inclusion analysis. Column 1 represents the marker and standard being compared; Column 2 the number of diseased samples, Column 3 the number of control samples, Columns 4 and 5 the AUCs for the markers and the standards, Column 6 the difference of AUCs, Column 7 the standard error of Diff AUC from DeLong analysis, Column 8 the p-value from DeLong test, Column 9 the adjusted p-value using Holm procedure for multiplicity in testing the glomerular markers also for tubular damage, and Column 10 the Benjamini and Hochberg false discovery rate for multiplicity in testing glomerular markers also for tubular damage. Note: p<0.05 are in bold font.

Inclusion Model: Tubular Damage S1-S3

Comparison	npos	nneg	marker AUC	CTL AUC	Diff AUC	SE	pvalue	p.Holm	q.BH
Clusterin-SCr	135	604	0.852	0.721	0.131	0.036	**2.30E-04**		
Clusterin-BUN	135	604	0.852	0.752	0.100	0.033	**2.18E-03**		
KIM1-SCr	135	595	0.888	0.721	0.168	0.034	**6.42E-07**		
KIM1-BUN	135	595	0.888	0.751	0.138	0.032	**1.43E-05**		

Inclusion Model: Glomerular Alterations / Damage

Comparison	npos	nneg	marker AUC	CTL AUC	Diff AUC	SE	pvalue	p.Holm	q.BH
Cystatin-SCr	41	698	0.904	0.550	0.354	0.073	**1.31E-06**	**2.63E-06**	**2.63E-06**
Cystatin-BUN	41	698	0.904	0.740	0.164	0.051	**1.30E-03**	**2.61E-03**	**2.61E-03**
B2Microglobulin-SCr	41	698	0.890	0.550	0.340	0.076	**6.78E-06**	**1.36E-05**	**1.36E-05**
B2Microglobulin-BUN	41	698	0.890	0.740	0.150	0.052	**3.99E-03**	**7.99E-03**	**7.99E-03**
Total Protein-SCr	41	698	0.858	0.550	0.308	0.071	**1.30E-05**	**2.60E-05**	**2.60E-05**
Total Protein-BUN	41	698	0.858	0.740	0.118	0.066	7.19E-02	1.44E-01	1.44E-01

f. **Individual Animal Data**

Individual animal data were provided separately in Excel format by Merck, Novartis and FDA. These documents tabulate histopathology, clinical chemistry, urinalysis, and biomarker data for individual animals by study. These documents are available electronically (restricted access) at \\Vgds.nctr.fda.gov\data\VXDS\VXDS15 PSTC June 2007.

Biomarker values for individual animals are shown in the PSTC figures below. Figure 5 shows individual animal values for BUN, sCr, KIM-1, and albumin relative to either control or diseased pathology in a Merck study of gentamicin. Individual animals in multiple studies are shown for the Novartis and FDA data in Figures 6 and 7, respectively. These figures group the animals by study and within each study by dose-level and time-point. The animals are colored by histopathology grade and are labeled by termination time-point. Figures showing biomarker values for individual animals in the multiple Merck studies were not provided.

Figure 5: PSTC Figures showing biomarker values for individual animals – Merck gentamicin study

Gentamicin treatment was perform at 0, 20, 80, and 240 mg/kg/day, and necropsies were performed on Day, 9, and 15 (12 for 240 mg/kg/day). Collected blood samples were analyzed by serum chemistry (Cr and BUN) and urine samples were tested for KIM-1 and albumin levels which are expressed as excreted (E) values that were normalized to total urine volume over the collection period. Excretion and UCr normalization are inversely correlated to a high degree. These data are discussed in the main text as UCr normalized fold-change. Note that the excretion fold change and UCr fold change are essentially the same and have been compared throughout.

The marker and corresponding excretion units are indicated above each plot. The x-axis indicates the Day_Dose. Note that the animals in the 240 mg/kg/day group were necropsied on Day 12, not Day 15, because of morbidity. Open squares indicate grade 0 pathology and yellow indicates pathology grade greater or equal to 1. Black circles indicate the mean for each group.

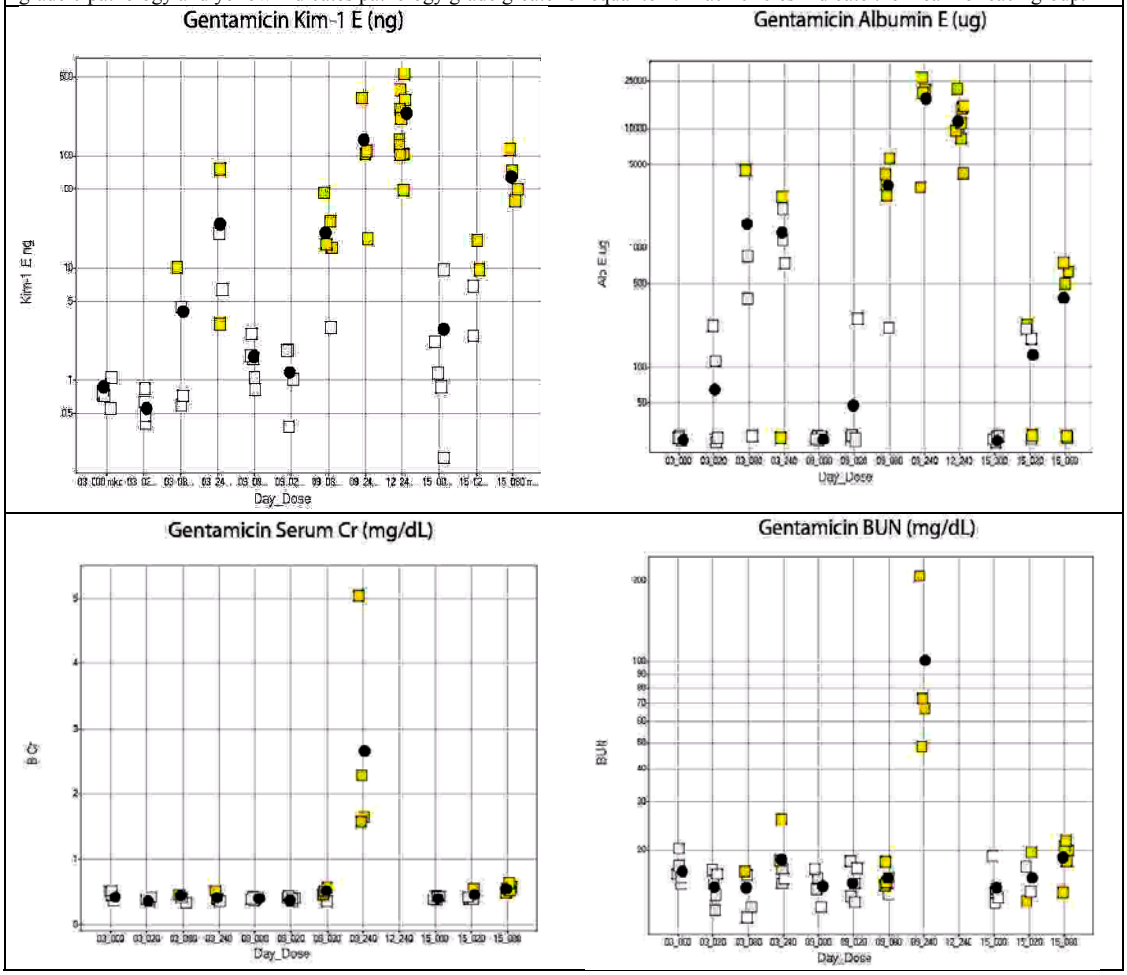


Figure 6: PSTC figures showing biomarker values for individual animals – multiple Novartis studies

Urinary concentration levels of KIM-1, clusterin, sCr and BUN (fold changes versus control animals) for the animals used in the ROC analysis relative to histopathology grade of tubular damage. The animals are ordered by study and within each study by dose-level and time-point. The animals are colored by histopathology grade and are labeled by termination time-point. The magenta line is the decision threshold for 95% specificity. The y-axis scales are logarithmic.

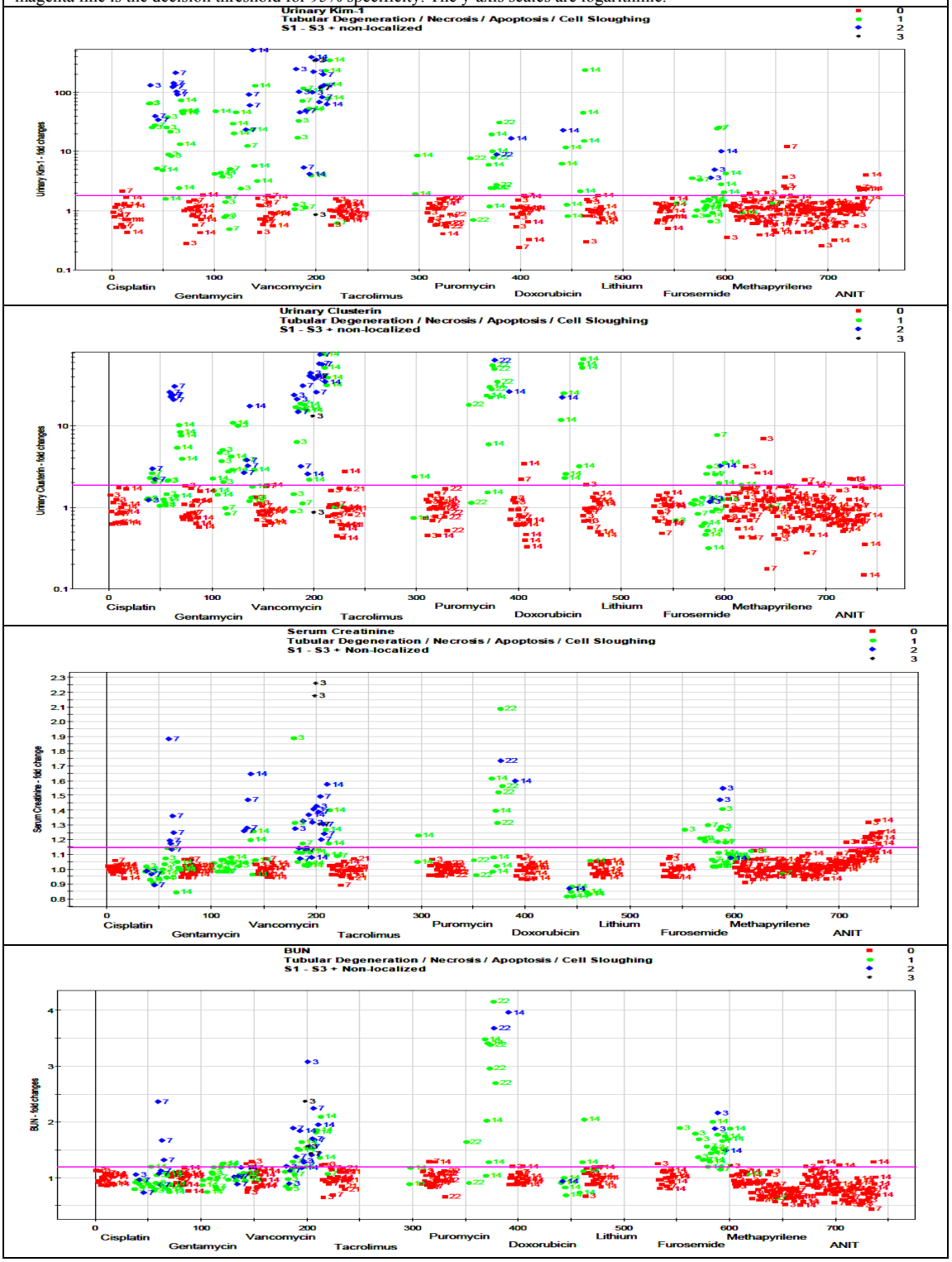
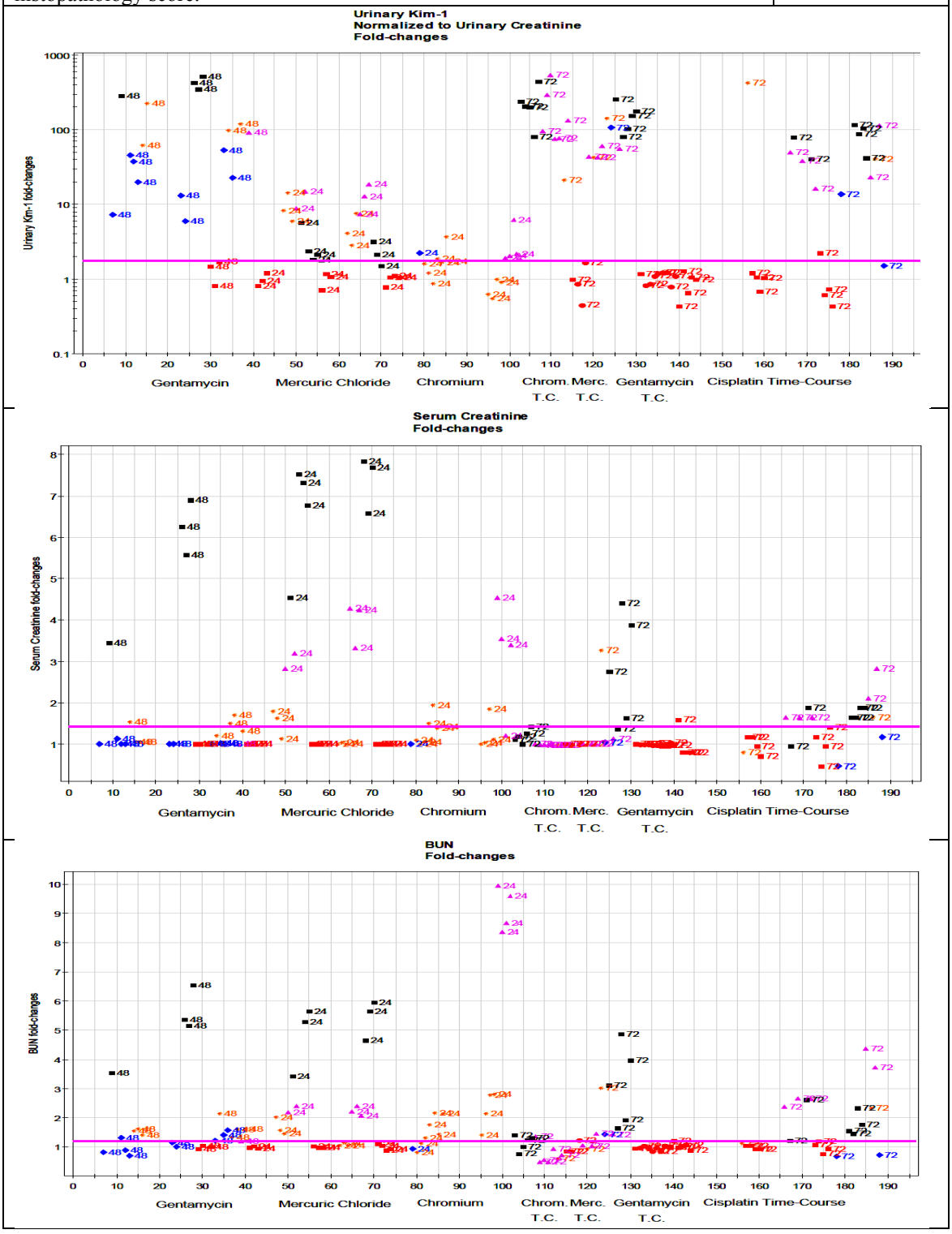


Figure 7: PSTC Figures showing biomarker values for individual animals – multiple FDA studies

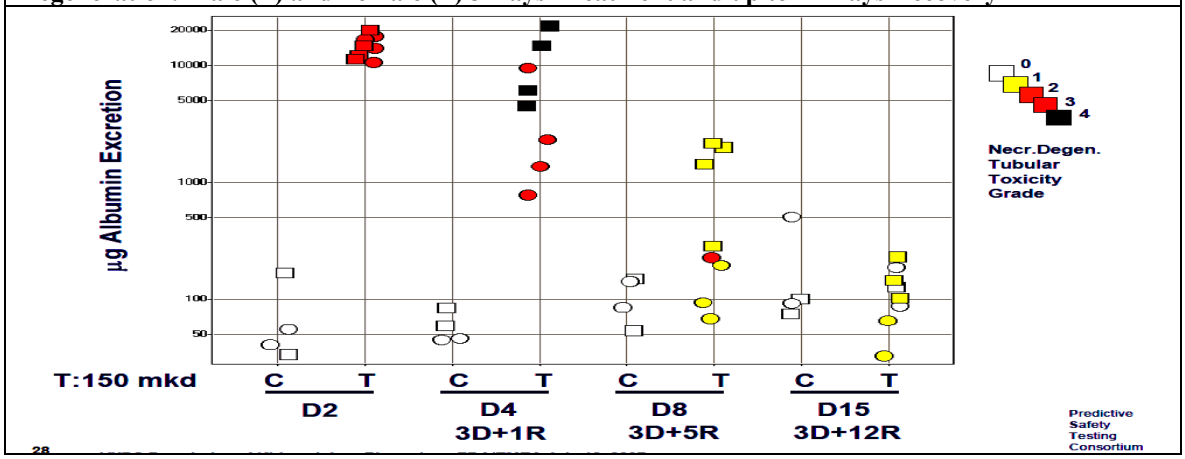
Biomarker levels of all samples expressed as fold-changes (compared to the average of control animals). The values for KIM-1 were normalized for urinary creatinine. The samples are ordered by study, labeled with termination time-point and colored by histopathology score.



g. Recovery Studies

Only two recovery studies were conducted. Figure 8 below shows a decrease in urinary albumin during the recovery phase after treatment with carbapenem A. The decreased urinary albumin values are associated with decreased histopathology scores in this Merck study. However, minimal histopathology was observed on Day 15 in five treated animals whose albumin values were similar to those in control animals. The PSTC did not provide similar plots/data for the other biomarkers.

Figure 8: PSTC Figure - Urinary Albumin (µg) Monitored Carbapenem A Induced Necrosis and Degeneration. Male (□) and Female (○) 3 Days Treatment and up to 12 Days Recovery



h. ROC analysis of different histopathology lesions

Novartis performed ROC analyses for all individual biomarkers by site of injury (as determined by histopathology), although it was not clear whether inclusion or exclusion analyses were performed. Merck did not provide similar analyses. Novartis provided for individual and compiled pathologies the corresponding AUC for each biomarker in a matrix table. Example tubular and glomerular pathologies along with a compilation of pathologies are shown in Table 12 below, while the complete Novartis results are shown in the Appendix 6.b.viii. For glomerular alteration or damage, the AUC values decreased in order from cystatin C, β2-microglobulin and total protein, and clusterin, to BUN, KIM-1 and sCr. For any tubular alteration, the AUC values for KIM-1 and clusterin were only slightly higher than AUC values for BUN and sCr, whereas the AUC values for cystatin C, β2-microglobulin and total protein were less than AUC values for BUN and sCr. For the specific pathologies of tubular necrosis and tubular degeneration/necrosis/apoptosis, the difference in AUC values was greater between KIM-1 or clusterin and BUN or sCr. No statistical analysis was provided for these comparisons.

Table 12: Novartis ROC analysis for different pathologies and biomarkers

Primary Histopathology Process	Secondary Lesion	Structural Element / Segment	Serum Creatinine	BUN	Urinary Creatinine	Urinary Protein	Ur. B2-Microglobulin	Urinary Cystatin C	Urinary Clusterin	Urinary Kim-1	N° Controls	N° Diseased
Example pathologies												
Tubular Cell Degeneration/Necrosis/Apoptosis	Necrosis	No precise localization possible	0.92	0.91	0.50	0.65	0.71	0.56	0.93	0.93	291	36
		Prox. convoluted tubule (PCT, s1-s2)	0.78	0.75	0.50	0.73	0.71	0.81	0.90	0.95	292	51
		Thick descending tubule (s3)	0.72	0.81	0.50	0.58	0.51	0.61	0.86	0.91	292	100
		Thick ascending tubule	0.74	0.83	0.50	0.51	0.53	0.67	0.96	0.93	292	41
		Distal convoluted tubule	0.55	0.61	0.50	0.88	0.99	0.99	0.83	0.70	292	2
		Collecting duct	0.98	0.85	0.50	0.70	0.68	0.54	0.93	0.88	292	18
	Apoptosis	No precise localization possible	0.52	0.59	0.50	0.67	0.65	0.74	0.85	0.71	290	6
		Prox. convoluted tubule (PCT, s1-s2)	0.79	0.67	0.50	0.71	0.84	0.85	0.91	0.86	291	17
		Thick descending tubule (s3)	0.92	0.95	0.50	0.81	0.56	0.60	0.98	0.88	291	5
		Thick ascending tubule	0.89	1.00	0.50	0.99	0.64	0.62	1.00	1.00	292	1
		Collecting duct	0.81	0.60	0.50	0.55	0.52	0.72	0.59	0.57	292	11
Glomerular Alteration	Mesangial proliferation/expan	Glomerulus	0.53	0.80	0.50	0.86	0.89	0.91	0.83	0.76	292	40
	Glomerular Vacuolation	Glomerulus	0.82	0.96	0.50	0.83	0.95	0.95	0.83	0.74	292	14
Overall compilation												
Any kidney lesion			0.64	0.64	0.50	0.55	0.51	0.56	0.65	0.67	209	373
Any tubular lesion			0.65	0.65	0.50	0.56	0.51	0.58	0.70	0.70	234	313
Glomerular Alteration / Damage (Mesang. prolif., Interstit. Bowmans Capsule fibrosis, Glom. vacuolation)			0.52	0.80	0.50	0.86	0.89	0.91	0.83	0.75	291	41
Tubular Necrosis, S1-S2 + nl			0.83	0.81	0.50	0.60	0.55	0.73	0.93	0.95	291	78
Tubular Degeneration/Necrosis/Apoptosis/Cell Sloughing/Basophilia/Mitosis/Dilatation, S1-S3 + nl			0.67	0.67	0.50	0.55	0.50	0.59	0.74	0.76	246	249
Tubular Degeneration/Necrosis/Apoptosis/Cell Sloughing, S1-S2 + nl			0.81	0.80	0.50	0.61	0.58	0.74	0.93	0.93	289	90
Tubular Degeneration/Necrosis/Apoptosis/Cell Sloughing, S1-S3 + nl			0.73	0.78	0.50	0.59	0.53	0.66	0.87	0.90	289	135

ROC analyses for different pathological lesions and biomarkers. The area under curves (AUCs) for the different markers are color-coded (orange represents AUCs>0.9, bright yellow AUCs>0.8 and light yellow AUCs>0.7). The number of controls used for the corresponding analysis and the number of animals with the corresponding pathology are shown. Pathologies with less than 20 cases are colored gray.

4. Reviewer Discussion of Qualification Data

a. The Qualification Process

Data submitted by the PSTC for biomarker qualification were received by the BQRT from June through November 2007 and included data contained in the original PSTC submission as well as additional data requested by the BQRT to bridge information gaps identified during the review process. These submissions and meetings with the BQRT are summarized in Table 13 below.

Table 13: Summary of PSTC submissions and meetings with the BQRT

Date	Description
06_15_07	Initial submission containing: Consolidated PSTC summary report containing overall conclusions, proposed next steps, and a list of focused questions for discussion Merck summary report providing data on three urinary protein biomarkers (KIM-1, albumin, and trefoil factor-3) Novartis summary report providing data on five urinary protein biomarkers (KIM-1, clusterin, cystatin C, β2-microglobulin, and total protein) FDA summary report providing additional corroborating data for urinary KIM-1 PSTC Charter describing aspects of the consortium Primary literature references used to support key claims
07_12_07	PSTC Meeting minutes of VXDS meeting with FDA/EMEA/PMDA along with FDA/EMEA Preliminary review comments and questions FDA Preliminary statistical review comments PSTC presentations from July 12 meeting

07_27_07	Slides for 07_27_07 VXDS teleconference and email summarizing status of open items at the end of the meeting.
08_20_07	VXDS01 - Supplementary submission #1: Toxicology study reports and/or histopathology and clinical chemistry data Individual animal data for histopathology and all analytes in Excel format Assay analytical validation reports
09_10_07	VXDS02 - Supplementary submission #2: All contents of original VXDS and VXDS01 PSTC slides and meeting minutes from July 12, July 27, August 14 meetings between PSTC and FDA/EMEA) Requested interference data (hemoglobin, bilirubin, metals, high salt) Requested statistical analyses Histopathology slide images BQRT response documents (UCr normalization, further prodromal clarification, Han Wistar/Sprague Dawley comparability from KIM-1 studies) Comprehensive summary addenda to supplement original summary
09_20_07	VXDS03 - Supplementary submission #3: Introductory summary and expert clinical reviews for five renal biomarkers (KIM-1, albumin, total protein, cystatin C, and β-2 microglobulin). Reports detailing the results of the selected blinded histopathology slide re-reads from both Merck and Novartis FDA KIM-1 Dataset Supplement complementing the 06_15_07 PSTC VXDS submission and the 07_12_07 PSTC VXDS presentation Merck diuresis research study protocol to supplement histopathology and clinical chemistry tabular results submitted within supplementary VXDS 01 Supporting primary literature references
10_01_07	VXDS04 - Supplementary submission #4: Corrections to Merck Overall summary document and relevant appendices Merck Datasheet: Histopathology, clinical chemistry, electrolytes and protein biomarkers on single-animal basis for all studies Merck Appendix II (Histopathology Master) Excel Sheets
10_03_07	VXDS05 - Supplementary submission #5: Introductory summary describing the rationale for the revised claims Revised claims for the proposed urinary biomarkers of acute drug-induced kidney injury Flowchart describing application of the biomarkers to bridge to clinical studies Four historical case study examples that describe application of biomarkers
10_08_07	VXDS06 – Supplementary submission #6 Additional ROC analyses evaluating the performance of the new renal safety biomarkers, SCr, and BUN with decreasing histopathology grades.
10_09_07	PSTC Meeting minutes of VXDS meeting with FDA/EMEA/PMDA along with Presentation on the status of open items Presentation on ROC curves using a subset of Merck data Analysis of ROC curves for Merck data based on histopathology grade Analysis of ROC curves for Novartis data based on histopathology grade FDA presentation concerning clinical development of biomarkers
10_15_07	Webex with Merck to Discuss Analysis Discrepancies – Email summary of discussion, Summary and Excel file of BRQT analysis
11_08_07	VXDS07 - Supplementary submission #7: Merck ROC summary statistics for a defined maximum composite injury score excluding interstitial inflammation
11_26_07	VXDS08 – Supplementary submission #8 Guidance on when new biomarkers should be applied and what preclinical testing should be performed Guidance on application of new biomarkers in early clinical trials Tabulated summaries of published clinical data

b. **Analytical Validation**

The purpose of bioanalytical method validation is to demonstrate that a particular method used for quantitative measurement of an analyte (in this case a biomarker) is reliable and reproducible for its intended use. The Bioanalytical Method Validation Guidance for Industry (www.fda.gov, May 2001) recommendations are not for new in vitro diagnostic tests, but instead apply to the methods to be used by a sponsor during drug development. Hence, this guidance seems to be an appropriate reference for biomarker development. According to the guidance, full validation is important when developing and implementing a bioanalytical method for the first time, for a new drug entity or when metabolites are added to an existing assay for quantification. A partial validation is a modification of already validated bioanalytical methods and is appropriate for bioanalytical method transfers between two laboratories, a change in the biological matrix of interest (e.g. rat plasma to rat urine), etc.

According to the Bioanalytical Method Validation Guidance, the key parameters for bioanalytical method validation are: accuracy, precision, selectivity, sensitivity, reproducibility and stability. Measurements of the biomarkers in the biological matrices should be validated and the stability of the biomarkers in spiked samples determined. The chemical identity and purity of the reference standard used to spike samples and to generate quality control samples is critical since validation data can be affected. With respect to the accuracy of the assay, the Bioanalytical Method Validation Guidance recommends that the mean value of replicate analyses of samples should be within 15% of the actual value except at the lower limit of quantification, where it should not deviate by more than 20%. The precision at each concentration level should not exceed 15% of the coefficient of variation (CV) except for the lower limit of quantification, where it should not exceed 20% of the CV.

In addition to the approach described in the Bioanalytical Method Validation Guidance, a “fit-for-purpose” approach has been proposed for biomarker method development and validation (Lee et al 2006). According to this approach, the regulatory requirements for method validation at a particular time during development of a biomarker should be based on the intended use of the biomarker. The proposed category of “exploratory method validation” has less rigorous criteria for validation than the criteria that would need to be met for a well-qualified biomarker. Lee et al (2006) indicate that a biomarker with only exploratory method validation would be suitable for use in an early phase clinical trial where it would be employed for less critical decision-making. As the biomarkers discussed in this review are to be used to aid in critical decision-making (e.g. to help determine the

safety of a drug at a given dose), it is the opinion of the BQRT that for biomarkers submitted for FDA qualification, the method validation for the biomarker should in general be at the advanced method stage.

Compared to the FDA guidance, Lee et al (2005) propose a less strict acceptance criteria for the imprecision and accuracy of a biomarker assay (in the range of 25-30% as compared to the usual limit of 15%). These more lenient criteria depend on a case-by-case basis on the analyte properties, type of assay, assay limitations, and intended use of the data. The BQRT may consider the qualification of biomarkers whose method validation yields values of greater than 15% for imprecision on a case-by-case basis. The decision will be based on the relationship of the imprecision to the proposed ranges for normal and "positive" values and how this imprecision will affect the sensitivity and specificity of the assay. An imprecision value of greater than 15% may be acceptable for a biomarker assay if there is a sufficiently large fold difference between normal and "positive" biomarker values and minimal to moderate biological variability of the normal population. However, such imprecision may not be acceptable when the difference between normal and "positive" values is small.

Immunoassays

In addition to the discussion above, unique issues are raised by the use of immunoassays to measure biomarker levels. Immunoassays carry other considerations such as cross-reactivity, non-specific binding and interfering endogenous substances. A comparator method such as LC-MS is useful for checking the consistency of an immunoassay. Fluorescent microsphere assays, e.g. the Luminex xMAP® platform used for several of the analytes, also carry unique considerations. Some of these considerations are:

1. Serum antibodies binding directly to the beads (Waterboer 2006)
2. Stochastic variance, inter-instrument calibration (Hanley 2007a)
3. Variation in size of microspheres affecting brightness (Hanley 2007b)
4. Carryover between wells (Hanley 2007c) of 2 types:
   a. small predictable, declines consistently
   b. potentially large, unpredictable, does not decline
5. Effect of binding site density on apparent affinity of interaction (Iannone 2006)

Merck Reports

The Merck Summary Reports indicate the rat KIM-1 assays were performed at the Brigham and Women's Hospital. These Summary Reports described some of the methods used, but did not include the

validation of the KIM-1 assay. The Merck Appendix IV, provided by Drs Vaidya and Bonventre of the Brigham and Women's Hospital, included a summary of the Luminex xMap assay development but not a full validation. A publication (Vaidya 2006) provided some assay validation data for the KIM-1 sandwich ELISA assay.

In the August 20 submission, Merck submitted an 18 page summary of the TFF3 and urinary albumin assays. Each of the individual study reports was checked for methods validation. A full description of validation was not found. The AssayMax rat albumin ELISA kit from AssayPro (a competitive ELISA) was used for detection of urinary albumin. The summary states that urine from Carbapenem A treated male rats was pooled, diluted 1:20 in diluent and used as the matrix for the validation studies. It is interesting that non-naïve animals were used to provide the matrix. The reasoning behind this was not apparent. The ELISA method was compared (cross-validated) with an immunoturbidimetric assay.

Novartis Reports

The assays for the urinary β2-microglobulin, cystatin C and clusterin were multiplexed sandwich and competitive immunoassays using the Luminex xMAP® platform. The KIM-1 assay was a sandwich immunoassay also on the xMAP® platform. Each analyte was measured using the Luminex 100 reader. The original submission provided only a summary of the methods validation. Validation reports were provided in a subsequent submission.

Urinary total protein was assayed by commercially available methods (Advia 1650). The other methods described were for multiplex immunoassays as noted above. Clusterin, cystatin C and osteopontin were multiplexed in one platform. β2-microglobulin, GSTα, GSTμ, NGAL, TIMP and VEGF were multiplexed into a second platform. Chemical analysis forms were not provided for the standards/calibrators.

A separate report was submitted for KIM-1. The KIM-1 capture antibody coupled to the fluorescently addressable microspheres and the biotinylated detection antibody, sample diluent buffer and standard were obtained from J. Bonventre, Harvard University. The rat plasma and urine samples were provided by Novartis. They were shipped to Rules-Based Medicine (RBM) on dry ice and the assays performed there. Analysis was performed in a Luminex 100 instrument and the resulting data stream was interpreted using proprietary software developed at RBM. The validation report submitted was generated at RBM. Chemical characterization and purity of the standards was not apparent. Validation of the proprietary software was not referenced.

With the exceptions already noted above, the validation of the assays followed accepted procedures and addressed the critical points that would be expected. The values that the sponsor assigned as acceptable were less rigorous than those recommended in the FDA guidance. However, values usually fell within more stringent ranges. A few points are worth further discussion.

Variability

The acceptance criterion for the newly developed assays was a deviation of ±20 % for the mean accuracy except at the LLOQ where the deviation was ±30%. Although this is less stringent than the guidance recommendations, the majority of data were within the more rigorous standard. One exception to this occurred in the assay for KIM-1. The sponsor's summary of LLOQ is shown below:

Table 5. Measurement of LLOQ for Urine KIM-1

	KIM-1						
	Rep 1	Rep2	Rep 3	Run Mean	Run CV	**Overall Mean**	**Overall CV**
Run No.	ng/ml					ng/mL	
501767	0.380	0.370	0.384	0.378	2%	0.355	**9%**
	0.204	0.214	0.188	0.202	6%	0.194	**7%**
	0.137	0.146	0.111	0.131	14%	0.113	**17%**
	0.082	0.079	0.085	0.082	4%	0.069	**19%**
	0.057	0.040	0.060	0.052	21%	0.036	**48%**
	LOW	0.032	0.036	0.034	8%	0.024	**78%**
	LOW	LOW	LOW				
501772	0.354	0.292	0.370	0.339	12%		
	0.185	0.167	0.206	0.186	11%		
	0.103	0.089	0.103	0.098	9%		
	0.071	0.059	0.047	0.059	20%		
	0.028	LOW	0.028	0.028	0%		
	LOW	LOW	LOW				
	LOW	LOW	LOW				
501775	0.372	0.350	0.320	0.347	8%		
	0.186	0.200	0.197	0.194	4%		
	0.124	0.115	0.092	0.110	15%		
	0.055	0.076	0.067	0.066	16%		
	0.033	0.010	LOW	0.021	77%		
	LOW	0.002	LOW	0.002			
	0.002	LOW	LOW				

The LLOQ for urine was interpolated based on the average of the imprecision that was obtained for 3 different experiments performed in triplicate over three different days and instruments. The LLOQ for urine was determined to be 0.058 ng/mL.

At the LLOQ defined above, the variability is somewhere between 19 and 48%. There were numerous samples reported to have values of urinary KIM-1 at the LLOQ. An interesting analysis would be to determine the correlation between histological findings and urinary KIM-1 at the LLOQ. Given that there was minimal histological data (1-2 sections per animal) in the current studies, such an analysis would not be particularly informative at this time.

It is expected that almost any assay will have its greatest variability in the low concentration ranges. The significance of the variability of the assay at the LLOQ depends upon the magnitude of the threshold of concern. That is, at what level of urinary KIM-1 do we say that there is a renal lesion of concern or the beginning of a lesion of concern? If the cutoff point for this level of concern is well above the LLOQ, the variability of the assay in the low concentration range becomes less significant.

Standards/Calibrators and Alternative Methodology

Comments that extend across the validation material submitted:

1. *Standards/calibrators used.* Commercial source and catalog numbers were provided. Purity and chemical characterization were not apparent in the reports.

2. *No methodology was listed as used to establish equivalence of methodology or to cross check.* Ideally, an alternate method should have been described or proposed for establishing accuracy of the tests.

Cross Reactants

The report for the clusterin, cystatin C, β2-microglobulin assays list 76 possible cross-reactants. The report for KIM-1 also listed these antigens. These potential cross-reactants were tested and the amount of cross-reactivity was determined. It is impossible to test every potential cross reactant. Since the reports did not define the epitopes of the antibodies employed in the assays, a systematic search of protein and gene sequence databases could not be performed.

Intra-site reproducibility of results

The calibration curves were reported to have been run with independently prepared reagents, run on different days using different instruments and operators. As the data are presented, it is apparent that different samples were processed on different days. Determination of which data were generated by different operators and with which lots of reagents is not possible from the materials as presented. The majority of samples for the analytes of interest fall within an acceptable range of variance.

Inter-site reproducibility of results

As noted in the Minutes of the July 12, 2007 meeting there was "remarkable similarity between the Novartis and Merck data for KIM-1." However, there was no comparison of inter-site reproducibility of the biomarker results. That is, for the KIM-1 assay, aliquots of the same samples were not analyzed at the two sites.

Matrix Interference

Matrix interference was examined by spiking hemoglobin, bilirubin and triglycerides into a mid-level control and determining the percent recovery (observed/expected). The results shown indicated that values were within ±6% of 100 % recovery for the clusterin, cystatin C and ±2% for the β2-microglobulin assays.

Triglycerides exert a greater effect over the KIM-1 assay.

Because of the question of variability at the low end of the concentration range, the sponsor may find it advantageous to repeat the triglyceride interference analysis with a low-level control for all of the analytes of interest.

Other matrix interference issues

The "Information Gaps in PSTC Qualification of Biomarkers of Nephrotoxicity" presented to the PSTC on July 2007 asked a specific question as to the effect of hemoglobin, high protein, bilirubin, high salt, metals (mercury, cadmium, lead, lithium, gadolinium,) on the assays. Hemoglobin and bilirubin were analyzed and reported as noted above. High salt per se was not directly addressed. However, samples were diluted to appropriate concentration and compared to further dilutions in sample diluent. This would address to some extent the effect of matrix (primarily urine) concentrations of salts. Urine samples diluted 1:10 and 1:80 showed essentially the same results for KIM-1. Plasma samples showed greater variability with dilution. Urinary analysis of clusterin showed variability with dilution, e.g diluted 1:25 (101% of expected recovery) versus 1:100 (78% of expected recovery) and 1:200 (56% of expected recovery). β2-microglobulin also showed some variability with urine dilution.

The September 10, 2007 submission included the results of interference tests conducted with cadmium, gadolinium, mercury and lead. The concentrations used in the tests corresponded to the maximum published concentrations observed in human urine. The interference tests were performed at RBM for the multiplex assays. Briefly, the sponsor spiked the urine samples with the heavy metals and compared the percent recovery of the desired biomarker to control (no added heavy metal). Calbindin and EGF were most affected by these substances. KIM-1 recovery was also disturbed by up to 15%. Because the acceptance criteria were set at a deviation of ±30%, the results were considered acceptable with the exception of calbindin and EGF.

Issues specific to multiplex and microbead assays

To increase the FDA's confidence in the dependability of the assays, the sponsor should optimize the dilution and size range of beads in the preparations used and assess well-carry over effects.

The analysis package states on page 8 of the RBM report states that the software used is available commercially, but not how it was validated and the reference to that information. It is unclear from the description if the same curve equation is used each time the assay is

used or if the same curve equation is used within a given assay. It is not clear that the same equation and analysis package used at Harvard was used at Mesoscale, and how this difference may have contributed to inter-laboratory variation.

c. **Correlation between Histopathology and Biomarker Data.**

i. Exclusion of animals from ROC analysis

The PSTC preferred drawing conclusions based on their "exclusion" analysis, in which nephrotoxicant treated animals with a histopathology score of zero were omitted from the analysis. The PSTC maintained that the exclusion analysis avoids incorrect conclusions in cases where the biomarker was positive and histopathology was not detected. The PSTC argued that these cases could represent either "prodromal" signals or situations where the histopathology of a single section of only one kidney failed to detect a potential focal lesion elsewhere in either of the two kidneys that could contribute to the appearance of biomarker in the urine from an overnight collection. The PSTC explained that the term 'prodromal' was used when a signal at a low dose/early timepoint (when no histopathology was observed) was confirmed by histopathology at higher doses/later timepoints. The discussion by the PSTC from the combined report submitted 09/05/07 is provided in the Appendix 6.b.vii.

However, no data established in a sufficient number of animals, evaluated with an adequate number of histopathology sections that positive values for a biomarker are predictive of subsequent histopathology. Therefore, the BQRT preferred to draw conclusions based on the PSTC "inclusion" analysis in which all animals were evaluated and biomarker positive animals were treated as false positives.

It can not be asserted than an increase in titer of a biomarker without a change in histopathology reflects occult toxic injury. This must be supported with a lesion at the electron-microscopic level, changes in immunostaining (protein loss), or a functional anomaly (e.g., erythropoietin production; renin-release; urine concentrating capacity, bicarbonate metabolism etc.).

ii. Blinding of histopathology analysis

Background

Histopathology was used as the "gold standard in the PSTC submission. Whether the biomarker and histopathology evaluations were assessed independently of each other (that is, by interpreters who were unaware of the results of the other investigation) is unclear. Since bias is unintentional and unconscious, knowledge of even the

dose group could potentially lead to detection bias (Ransohoff 2005), especially in the histopathology evaluation of less severe lesions. Such bias could result in falsely elevated estimates of biomarker accuracy.

Histopathology slides from pre-clinical safety assessment studies are usually evaluated with the knowledge of the treatment group of origin. The Standards of Practice document written by the Society for Toxicologic Pathology (Crissman et al 2004) is for the application of histopathology as an "integral component of safety/toxicology studies," which are reviewed by regulatory agencies. The recommended approaches allow the pathologist "to find important, and sometimes subtle, differences between the tissues of treated and untreated animals" through a constant comparison with the control sections and full knowledge of the study design and other study results. However, the document acknowledges that histopathology is a "descriptive and interpretive science", which has "an element of subjectivity." Although these recommended practices may be useful in detecting treatment effects important for safety determinations in drug development programs, the BQRT feels this approach does not provide a completely unbiased evaluation of histopathology for biomarker qualification.

In the context of evaluating a new biomarker rather than a drug, the BQRT feels that knowledge of the treatment assignment or other aspects of study design or results that could potentially unblind the pathologist and potentially lead to bias. Even if a pathologist is "blinded" to specific novel biomarker results, any additional data, such as comparator biomarker results, provided to the pathologist may impart clues that could influence, consciously or subconsciously, the evaluation of the slides. The BQRT thus holds the evaluation of the histopathology for biomarker qualification to a different standard than that used for safety determinations in drug development programs. For this reason, the BQRT believes that it is critical that blinded evaluations of both histopathology and biomarkers be used in prospective studies of these biomarkers.

Timeline of events

In 2007, the PSTC initial submissions indicated that all the biomarker analyses were conducted in a blinded manner, whereas the histopathology evaluation was not blinded at any site. According to these initial PSTC submissions and discussions with the BQRT, pathologists at both sites followed standard procedures for evaluation of tissue samples from toxicology studies. The pathologists first evaluated slides from the control and high dose groups prior to evaluation of the low and mid-dose groups. It also was not clear

whether the pathologists had knowledge of the novel or comparator biomarker results prior to the histopathology evaluation.

Since the pathologists had full knowledge of the study design and dose groups, the BQRT felt there was potential for bias in the evaluation of histopathology. To address this issue, the BQRT proposed a reciprocal exchange and evaluation of slides between Merck and Novartis. This would have addressed discrepancies in the histological lexicons used and operator bias in the pathologists' evaluations. In response, the PSTC arranged for an evaluation of a sub-set of the slides by an independent pathologist. SRI International (David Fairchild, D.V.M., D.A.C.V.P., San Rafael, CA, USA) conducted this independent reading of the slides for the cisplatin and gentamicin studies (115 slides from Merck cisplatin TT04-2530 and gentamicin TT04-2530 and 141 slides from Novartis cisplatin 29524 and gentamicin 29755 studies).

Since the companies used different rat strains, the PSTC insisted that some control slides be identified as control slides even though the studies that used different rat strains were to be evaluated separately and independently of each other. The remaining slides were to be recoded in a random manner. The SRI pathologist was provided with the PSTC histopathology lexicon and informed of the identity of three slides from control (vehicle-treated) animals for each study. These control slides were selected randomly by the BQRT by picking identification numbers. The study group identity of the remaining slides was blinded to the SRI pathologist. Despite a BQRT request that the remaining slides be recoded in a random manner, the Novartis slides were not recoded or randomized. Although the Merck slides were recoded, they were not completely randomized since half of each study group was numbered contiguously. Since the dose group could be potentially identified in both Merck and Novartis slide sets, the independent histopathology evaluation was only performed under partially blinded conditions.

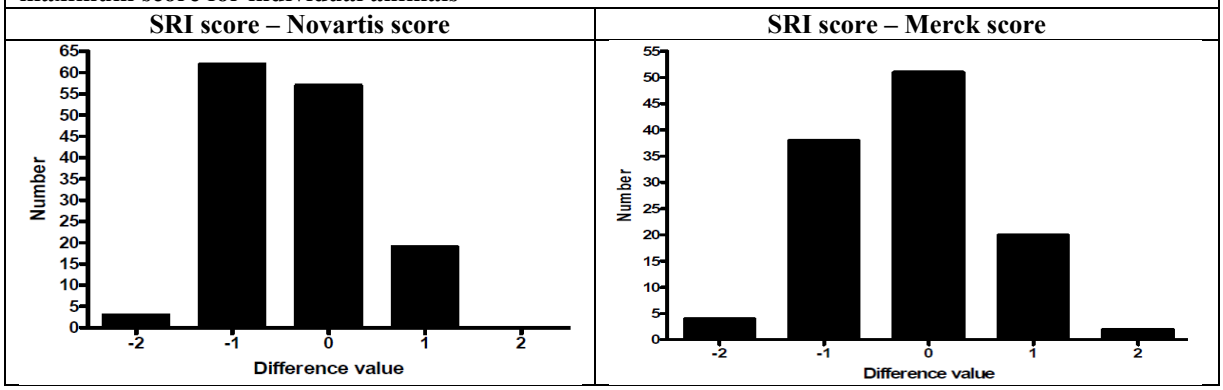
In June 2008 after many discussions between the PSTC and the BQRT, the PSTC provided additional descriptions of the histopathology evaluations (see Appendix 6.b.ix).

Results of the independent histopathology evaluation

Despite a specific PSTC request to SRI, the independent slide reading was not performed using the PSTC histopathology lexicon. Furthermore, the SRI pathologist used a severity grading scale of 0 to 4 rather than the PSTC histopathology severity grading scale of 0 to 5 used by the Merck and Novartis pathologists. Therefore, comparison of the SRI histopathology assessment to the histopathology assessment performed by Merck and Novartis is not straight forward.

If the histopathology assessments were identical, the BQTR expected a difference in score of zero. If the grading system was identical and the variation in reading between SRI and PSTC was normally distributed, we expected a normal distribution centering on a difference of zero. However, the SRI pathology evaluation used a scale of 0 to 4 instead of the scale of 0 to 5 used by the PSTC. Therefore, the skewing of the distributions to -1 observed in Figure 9 below was expected.

Figure 9: Reviewer's analysis – distribution of differences between PSTC and SRI histopathology maximum score for individual animals



A more detailed analysis of these differences is shown in Figure 10 using 4 X 5 tables to compare SRI and Merck and Novartis grades. For the Merck data, the differences were due in large part to the SRI pathologist designating a "+1"grade for cortical interstitial inflammation in the absence of other histopathology findings. Merck maintained that interstitial cellular inflammation was a background histopathologic feature, based their argument on the historical frequency of this finding and its presence in untreated control rats; hence Merck did not include such lesions in grading histopathology in control and treated rats. In the gentamicin study, the incidence of cortical interstitial inflammation in the absence of other findings was greater in treated groups (Control: 1; 20 mg/kg: 4; 80 mg/kg: 4) and hence the difference between the SRI and Merck histopathology evaluations were most apparent in this study. No incidences of interstitial inflammation were noted by the SRI pathologist in the control and low dose groups of the Merck cisplatin study. Since Novartis did not provide a 4 x 5 table, no detailed comparison can be made.

Figure 10: Summary of differences between SRI and PSTC histopathology evaluations

Reviewer analysis - Novartis versus SRI data

Novartis grade	SRI grade 0	1	2	3	4	Total
0	4	50	3	0	0	57
1	4	34	10	0	0	48
2	0	9	8	2	0	19
3	0	0	3	12	0	15
4	0	0	0	2	0	2
5	0	0	0	0	0	0
Total	8	93	24	16	0	141

Reviewer analysis - Merck versus SRI data

Merck grade	SRI grade 0	1	2	3	4	Total
0	35	18	3	0	0	56
1	0	3	2	1	0	6
2	0	13	4	1	0	18
3	0	0	0	2	1	3
4	0	0	0	5	4	9
5	0	0	0	4	19	23
Total	35	34	9	13	24	115

Merck analysis

Max Composite: Merck vs. SRI Severity Grade

Merck Grade \ SRI Grade	0	1	2	3	4	Total
0	**49**	7	0	0	0	56
1	3	**2**	1	0	0	6
2	0	14	**3**	1	0	18
3	0	0	1	**1**	1	3
4	0	0	0	5	**4**	9
5	0	0	0	4	19	23
Total	52	23	5	11	24	115

Reviewer analysis - Merck versus SRI data using only Merck included lesions

Merck grade	SRI grade 0	1	2	3	4	Total
0	48	8	0	0	0	56
1	3	2	1	0	0	6
2	0	14	3	1	0	18
3	0	0	1	1	1	3
4	0	0	0	5	4	9
5	0	0	0	4	19	23
Total	51	24	5	11	24	115

The Novartis report on the semi-blinded SRI independent evaluation indicated a discrepancy of 42% animals in the assignments to non-injured and injured group between the SRI assessment and the Novartis assessment of the subset of the Novartis slides. Factors potentially contributing to this discrepancy include blind reading versus non-blinded reading, slide quality, labeling issues, different level of detail of histopathology assessment, inter-pathologist variation and limited sub-set of data. However, the PSTC maintains that this difference of pathology assessment only moderately affected the results of the ROC analysis. Although the absolute performance (AUC) values of KIM-1 and clusterin decreased for the SRI data compared to the Novartis data using exclusion analysis in Table 14 below, the resulting AUC values were still greater than the corresponding AUC values for the accepted biomarkers of BUN and sCr. Novartis did not provide inclusion analysis of this data. The conclusions that the new biomarkers of KIM-1 and clusterin out perform the accepted biomarkers appear to be still valid based on the exclusion analysis of the SRI evaluation of the Novartis studies. However, the AUC values for the new biomarkers decreased more than the AUC values for the accepted biomarkers in the SRI independent analysis compared to the Novartis analysis. *Therefore, blinded assessment of histopathology is recommended in the qualification of new biomarkers.*

Table 14: Cisplatin and Gentamicin studies only – Exclusion analysis

ROC analysis - Novartis pathology-"Proximal Tubular Damage"

	Pathology	AUC	Thr	Spec %	Sens %	Direct	N° Ctrls	N° Dis
Kim-1	Necr Apopt Cell Sloughing S1-S3 + non-loc.	0.95 ± 0.02	2.17	97	88	+	35	50
Clusterin	Necr Apopt Cell Sloughing S1-S3 + non-loc.	0.93 ± 0.03	1.81	97	70	+	35	50
Creatinine	Necr Apopt Cell Sloughing S1-S3 + non-loc.	0.66 ± 0.06	1.07	97	30	+	35	50
BUN	Necr Apopt Cell Sloughing S1-S3 + non-loc.	0.54 ± 0.06	0.85	97	24	-	35	50

ROC analysis for SRI pathology "PCT Necrosis"

	Pathology	AUC	Thr	Spec %	Sens %	Direct	N° Ctrls	N° Dis
Kim-1	SRI: Necrosis PCT	0.82 ± 0.04	1.51	96	70	+	24	64
Clusterin	SRI: Necrosis PCT	0.84 ± 0.04	1.81	96	42	+	24	64
Creatinine	SRI: Necrosis PCT	0.62 ± 0.06	1.15	96	45	+	24	64
BUN	SRI: Necrosis PCT	0.53 ± 0.07	0.87	96	28	-	24	64

After a discussion of the differences between the Merck and BQRT analysis (10/15/07), Merck performed ROC analysis of the SRI histopathology evaluation of the Merck data that both included and did not include the finding of interstitial inflammation. Similar to that observed with the Novartis SRI comparison above, the AUC values for the SRI histopathology with and without interstitial inflammation are generally lower than the AUC values for the Merck histopathology for all the biomarkers. Both the Merck and SRI AUC values for KIM-1 are still greater than the corresponding AUC values for the accepted biomarkers of BUN and sCr. However, the Merck and SRI AUC values for albumin are either lower than or similar to the corresponding AUC values for BUN and sCr. Thus, the conclusion that KIM-1 "*out performs*" the accepted biomarkers is still valid based on the SRI evaluation of the Merck data. However, the conclusion that albumin also outperforms BUN and sCr biomarkers can not be established solely on the SRI evaluation of the Merck data using the inclusion analysis. Since SRI evaluated both the Merck and Novartis slides in the same manner, the possibility exists that the evaluation was influenced by the lack of randomization for the Novartis slides and the partial randomization of the Merck slides.

Table 15: Compilation of the analysis of Merck versus SRI histopathology results

Initial ROC analysis by Merck - Inclusion model – Merck VXDS03 – 8/20/07

Merck Histopathology

Marker	AUC	SE	log2 cutoff	fold cutoff	FPR	TPR	npos	nneg
Kim-1.ucr	0.983	0.015	3.19	9.16	0.043	0.95	40	47
S.Cr(Kim-1) mg/dL	0.959	0.023	0.26	1.20	0.043	0.88	40	47
BUN(Kim-1) mg/dL	0.903	0.035	0.32	1.25	0.043	0.75	40	47
Albumin.ucr	0.944	0.025	2.63	6.18	0.037	0.75	44	54
S.Cr mg/dL	0.964	0.020	0.26	1.20	0.037	0.89	44	54
BUN mg/dL	0.903	0.033	0.40	1.32	0.037	0.73	44	54

Subsequent Merck analysis of SRI data– VXDS03 and Supplement 11/08/07

Merck Maximum Composite Tubular Injury Score (SRI assessment)

Marker	AUC	SE	log2 cutoff	fold cutoff	FPR	TPR	npos	nneg
Kim-1.ucr	0.935	0.028	3.19	9.16	0.045	0.88	43	44
SCr(Kim-1) mg/dL	0.890	0.036	0.26	1.20	0.045	0.81	43	44
BUN(Kim-1) mg/dL	0.899	0.035	0.32	1.25	0.045	0.70	43	44
Albumin.ucr	0.863	0.038	2.63	6.18	0.040	0.69	48	50
S.Cr mg/dL	0.885	0.035	0.26	1.20	0.040	0.81	48	50
BUN mg/dL	0.904	0.032	0.32	1.25	0.040	0.71	48	50

BQRT Maximum Composite Tubular Injury Score (SRI assessment) (includes interstitial inflammation)

Marker	AUC	SE	log2 cutoff	fold cutoff	FPR	TPR	npos	nneg
Kim-1.ucr	0.921	0.030	3.19	9.16	0.026	0.80	49	38
SCr(Kim-1) mg/dL	0.853	0.040	0.32	1.25	0.026	0.63	49	38
BUN(Kim-1) mg/dL	0.842	0.042	0.40	1.32	0.026	0.59	49	38
Albumin.ucr	0.865	0.036	2.56	5.88	0.048	0.61	56	42
S.Cr mg/dL	0.842	0.039	0.26	1.20	0.048	0.70	56	42
BUN mg/dL	0.844	0.039	0.32	1.25	0.048	0.61	56	42

AUC: area under the curve, se: standard error, fold.cutoff: fold change, FPR: specificity (~0.05 error), TPR: sensitivity, npos: number positive samples by histopathology, nneg: number negative samples by histopathology. Maximum composite score was used for histopathology. Note there were 11 samples for which KIM-1 was not measured. BUN(KIM-1) and S.Cr(KIM-1) are values for an analysis of BUN and S.Cr with these 11 samples

Subsequent Merck analysis of SRI data– VXDS03 and Supplement 11/08/07

Merck Maximum Composite Tubular Injury Score (SRI assessment)

Marker	AUC	SE	log2 cutoff	fold cutoff	FPR	TPR	npos	nneg
Kim-1.ucr	0.935	0.028	3.19	9.16	0.045	0.88	43	44
SCr(Kim-1) mg/dL	0.890	0.036	0.26	1.20	0.045	0.81	43	44
BUN(Kim-1) mg/dL	0.899	0.035	0.32	1.25	0.045	0.70	43	44
Albumin.ucr	0.863	0.038	2.63	6.18	0.040	0.69	48	50
S.Cr mg/dL	0.885	0.035	0.26	1.20	0.040	0.81	48	50
BUN mg/dL	0.904	0.032	0.32	1.25	0.040	0.71	48	50

Prior to receipt of the SRI results, the BQRT decided that a statistical difference in the ROC AUC values would be the criteria for deciding that the independent histopathology evaluation produced a different result from the PSTC histopathology evaluation. The table below summarizes the AUC values from the inclusion and exclusion ROC analysis of the Merck versus the SRI histopathology evaluations. The AUC values decreased for the SRI evaluations compared to the AUC for the Merck evaluation, except for BUN for the SRI evaluation without interstitial inflammation (SRI/Merck). Examination of the confidence interval for each point estimate for the AUC values suggests that only the inclusion ROC analysis for BUN using the SRI

evaluation with interstitial inflammation (SRI/BQRT: 0.92 – 0.764) was significantly different from the Merck evaluation (1.004 – 0.924). However, in the Novartis exclusion analysis, the AUC values for KIM-1 and clusterin appear to be significantly higher than the corresponding AUC values using the SRI evaluation, while the AUC values for BUN and sCr are not. The ROC analysis of the Novartis and SRI data using the inclusion model were not submitted.

Table 16: Reviewer's summary of AUC values from ROC analysis of Merck/Novartis versus SRI [mean (SE)]

	Inclusion analysis			**Exclusion analysis**		
Biomarker	Merck	SRI/Merck	SRI/BQRT	Merck	SRI/Merck	SRI/BQRT
KIM-1.ucr	0.983 (0.015)	0.935 (0.028)	0.921 (0.030)	1.000 (NA)	0.978 (0.017)	0.975 (0.017)
SCr (KIM-1)	0.959 (0.023)	0.890 (0.036)	0.853 (0.040)	0.946 (0.028)	0.909 (0.037)	0.865 (0.044)
BUN (KIM-1)	0.903 (0.035)	0.899 (0.035)	0.842 (0.042)	0.900 (0.038)	0.904 (0.038)	0.853 (0.046)
Albumin.ucr	0.944 (0.025)	0.863 (0.038)	0.865 (0.036)	0.987 (0.013)	0.958 (0.023)	0.960 (0.021)
SCr	0.964 (0.020)	0.885 (0.035)	0.842 (0.039)	0.950 (0.025)	0.899 (0.037)	0.844 (0.045)
BUN	0.903 (0.033)	0.904 (0.032)	0.844 (0.039)	0.901 (0.037)	0.906 (0.035)	0.847 (0.045)
				Novartis	SRI/Novartis	
KIM-1.ucr				0.95 (0.02)	0.82 (0.04)	
Clusterin				0.93 (0.03)	0.84 (0.04)	
SCr				0.66 (0.06)	0.62 (0.06)	
BUN				0.54 (0.07)	0.53 (0.07)	

BQRT recommendations concerning histopathology in biomarker qualification

If a study is designed specifically for biomarker qualification, the BQRT believes that blinded histopathology evaluation should be prespecified in the protocol. If the study has objectives other than biomarker qualification, then the protocol should define the data that the pathologist will have access to at the initial and any subsequent histopathology evaluations. For minimal qualification of biomarkers, retrospective studies in which the pathologist is only blinded to both the new and any standard comparator biomarker results may be included in the submission for qualification. In all biomarker submissions, the study report should outline the data the pathologist had knowledge of during the histopathology evaluation.

The BQRT believes histopathology assessments should be conducted differently in drug safety and biomarker qualification studies. At this time, the BQRT recommends that future biomarker qualification studies ideally do the following:

(1) assess the impact of "background lesions" and morphologic variations on biomarker performance,
(2) ensure adequate tissue sampling, and
(3) maintain blinding of the reviewing pathologist to biomarker results, treatment assignment and other aspects of the study design or results that could potentially unblind the reviewing pathologist to this information.

iii. ROC curves generated for different histopathology grades

The PSTC Nephrotoxicity Working Group originally dosed rats with over forty compounds and generated histopathologic, clinical chemistry, and biomarker data limited primarily to the kidney. Novartis Pharma AG and Merck Research Laboratories analyzed the data from these studies using receiver operating characteristic analyses (ROC). These ROC curves graphically present the sensitivity and specificity of the biomarkers for nephrohistopathologic findings across the compounds tested. Initially, all of the histopathologic scores for a given compound and dose of a compound were used to generate the ROC curves; that is, the biomarker characteristics (sensitivity and specificity) were plotted for pathology findings up to and including the greatest damage (scores of 5 for the Novartis data and 4 for the Merck data) in Figures 1 and 2, respectively. These ROC curves showed that, generally, for a subset of new biomarkers studied (including KIM-1 and albumin), greater sensitivity and specificity was achieved than for the traditional renal toxicity biomarkers, blood urea nitrogen (BUN) and sCr.

However, the FDA biomarkers qualification review team (BQRT) noted that while there is interest in the performance of the biomarkers under worst case (highest pathology grade) conditions, there is interest as well in the comparative performance of the biomarkers where the renal damage is slight or moderate. To what extent were the results weighted by pathology effects at the high end? For example, does the performance of the novel biomarkers become indistinguishable from that of BUN or sCr when the maximum histopathology is limited to a score of 3 or lower?

Consequently, the BQRT requested the PSTC Nephrotoxicity Working Group generate ROC curves for specific biomarkers at additional (lower) intervals of histopathology. These ROC curves in Figure 3 were generated by Merck and Novartis and presented to the BQRT during October 2007. Merck examined three additional histopathologic severity grade intervals (0 to 3, 0 to 2, and 0 to 1) and concluded performance of urinary biomarkers KIM-1, albumin, and TFF3 increased or remained the same relative to the appropriate comparator (BUN and sCr). Novartis also examined the same intervals and concluded the performance of urinary biomarkers KIM-1, clusterin, cystatin C, and β2-microglobulin exceeded that of BUN and sCr in all analyses.

d. **Performance of Proposed Biomarkers Compared with Accessible Biomarkers in Current Use.**

i. Collection of samples

The separate descriptions of sample collection and handling provided by Merck and Novartis were complex. At the request of the BQRT, the PSTC provided a simplified flow chart of sample handling (see Appendix 6.b.vi).

However, the BQRT notes some differences in the Merck and Novartis reports concerning sample collection. Although both Merck and Novartis collected urine from fasting animals in metabolism cages for 16 and 16-20 hours, respectively, the BQRT notes that Merck collected the urine into containers on dry ice, while Novartis collected urine into containers on wet ice. Additionally, the Novartis animals had access to food during the period between the end of urine collection at 6 AM and necropsy later that day, while the Merck animals were fasted until necropsy. The analysis of the TFF-3 biomarker was especially complicated since this biomarker is unstable in urine and its analysis required a concentration step not required by other biomarkers.

ii. Blinding of biomarker assays

The PSTC stated that the biomarker assays were blinded. However, a description of the blinding process was not provided. It is unclear whether the samples were randomized prior to blinding.

iii. Urinary creatinine normalization

The PSTC used urinary creatinine (UCr) as the reference standard for normalization. The BQRT requested a discussion of the limitations of creatinine normalization. The PSTC submitted a discussion of the limits and effects upon urinary creatinine normalization procedure in the Merck report submitted 09/05/07. This discussion examined the effects of exercise, sex differences, hypohydration, urinary acidity, specific activity, and collection time, and diet on urinary creatinine excretion, creatinine clearance, and most importantly urinary creatinine normalization. However, most of the discussion focused on creatinine normalization in humans, not animals. The PSTC did note that dietary restriction in rats (reduced protein intake) is associated with lower excreted urinary creatinine levels compared to ad libitum fed animals (Tucker et al 1976). The PSTC concluded that the urinary creatinine measurements in the VXDS submission were not subject to many of the discussed effects since the sex, diet, and age of the animals and the duration of urine collection were highly controlled in the study protocols.

The literature search by the BQRT confirmed that dietary restriction can alter creatinine excretion. Overnight collection of urine from fasting rats resulted in decreased creatinine clearance when compared to overnight collection of urine from fed rats (Van Liew et al 1993). Similarly, starvation of rats resulted in decreased water intake, urinary creatinine excretion and creatinine clearance (Thompson et al 1987), although during the first day following starvation urinary output paradoxically increased, but fell dramatically thereafter. (Van Liew et al 1993; Verbaeys 1995). Interpretation of creatinine clearance data may be complicated by the detection method used, since nonspecific chromagens in plasma and urine

may interfere with the determination of creatinine, overestimating plasma and urinary concentrations (Tartoff 1996).

Since creatinine excretion is a function of muscle mass (Wassner et al 1977), indirect effects on muscle degradation can also alter creatinine excretion. An increase in creatinine excretion concomitant with an increase in sCr was indicative of muscle damage immediately post-flight in space-flown rats (Wade et al 1998).

Creatinine may undergo tubular secretion in the kidney, thus overestimate the glomerular filtration rate (GFR) to the extent that tubular secretion occurs. The creatinine clearance adequately reflected GFR in three rat strains (Wistar, Wistar-Kyoto and the Spontaneously Hypertensive Rat), but not in two strains of the Biobreeding/Worcester rat in which creatinine clearance was lower than the inulin clearance (Van Liew et al 1993). The depressed clearance of creatinine in the Biobreeding rat suggested tubular reabsorption of filtered creatinine. Evidence for 30-50% reabsorption of endogenous creatinine along the tubulule has been described previously in the Fisher male rat (Namnum et al. 1983), in female Lewis rats (Darling and Morris 1991), and in an unknown strain of male albino rats Harvey and Malvin [1965]. In contrast, in the female Spraque-Dawley rat, creatinine and inulin clearances are identical over a wide range of filtration rate and in a variety of experimental renal diseases [Zager 1987].

The renal clearance of endogenous creatinine is widely used to assess glomerular filtration rate (GFR) and renal function in animal investigations. Urine creatinine excretion determined over a specific time period has frequently been used to normalize renal function data in Brown-Norway and Sprague Dawley rats (Weichert-Jacobsen 1999, Ngai 2006). Creatinine clearance normalization has been used in characterizing drug-induced nephrotoxicity in rats. For example, urinary creatinine excretion progressively decreased beginning 4 days after the start of gentamicin administration to Sprague Dawley rats concomitant with progressive increases in the excretion of sodium and serum urea nitrogen that peaked on day 10 and 12, respectively. However, despite continued gentamicin administration, the urinary creatinine excretion gradually increased (Soejima et al 1998). Using spontaneously voiding, freely moving rats, Haas et al (1997) found that creatinine excretion, which is supposed to be constant over the day, showed fluctuations paralleling the variation in urinary flow suggesting incomplete bladder emptying. The use of creatinine excretion reduced the variation coefficient of sodium excretion from 61 +/- 17% to 29 +/- 5% during normal diuresis. Hass et al (1997) indicated that normalization by creatinine excretion allows detection of changes in a urinary parameter if the change exceeds a 40% deviation of the normal value.

iv. Background (control range) biomarker levels

How do age, gender and rat strain affect the background control ranges? The following discussion is from the combined PSTC report submitted 09/05/07.

> *It is known that different rat strains can react differently to toxic agents leading to different histopathological outcomes for the same doses (Kulokarni et al 1999; Tuomisto and Pohjanvirta 1991). In the context of this submission, histopathologic findings were used as the 'gold standard' to qualify the biomarkers. Therefore, biomarker levels were directly compared to histopathological assessments in individual animals. If rat strains show different sensitivity to toxic treatment, this might be reflected in varied molecular response and histopathologic outcome. Therefore, different sensitivities of rat strains to toxic treatment are not predicted/expected to influence the qualification, as long as molecular response and histopathologic change have the same biological relationship in both rat strains.*

To investigate if the relationships between molecular responses of the biomarkers and histopathology outcome differ from strain to strain (e.g. a lesion of grade 1 corresponds to a 3-fold increase of a biomarker in the first strain and a 10-fold increase in the second strain), a completely matched study design including both strains (same food, housing, dose-levels, administration regimen, measurements of same biomarkers, assessment of histopathology by same pathologist etc.) would be an ideal experimental approach, to exclude any additional possible sources of variation. Such studies were not available for the current submission.

v. Threshold definition

The biomarker Kim-l was measured in both strains in different studies, whereby separate ROC analyses were performed for the Novartis studies, the FDA studies and the Merck studies in the context of the pathology "tubular damage". When comparing the thresholds for 95% specificity of these three analyses, the PSTC determined that the thresholds for Kim-l were highly comparable, with a 1.87-fold increase for Novartis (418 Han Wistar rats), a 1.88-fold increase for Merck (114 Sprague-Dawley rats) and a l.77-fold increase for the FDA studies (168 Sprague-Dawley rats, even though a colony infection was suspected). These study data suggested that there are no major differences for Kim-l threshold, although different rat strains were used. In addition, the published literature for KIM-1 does not indicate differences of molecular mechanisms and biological responses amongst strains of rat. Furthermore, the mechanisms of Kim-l expression and excretion into urine are reported to be similar for humans and rats (Han et al 2004; Ichimura 2004). The routine application of these biomarkers to different rat strains will provide further evidence on the impact of strain variation to biomarker thresholds and performance.

The background ranges for the biomarkers as provided by the PSTC are summarized in Table 17 below.

Table 17: Normal ranges

Merck – Sprague Dawley

Concentration/UCmn.mg.dL

	n	min	q05	mean	median	q95	max	sd
TFF3.ng.mL	101	0.25	0.90	4.95	4.60	9.49	14.09	2.69
Albumin.EIA.ug.ml	88	0.02	0.07	0.21	0.15	0.46	1.00	0.16
Albumin.Roche.ug.mL	114	0.04	0.07	0.17	0.11	0.43	1.74	0.20
Kim.1.pg.mL	46	0.29	0.51	1.73	1.27	3.72	10.21	1.57

all values are per U.Cmn.mg.dL

All Sprague-Dawley Males (n=302)

Rprogram 107

Norvatis

Normal Levels:
Kim-1: 190 pg/g Creatinine
Clusterin: 256 ng/g Creatinine

Threshold 95% Specificity:
Kim-1: 355 pg/g Creatinine
Clusterin: 473 ng/g Creatinine

Normal Levels:
Cystatin C: 442 ng/g Creatinine
B2-Microglobulin: 8.6 µg/g Creatinine
Total Protein: 770 µg/g Creatinine

Thresholds 99% Specificity:
Cystatin C: 1376 ng/g Creatinine
B2-Microglobulin: 31 µg/g Creatinine
Total Protein: 1462 µg/g Creatinine

vi. Specificity and secondary organ effects

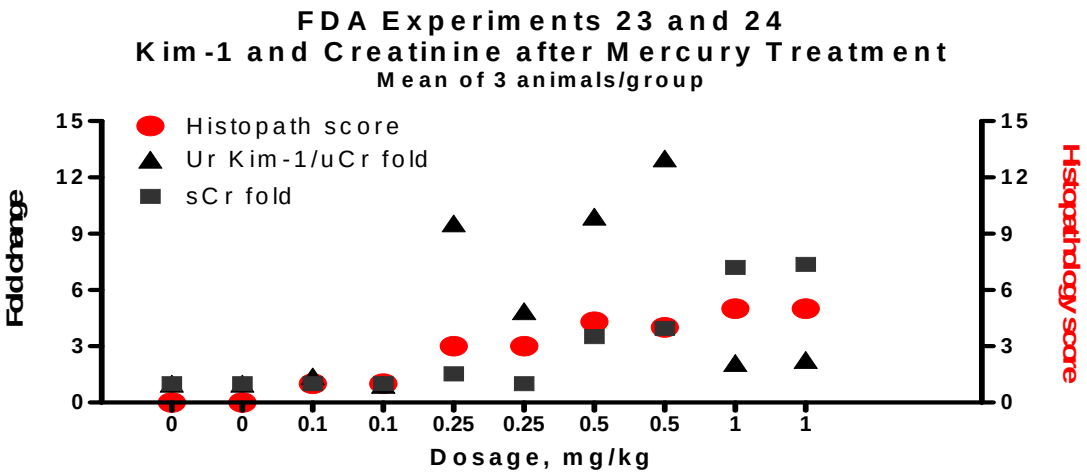
Although results for eleven non-nephrotoxicant studies were provided, not all biomarkers were measured in these studies. For instance, KIM-1 was measured in two non-nephrotoxicant Novartis studies, but not in the nine non-nephrotoxicant Merck studies. Therefore, the specificity of the biomarkers has been inadequately characterized. Furthermore, examination of the individual values for animals treated with non-nephrotoxicants methapyrilene or ANIT in the Novartis data indicates some values above the decision threshold for 95% specificity. These high values were in animals treated with the high dose of each compound. Similarly, in the Merck data for albumin, values for some animals show a dose-related increase with some non-nephrotoxicants, including cerivastatin, diuresis with 4% sucrose, furan, and genipin.

vi. Unusual findings:

An unexpected result was observed with mercury (Hg) as a nephrotoxicant. As illustrated in Figure 11 below, at doses up to 0.5 mg/kg of Hg, the fold change in KIM-1 was greater than the fold change for sCr. However, at the highest dose of Hg (1 mg/kg) the fold change for KIM-1 was below that of sCr despite a histopathology score of 5. In contrast, kidney KIM-1 mRNA expression was increased by 71-, 187-, and 167-fold at Hg doses of 0.25, 0.5 and 1 mg/kg (Zhou et al 2008). One explanation for the markedly reduced appearance of urinary KIM-1 may be related to decreased KIM-1 protein synthesis in the necrotic tubular cells. A second explanation involves the binding of urinary excreted Hg to cysteine residues in KIM-1 resulting in an altered KIM-1 structure that is not

recognized by the monoclonal antibodies used in the microbead assay. A third explanation involves an altered KIM-1 structure that can not be recognized by the protease that cleaves the KIM-1 ectodomain allowing excretion into the urine. A fourth explanation involves direct inhibition by Hg of the protease that cleaves the KIM-1 ectodomain resulting in reduced KIM-1 excretion into urine.

Figure 11: Reviewer's analysis of FDA experiments 23 and 24



f. **Format issues**

i. Data for individual animals

Initially, data for individual animals was not provided. The PSTC complied with BQRT request for this data by supplying separately for each site (FDA, Merck, and Novartis) an Excel spreadsheet that contained animal ID, histopathology score according to the lexicon, biomarker data, clinical chemistry data, and any data transformations. Although somewhat unmanageable, this format provided the requested information and was acceptable. However, any reviewer analysis of this data required considerable time to extract critical information for a particular study. Submission of additional Excel spreadsheets by study or compound would be helpful.

ii. Data from the independent histopathologist

Data from the SRI pathologist was transmitted as Word documents in which the lesions for each animal could not be properly identified until considerable time was spent reformatting the document. Then the data needed to be manually inserted into an Excel document for comparison with the PSTC data. Because of inefficient use of reviewer time, this format is unacceptable and is not recommended in future VXDS submissions.

5. Qualification Conclusions

a. ***The BQRT concludes after review of the PSTC Biomarker Qualification Data Package that:***

The urinary kidney biomarkers (KIM-1, albumin, total protein, β2-microglobulin, cystatin C, clusterin and trefoil factor-3) are considered acceptable regulatory tools in the context of nonclinical drug development for the detection of acute drug-induced nephrotoxicity. These biomarkers may be used voluntarily as additional evidence of nephrotoxicity to complement the standard data reported in nonclinical safety assessment studies. These biomarkers provide additional and complementary information to BUN and sCr that correlates with histopathological alterations considered to be the gold standard. **The ROC analyses showed that some of these biomarkers have better sensitivity and specificity than BUN and sCr when tested with a limited number of nephrotoxicant and control compounds.** While further studies are needed to support a broader application claim, the data submitted thus far appear to be sufficient to support the voluntary testing proposed by the PSTC.

It is worthwhile to explore further the potential in early clinical trials of KIM-1, albumin, total protein, β2-microglobulin, cystatin C, clusterin and trefoil factor-3 as clinical biomarkers for acute drug-induced kidney injury. The BQRT considered, however, that until further data are available to correlate these biomarkers with the evolution of the nephrotoxic alterations and their reversibility, their use for monitoring nephrotoxicity in the clinical setting cannot be recommended.

b. **Recommended application context for the voluntary use of these urinary biomarkers:**

KIM-1 is a biomarker that may be used by sponsors to detect acute drug-induced kidney tubular alterations in rats and can be included along with traditional clinical chemistry markers and histopathology in GLP toxicology studies which are used to support renal safety in clinical trials.

Albumin is a biomarker that may be used by sponsors to detect acute drug-induced kidney tubular alterations in rats and can be included along with traditional clinical chemistry markers and histopathology in GLP toxicology studies which are used to support renal safety in clinical trials.

Total Protein is a biomarker that may be used by sponsors to detect acute drug-induced glomerular alterations/damage and/or impairment of kidney tubular reabsorption in rats and can be included along with traditional clinical chemistry markers and histopathology in GLP toxicology studies which are used to support renal safety in clinical trials.

β2 microglobulin is a biomarker that may be used by sponsors to detect acute drug-induced glomerular alterations/damage and/or impairment of kidney tubular reabsorption in rats and can be included along with

traditional clinical chemistry markers and histopathology in GLP toxicology studies which are used to support renal safety in clinical trials.

Cystatin C is a biomarker that may be used by sponsors to detect acute drug-induced glomerular alterations/damage and/or impairment of kidney tubular reabsorption in rats and can be included along with traditional clinical chemistry markers and histopathology in GLP toxicology studies which are used to support renal safety in clinical trials.

Clusterin is a biomarker that may be used by sponsors to detect acute drug-induced kidney tubular alterations in rats and can be included along with traditional clinical chemistry markers and histopathology in GLP toxicology studies which are used to support renal safety in clinical trials.

Trefoil Factor-3 is a biomarker that may be used by sponsors to detect acute drug-induced kidney tubular alterations in rats and can be included along with traditional clinical chemistry markers and histopathology in GLP toxicology studies which are used to support renal safety in clinical trials.

c. Strengths and Limitations of Data Submitted

The PSTC submission included multiple studies that were combined to explore the sensitivity and specificity of seven biomarkers of nephrotoxicity in rats. This submission is an example of a focused, context-dependent qualification proposal, suitable for evaluation by the pilot FDA qualification process. However, there were deficiencies in the design of the studies that limit our confidence in the PSTC's conclusions. These limitations are described below.

1. Only two nephrotoxicant drugs and no non-nephrotoxicant drugs were assessed by all three sites, providing a very limited database for assessing the capacity of these biomarkers to reliably demonstrate chemical-induced toxicity. Data are hence insufficient to fully evaluate the performance characteristics of these biomarkers. Future submissions should include an evaluation of performance characteristics by exposing the same strains of rats to additional nephrotoxicant drugs and non-nephrotoxic drugs from different mechanistic classes and in altered physiologic conditions.
2. Limited data were submitted on the analytical validation of these assays. Future submissions should include a detailed section on the methods and results of analytical validation of assays, including specificity, biomarker stability and sample handling.
3. The provided animal data lacked critical information on the body weight, food consumption and water consumption of individual animals. This is problematic because these factors affect levels of

urinary creatinine. Future submissions should include data on these variables.

4. Histopathology, used as the gold standard in this submission, was not evaluated in a fully blinded manner. According to the PSTC, pathologists at both sites were blinded to the results of the novel biomarker. However, most of the pathology assessments were conducted with knowledge of treatment groups. Additionally, it remains unclear if the pathologists were aware of the traditional biomarker (BUN and sCr) values at the time of the pathology assessment. Incomplete blinding can lead to bias especially in the histopathology evaluation of less severe lesions, resulting in falsely elevated estimates of biomarker performance. To avoid bias, the BQRT strongly recommends that the evaluation of histopathology and biomarker results be conducted in a fully blinded manner in future biomarker studies.
5. The PSTC based the histopathology scores on the evaluation of only one 5 μm section of one kidney per animal. In the absence of data establishing the ability of a single section to accurately characterize the presence, extent, maximum severity, and location(s) of injury, we recommend that multiple histopathology sections be taken in biomarker qualification studies
6. Variation in biomarker levels in control and treated animals may be influenced by so-called "background lesions" and morphologic variations. Future biomarker qualification studies should assess the impact of "background" lesions and morphologic variations on biomarker performance and include a list by animal of all the variations (common as well as uncommon lesions) in the target tissue.
7. Biomarkers were separated into tubular-specific and glomerular-specific based on the association of the biomarker with either tubular or glomerular histopathologic lesions. Data suggested that KIM-1, albumin, TFF-3, and clusterin detect tubular lesions, while cystatin C, β2-microglobulin, and total protein detect glomerular lesions. Conclusions from these data are limited by the following:
    a. Limited number and type of nephrotoxic compounds
    b. Limited histopathological exploration of the tissue
    c. Use of high doses of nephrotoxicants limited conclusions because by the time the histopathology specimens were taken, the damage was extensive
    d. Absence of immunohistochemistry or other data showing localization of the biomarker to specific regions of the kidney.

To support conclusions based on the association with either tubular or glomerular histopathologic lesions, future submissions should include:

a. Data on behavior of the novel biomarkers using multiple nephrotoxic and nonnephrotoxic compounds to broaden our understanding of the generalizability of conclusions about the ability of the biomarkers to detect localizable lesions

b. Use of lower doses of nephrotoxic compounds so that histopathology specimens can be gathered when the injury is more localized

c. Immunohistochemistry or other data to demonstrate the localization of the biomarker to the damaged areas of the kidney.

8. Data were limited to injury detectable by histopathology. The data excluded other types of kidney injury, such as functional changes or inhibition of transporters in the proximal or distal tubule, resulting in glycosuria, aminoaciduria, phosphaturia, etc. Future studies should attempt to evaluate the behavior of the biomarkers for these other types of injury.

9. Data to adequately address the temporal relationship between biomarker levels and the emergence and recovery of the histopathologic alterations were not provided. This information is critical to establish that biomarker levels reflect the injury and recovery pattern observed by histopathology.

Together the above limitations indicate that application of these biomarkers to monitor renal toxicity has not yet been sufficiently demonstrated to stand on its own without histopathology as a measure of renal toxicity.

These novel renal biomarkers are only qualified for use in the rat. The data presented do not include other animal species. Development of assays for these novel biomarkers in other animal species would expand their utility in the nonclinical evaluation of nephrotoxic drugs.

Many of the submitted studies appeared to be exploratory in nature. Our confidence in these data would be greatly strengthened by prospectively designed hypothesis driven studies. Moreover, the prospective design of these studies would ensure the collection of data critical to evaluate the performance characteristics of these biomarkers.

d. Regulatory Recommendations

The submitted data support the voluntary use of seven urinary biomarkers in preclinical research alongside histopathology to identify drug-induced acute kidney injury in the rat. The recommended application contexts for the use of these urinary biomarkers using the submitted test parameters are as follows:

KIM-1, albumin, clusterin and trefoil factor-3 can be included along with histopathology and the standard renal clinical chemistries as biomarkers of drug-induced acute kidney tubular alterations in Good Laboratory Practice (GLP) rat studies used to support clinical trials.

Total protein, β2 microglobulin and cystatin C can be included as biomarkers of acute drug-induced glomerular alterations/damage and/or impairment of kidney tubular reabsorption in GLP rat studies used to support clinical trials.

However, further studies are needed to improve our understanding of how these markers respond in different animal models and with different drugs, and how best to interpret different biomarker levels. In order to gain useful information about the biomarker performance in different contexts, including the clinical setting, we recommend the following:

1. A standardized format for submitting preclinical and clinical data is needed to allow for an efficient and accurate review.

2. Consistency in approach, analysis, and presentation is a goal for biomarker qualification submissions. The achievement of this goal will allow efficient comparisons among studies for biomarkers of renal toxicity. For example, this will allow comparison of data on different biomarkers submitted by one member of the consortium. It will also allow comparison of data on a single biomarker submitted by different consortium members. Moreover, it will help with the creation of databanks that will ultimately expand the qualified context of use of these biomarkers.

3. The characterization of an endogenous substance in blood or urine requires a different testing paradigm than characterizing the effects of a xenobiotic. The data that was used in this submission were collected from studies that were designed for the characterization of a xenobiotic. Future studies will be more informative if designed for the purpose of assessing the putative biomarker.

4. In concurrence with the PSTC proposal, we agree that specific preclinical studies to support drug development should demonstrate that the novel biomarkers can detect early drug induced renal injury and reversibility of injury after drug cessation before proceeding to clinical studies.

5. Prospective preclinical studies are needed to address the correlation between biomarker levels and evolution of lesions with secondary

confirmation using appropriate techniques, such as immunohistochemistry, in-situ hybridization and/or electron microscopy, when appropriate relative to the biology of the biomarker and any claims concerning localization of injury.

6. In future biomarker qualification studies, pathologists need to be blinded to the results of biomarker analyses (including novel and traditional biomarkers such as BUN or sCr) at a minimum. Ideally, the pathologist will also be blinded to other aspects of study design or results that could potentially unblind the pathologist to treatment assignment or biomarker level.

7. While novel renal biomarkers should be tested in humans, they are not currently qualified to be used as primary renal injury monitoring tests or dose-stopping criteria. For the time being, the sponsor and regulatory division will decide on a case by case basis how best to implement these biomarkers in the clinical development program. Demonstration that a biomarker or a panel of biomarkers consistently detects toxicity at an early stage in animal models may justify incorporating them into clinical studies as sentinels for toxicity. Using novel renal biomarkers in early clinical trials for renal toxicity monitoring may represent a reasonable risk for the development of promising therapies which would otherwise be abandoned Use of a particular biomarker in a clinical trial will be dependent on demonstration of reversibility of both biomarker levels and histopathology and establishment of a prespecified cut-off value of abnormality.

6. Appendices

This section includes detailed information referenced in the main text of this review, including additional background information, as well as data submitted by the PSTC to support qualification of the proposed biomarkers of nephrotoxicity.

a. Background information about the proposed biomarkers submitted by the PSTC.

Previously published data on genomic biomarkers of nephrotoxicity (Han et al 2002, Silkensen et al 1997, Verstrepen et al 2001, Amin et al 2004, Thompson et al 2004) support the investigation of a number of accessible protein biomarkers as exploratory biomarkers with a high probability of success in diagnosing nephrotoxicity in rat (Han et al 2002) and monkeys (Davis et al 2004). In total, twenty-two (22) biomarkers were considered (Table 1) but for most of them there is still insufficient data to support a claim for qualification.

The C-Path Predictive Safety Testing Consortium (PSTC) made an original and eight supplementary submissions to the FDA and the EMEA between June and November 2007 to support the qualification of seven biomarkers of drug-induced acute kidney toxicity. A review of the scientific literature pertaining to exploratory studies in human clinical context relevant to these seven biomarkers is presented below.

i. Urinary KIM-1

Kidney injury molecule-1 (KIM-1) (see Table 1) is a type I cell membrane glycoprotein containing a unique six-cystein immunoglobulin-like domain and a mucin domain in its extracellular domain. Rat and human cDNAs encoding KIM-1 (KIM-1 in the rat) were initially identified by differential expression between normal and regenerating kidneys following ischemia/reperfusion (I/R) injury (Ichimura T et al. 1998). KIM-1 mRNA levels increase more than any other known gene after kidney injury. The ectodomain of KIM-1 is shed from cells in vitro (Bailly et al. 2002) and in vivo into the urine in rodents (Vaidya et al. 2006) and humans (Han et al. 2002) after proximal tubular kidney injury (Table 3). In preclinical and clinical studies using several mechanistically different models of kidney injury, urinary KIM-1 serves as an earlier diagnostic indicator of kidney injury when compared to any of the conventional biomarkers, e.g. plasma creatinine, BUN, glycosuria, increased proteinuria or increased urinary NAG, γ-GT, AP levels (Vaidya et al. 2006 and Han et al., 2002). An ELISA assay was first developed to measure KIM-1 in rodent and human urine samples, followed by a more sensitive, high throughput microbead-based assay to quantitate KIM-1 in rat urine (Vaidya et al. 2006). The microbead assay has a greater dynamic range and requires less urine volume (30 µl) and reagents than the conventional ELISA (Vaidya et al, 2006).

In human subjects, after adjustment for age, gender, and length of time delay between insult and sampling, a one-unit increase in normalized KIM-1 was associated with a subsequent greater than 12-fold increased

risk for the presence of acute tubular necrosis (ATN), whereas total protein, γ-glutamyltransferase, and alkaline phosphatase did not correlate with ATN (Han et al. 2002). In another clinical study, urinary KIM-1 was increased in affected renal patients (n=53) versus controls (n=11) ($p < 0.001$) after excluding minimal renal changes to patients that had benign outcomes and no observed renal structural damage (van Timmeren et al, 2007). Structural renal damage from biopsy such as interstitial fibrosis and macrophages, correlated positively with urinary KIM-1, while no correlation was found with proteinuria (van Timmeren et al., 2007). These observations indicate that urinary KIM-1, at a more stringent threshold than the PSTC KIM-1 data, does detect definitively diagnosed renal injuries in humans. Liangos et al (2007) prospectively evaluated the relationship between urinary KIM-1 and N-acetyl-beta-(D)-glucosaminidase activity (NAG) levels and adverse clinical outcomes in a cohort of 201 hospitalized patients with acute renal failure (ARF). Scores for Mean Acute Physiology, Age, Chronic Health Evaluation II (APACHE II) and Multiple Organ Failure (MOF) were used to evaluate patient outcome relative to biomarker values. Urinary KIM-1 and NAG increased in tandem with APACHE II and MOF scores. KIM-1 or NAG in combination with the covariates cirrhosis, sepsis, oliguria, and mechanical ventilation yielded an area under the receiver operator characteristic curve of 0.78 (95% CI 0.71 to 0.84) in predicting the composite outcome. In conclusion, common clinically used severity indices, such as sCr and urine output, had inferior prognostic value than did NAG and KIM-1.

Kidney Injury Molecule (KIM-1) properties.

PARAMETERS	DESCRIPTION	
Other names	Hepatitis A Virus Cellular Receptor-1 (HAVCR-1), T Cell Immunoglobulin like molecule-1 (TIM-1), Cochlear Injury Molecule-1 (CIM-1)	
Chromosomal location	Human (5q31), Chimpanzee (5), Rat (10q21), Mouse (11B1.1)	
Molecular Mass	Human (359 aa); Chimpanzee (predicted: 357 aa); Rat (307 aa); Mouse (305 aa)	
Structure	Immunoglobulin domain, threonine-serine-proline rich mucin domain, transmembrane domain, short cytoplasmic tail	
Number of KIM-1 homologues	Human (three: KIM-1, KIM-3, KIM-4), Rat (six: Kim-1 to Kim-6), Mouse (eight: Kim-1 to Kim-8)	
% Identity of human KIM-1 protein	Chimpanzee (79%), Rat (39%), Mouse (37%)	
Upregulation in kidney injury/toxicity/dedifferentiation models Rat: R Mouse:M Human:Hu	Adriamycin-nephrosis (1) R	Ischemia/reperfusion (8) R&Hu
	Brain dead donor rat kidney (2) R	Kidney fibrosis (9) R
	Cadmium (3) R	Mercuric chloride (4) R
	Chromium (4) R	Ochratoxin (10) R
	Cisplatin (5) R	Polycystic kidney disease (11) M
	Cyclosporine (6) R	Protein-overload (12) R
	Endotoxin	Renal cell carcinoma (13) Hu
	Folic acid (7) R	Sevoflurane (14) R
	Gentamicin (4) R	S-(1,1,2,2-tetrafluoroethyl)-l-cysteine (TFEC) (7) R

ii. Urinary Albumin

Albumin is often the major constituent of abnormal urine protein excretion and has been reported to provide superior diagnostic information compared with urine total protein measurements (Lydakis and Lip 1998). Clinically, the most significant use of urinary albumin is as a marker of

diabetic nephropathy and as a marker of the effectiveness of anti-hypertensive therapy in delaying end-organ disease progression (Lane 2003, Bakris 2004, Weir 2007). Also, albuminuria is a marker of cardiovascular disease in patients with type 1 and type 2 diabetes mellitus (Lane 2003, Bakris 2004, Weir 2007), increased intraglomerular pressure, and left ventricular hypertrophy (Lydakis and Lip 1998, Pedrinelli 2002, Lane 2003, Weir 2007).

Albumin values are generally expressed as excretion per timed urine specimen collection (24 hr) or per milligram of urine creatinine for a spot collection (Lane 2003). The critical range for measurement of urinary albumin, sometimes called "microalbuminuria", falls below the sensitivity of the traditional dipstick protein methods, reducing the utility of this methodology. Therefore, more sensitive antibody-based methods including ELISA, immunoturbidimetry, RIA and nephelometry or HPLC methods are required for interpretable albumin measurements (Weir 2007).

Urinary albumin has been reported to correlate with both glomerular and proximal tubular toxicity in the rat. Treatment effect on glomerular cell proliferation and inflammation has been tracked with urinary albumin measurements (Wagner 2002). Increases in glomerular filtration rate (GFR) typical of diabetic or hypertensive states cause increased tubular flow rate and decreases the overall reabsorption of albumin by the proximal tubules. Defective production of proteoglycans in the glomeruli of diabetic rats have been linked to a reduced charge selectivity of the glomerular barrier, and in obese rats changes in GFR are linked to increased proteinuria (Ruggenenti and Remuzzi 2006).

Ultrafiltered albumin is rapidly reabsorbed in the proximal tubule through binding with megalin and cubulin that are co-localized in the endocytic pits. Absence or dysfunction of this receptor mechanism would result in 90% of filtered albumin being lost into the urine matrix. (Ruggenenti and Remuzzi 2006, Christensen and Birn 2001, Russo et al 2007). Nephrosis in the rat, caused by aminonucleoside administration, results in increased albumin concentration in the proximal tubular fluid, increased total excretion in the urine and correlates with decreased expression of megalin in the proximal tubules (Oken and Flamenbaum, 1971, Russo et al 2007). Excess albumin has been shown to cause endoplasmic reticulum stress in proximal tubular cells (Ohse, et al 2006).

iii. Urinary Total Protein

Proteinuria has been highlighted both as a clinical prognostic marker and as a factor predicting progressive loss of renal function. Alteration of the glomerular filtration barrier is usually associated with damage of the glomerular podocyte and leads to leakage of proteins into the ultrafiltrate (Schmid et al 2003). Excess urinary protein leakage can in turn have a negative impact on tubular function as in the case of protein overload and thereby become also a marker of tubular dysfunction (Guder and

Hoffmann 1992). The normal glomerular filtrate contains 10 mg protein/L, but only approximately 1% is normally present in the urine because of the strong reabsorption capacity of the proximal tubule. If this reabsorption reaches a saturation point or if the tubule is damaged by toxic agents, proteinuria can be observed despite normal glomerular filtration. In progressive glomerular disease, dysfunctions of glomerular filtration and of tubular reabsorption are found together as tubulo-glomerular proteinuria. Proteinuria is diagnosed when total urinary protein excretion is greater than 300 mg/24h. This is the case for 10-15% of patients with hypertension (Rodicio and Rulope 1995). It has been postulated that the ratio between low and high molecular weight proteins in urine would allow a better prediction of the damage severity and grade of the glomerular lesions than the quantity of proteinuria (D'Amico and Bazzi 2003, Schieppati and Remuzzi 2003). Since albumin accounts for most of the protein in urine in proteinuria due to glomerular injury, added value of total urinary protein measurement compared to albumin and vice-versa is currently under debate (Eknoyan 2003, Eddy 2004). With normal tubular function however, proteolytic processing of albumin by the proximal tubule leads to urinary excretion of variably sized fragments that are not measurable by most immunologic methods, leading to underestimation of total albumin content in urine unless RIA or HPLC are utilized (Compton 2005). While some authors have shown that total urine protein is a good predictor of the podocyte injury and subsequent glomerulosclerosis (Shankland 2006), others have not found correlations between the podocyte effacement and the level of proteinuria in human glomerulopathies (van den berg et al 2004).

iv. Urinary β2 Microglobulin

β2-microglobulin is a single polypeptide chain of 12k Daltons and is small compared to albumin. It is part of the class I major histocompatibility complex and is present at the membrane of all nucleated cells. In healthy subjects, 150-200 mg of β2-microglobulin is synthesized per day. β2-microglobulin readily crosses the glomerular filtration barrier and is almost completely reabsorbed and metabolized by the tubules. Only 0.3% of the β2-microglobulin in the glomerular filtrate is normally excreted into urine (Miyata T, Jadoul M, Kurokaw K, et al 1998). It has been shown that impairment of tubular uptake causes an increase of urinary excretion of β2-microglobulin of up to several hundred fold. Three major mechanisms have been identified causing this impairment:

A. Glomerular damage can cause a high protein load in the tubules as higher molecular weight proteins pass through the glomerular filtration barrier. As a consequence, high molecular proteins such as albumin compete for common transport mechanisms, decreasing the tubular uptake and increasing the excretion of β2-microglobulin into urine. (Branten AJW, Buf-Vereeijken PW, Klasen Is et al 2005).

B. Direct functional impairment of the tubular reabsorption capability, for example, caused by treatment with tenofovir, disoproxil, fumarate or by different tubular diseases can result in decreased reabsorption and hence increased excretion of β2-microglobulin (Gatanaga H, Tachikawa N, Kikuchi Y et al 2006, Thielemans N, Lauwerys R, Bernard A 1994).

C. Competition for tubular reabsorption processes with drugs filtered through glomeruli can incease β2-microglobulin excretion. For example, gentamicin causes inhibition of protein reabsorption by the tubules by a mechanism similar to that of polycationic proteins. Even at therapeutic doses of gentamicin, urinary β2-microglobulin levels are increased in the absence of kidney injury and without increases in other markers of kidney-injury (Bernard A, Viau C, Ouled A, et al 1986, Gatanaga H, Tachikawa N, Kikuchi Y 2006, Kaye WA, Griffiths WC, Camara PD, et al 1981, Rybak MJ, Frankowski JJ, Edwards DJ et al 1987, Sorensen PG, Nissen MH, Groth S, et al 1985). As the tubular reabsorption rates of small-molecular proteins like β2-microglobulin are greater than those of high-molecular proteins, the current opinion is that the low molecular weight proteins are a lot more sensitive for detection glomerular alterations and changes in tubular reabsorption (Thielemans N, Lauwerys R, Bernard A 1994).

Many applications of β2-microglobulin as a marker for renal injury have been reported, such as monitoring kidney injury in populations exposed to heavy metals (Aoyagi T, Hayakawa K, Miyaji K, et al 2003), prediction of the prognosis of patients with idiopathic membraneous nephropathy (Gatanaga H, Tachikawa N, Kikuchi Y et al 2006), use as part of an identification and differentiation process for various renal diseases (Guder WG, Hofmann W 1992), and monitoring for kidney injury in patients treated with tenofovir, cisplatin or other nephrotoxic drugs (Gatanaga H, Tachikawa N, Kikuchi Y et al 2006, Trollfors B, Bergmark J, Hiesche K, Jagenburg R 1984, Trof RJ, Di Maggio F, Leemreis J, et al 2006).

v. Urinary Cystatin C

Cystatin C, called also γ-trace, is a non-glycosylated low-molecular protein with a molecular weight of 13,360 Daltons. Cystatin C is a cysteine protease inhibitor and is a member of the human cystatin family. It is continuously produced by all human nucleated cells. (Mussap M, Plebani M 2004). Cystatin C is freely filtered from blood at the level of the glomerulus. In rats, 99.5% of the filtered cystatin C is reabsorbed and metabolized by renal tubules (Tenstad O, Roald AB, Grubb A et al 1996).

Independence of changes in cystatin C from effects of age and muscle mass, the free filtration by glomeruli and the absence of tubular secretion or extrarenal clearance makes serum cystatin C an ideal estimator of the glomerular filtration rate (GFR). It has also been shown in numerous studies that serum cystatin C is comparable to many of the most reliable methods for estimating GFR (Madero M, Sarnak MJ, Stevens LA 2006).

The lack of re-absorption of cystatin C in the tubules can lead to a tremendous increase of urinary levels, reaching maxima of several hundred fold baseline levels in both humans and rats (Löfberg H, Gubb AO 1979, Uchida K, Gotoh A 2002). With high tubular protein loads secondary to glomerular damage or other causes of proteinuria, high molecular weight proteins such as albumin compete for common transport mechanisms, decreasing the tubular uptake and increasing excretion of cystatin C into urine. As tubular reabsorption rates of small-molecular proteins are greater than those of high-molecular proteins, low-molecular proteins are more sensitive markers for detection of glomerular alterations and changes of the tubular reabsorption (Thielemans N, Lauwerys R, Bernard A 1994). It has also been shown in humans that specific damage to the reabsorption system in tubular diseases such as polycystic kidney disease or pyelonephritis, as well as in diseases with a high tubular protein load such as the nephrotic syndrome, urinary excretion of cystatin C is increased (Tkaczyk M, Nowicki M, Lukamowicz J 2004, Conti M, Moutereau S, Zater M et al 2006).

vi. Clusterin

Clusterin is known by many names including sulfated glycoprotein 2 (SGP-2), glycoprotein III, testosterone-repressed prostate message 2 (TRPM-2), glycoprotein of 80 kDa, cytolysis inhibitor, complement lysis inhibitor (CLI), apolipoprotein J (ApoJ), and secreted protein 40,40. Both a secreted and a nuclear isoform of clusterin are described. To date, the non-extensively glycosylated 49kDa nuclear isoform is not considered relevant in the context of kidney injury. The secreted isoform is a 76-80 kDa glycosylated protein with extensive post-translational modifications. It is a disulfide-linked heterodimer consisting of a- and ß-subunits which is synthesized in many tissues and found in plasma, serum and CSF. In kidney, it is highly expressed during early stages of renal development. Secreted clusterin has been suggested to play an anti-apoptotic role, to be involved in cell protection (as scavenger for hydrophobic products), lipid recycling, cell aggregation and cell attachment (Rosenberg ME, Silkensen J 1995). Clusterin gene over-expression is induced by different types of kidney injury in glomeruli, tubules and papilla of animals e,g, in rats after nephrectomy (Correa-Rotter et al 1992), renal ischemia-reperfusion (Yoshida et al 2002, and treatment with puromycin (Correa-Rotter et al 1998) or sevoflurane (Kharasch et al 2006) and dogs after treatment with nefiracetam (Tsuchiya et al 2005). In humans, clusterin has been found in the glomeruli of patients with glomerulonephritis. Also in human renal tubular injuries, such as acute and chronic transplant rejection and renal dysplasia, increased immunostainable clusterin was reported (Rosenberg ME, Silkensen J 1995).

vii. Trefoil Factor 3

Trefoil factor 3 (TFF3) is a member of a three mucin-associated peptide family (Suemori et al. 1991). TFF3 was discovered as a robust and novel biomarker of proximal tubule injury in male rats through a large transcriptional effort (profiling and Taqman) on over forty exploratory and Merck developmental compound studies. TFF3 protein has also been detected in epithelial cells from other tissues, and was originally identified as playing a role in maintenance and restitution of epithelial barrier function in the intestines (Suemori, et al.1991). The male rat kidney shows substantial expression of TFF3 which decreases with age (Debata et al. 2007). Immunohistochemical studies have localized TFF3 to collecting duct epithelial cells in normal kidney (Figueroa et al. 2007). It is interesting that this site in the nephron is anatomically downstream of the proximal tubule which is the most common site of renal injury in the rat. Elucidating the robust relationship between TFF3's expression and localization, with collecting duct and proximal tubular injury, will provide further insight into the regulatory mechanisms that occur during acute nephrotoxicity and refine the utility of this novel biomarker.

b. Additional data supporting qualification of proposed biomarkers submitted by the PSTC

i. PSTC Standardized kidney histopathology lexicon

PSTC Standardized kidney histopathology lexicon

Primary Histopathology Process	Secondary Histopathology Lesion	Structural element / Segment
Tubular Cell Degeneration/Necrosis/Apoptosis	Necrosis	No precise localization possible Prox. convoluted tubule (PCT, s1-s2) Thick descending tubule (s3) Loop of Henle Thick ascending tubule Distal convoluted tubule Collecting duct
	Apoptosis	No precise localization possible Prox. convoluted tubule (PCT, s1-s2) Thick descending tubule (s3) Loop of Henle Thick ascending tubule Distal convoluted tubule Collecting duct
Necrosis/Infarction		Cortex Medulla Papilla
Tubular Cell Regeneration	Basophilia	No precise localization possible Prox. convoluted tubule (PCT, s1-s2) Thick descending tubule (s3) Loop of Henle Thick ascending tubule Distal convoluted tubule Collecting duct
	Mitosis increase	No precise localization possible Prox. convoluted tubule (PCT, s1-s2) Thick descending tubule (s3) Loop of Henle Thick ascending tubule Distal convoluted tubule
Tubular Cell Alterations	Hyaline droplet formation	No precise localization possible Prox. convoluted tubule (PCT, s1-s2) Thick descending tubule (s3) Loop of Henle Thick ascending tubule Distal convoluted tubule Collecting duct
	Hypertrophy/Enlargement	No precise localization possible Prox. convoluted tubule (PCT, s1-s2) Thick descending tubule (s3) Loop of Henle Thick ascending tubule Distal convoluted tubule Collecting duct
	Nuclear change	No precise localization possible Prox. convoluted tubule (PCT, s1-s2) Thick descending tubule (s3) Loop of Henle Thick ascending tubule Distal convoluted tubule Collecting duct
	Cellular sloughing	No precise localization possible Prox. convoluted tubule (PCT, s1-s2) Thick descending tubule (s3) Loop of Henle Thick ascending tubule Distal convoluted tubule Collecting duct
	Pigmentation accumulation	No precise localization possible Prox. convoluted tubule (PCT, s1-s2) Thick descending tubule (s3) Thick ascending tubule Distal convoluted tubule Collecting duct
	Vacuolation	No precise localization possible Prox. convoluted tubule (PCT, s1-s2) Thick descending tubule (s3) Loop of Henle Thick ascending tubule Distal convoluted tubule Collecting duct
Tubular Dilatation	Tubular Dilatation	Cortex Medulla Papilla
	Tubular Cystic Dilatation / Tubular Cyst(s)	Cortex Medulla Papilla
Pelvis Dilatation		Pelvis

Primary Histopathology Process	Secondary Histopathology Lesion	Structural element / Segment
Intratubular Casts	Crystalline	No precise localization possible Prox. convoluted tubule (PCT, s1-s2) Thick descending tubule (s3) Loop of Henle Thick ascending tubule Distal convoluted tubule Collecting duct
	Granular	No precise localization possible Prox. convoluted tubule (PCT, s1-s2) Thick descending tubule (s3) Loop of Henle Thick ascending tubule Distal convoluted tubule Collecting duct
	Hyaline (proteinaceous, pigmented)	No precise localization possible Prox. convoluted tubule (PCT, s1-s2) Thick descending tubule (s3) Loop of Henle Thick ascending tubule Distal convoluted tubule Collecting duct
	Leukocytic	No precise localization possible Prox. convoluted tubule (PCT, s1-s2) Thick descending tubule (s3) Loop of Henle Thick ascending tubule Distal convoluted tubule Collecting duct
	Mineralization	No precise localization possible Prox. convoluted tubule (PCT, s1-s2) Thick descending tubule (s3) Loop of Henle Thick ascending tubule Distal convoluted tubule Collecting duct
Inflammation	Interstitial,acute	Cortex Medulla Papilla Pelvis
	Interstitial,chronic	Cortex Medulla Papilla Pelvis
	Acute	Glomerulus
	Chronic	Glomerulus
Fibrosis	Perivascular	Cortex Medulla Papilla
	Interstitial	Cortex Medulla Papilla
	Interstitial Bowman's capsule	Glomerulus
	Fibrosis of glomerulus / Glomerulosclerosis	Glomerulus
Glomerular Alteration	Enlargement of Bowman's space	Glomerulus
	Decrease of Bowman's space	Glomerulus
	Mesangial proliferation/expansion	Glomerulus
	Glomerular Vacuolation	Glomerulus
Edema		Diffuse Cortex Medulla Papilla
Vascular alteration	Vasculitis	Diffuse Cortex Medulla Hilum
	Medial hypertrophy	Diffuse Cortex Medulla
	Necrosis	Diffuse Cortex Medulla Hilum
	Thrombosis-thrombus	Diffuse Cortex Medulla Hilum
Mineralisation-parenchymal		Cortex Cortico-medullary junction Medulla Papilla
Urothelial hypertrophy-hyperplasia		Papilla Pelvis
Juxtaglomerular Apparatus Hypertrophy		Juxtaglomerular
Concentric Lamellar Bodies		Cortex Cortico-medullary junction Medulla Papilla

ii. Summary of individual studies

PSTC detailed summary of Merck and Novartis studies

Merck studies

Compound/Study	Route	Dose	Necropsy day/ (Urine collection days)	Sex/ strain	Number / Group	Notes
Cisplatin	i.p. (5 ml/kg)	0, 0.5, 3.5 or 7 mk	3/8 (3/8)	M/SD	5	1 dose
Gentamicin	i.p. (5 ml/kg)	0, 20, 80, or 240 mkd	3/9/15 (3/9/15)	M/SD	5	240 mkd terminated day 12
Gentamicin-TC (time course)	i.p. (5 ml/kg)	0, 40, or 120 mkd	39 (10/39)	M/SD		10 day dosing with a 29 day recovery period
Carbapenem A-TC (time course)	i.v. (5 ml/kg)	0 or 150 mkd	2/4/8/18 (2/4/8/15)	MF/ SD	10/20	3 day study; up to 15 day recovery period
Carbapenem A	i.v. (5 ml/kg)	0, 75, 150, or 225 mkd	3/9/14 (3/9/14)	M/SD	5	Several animals treated with the 225 mkd dose were early sacrificed
Carbapenem A –TC (8D)	i.p. (5 ml/kg)	0 or 150 mkd	8 (daily)	M/SD	6	No pathology performed
Cyclosporin A	s.c. (5 ml/kg)	0, 6, 30, or 60 mkd	3/9/15 (3/9/15)	M/SD	5	
Thioacetamide	p.o. (5 ml/kg)	0, 50, 100, or 200 mk	2/3 (2/3)	M/SD	5	Liver & Kidney toxin
Hexachlorobutadiene (HCB)	i.p.	0, 7.5, 40, or 100 mkd	4/ 8/15 (4/8/15)	M/SD	4	CRL CRO
Allopurinol	i.p.	0, 6, 30, or 100 mkd	4/ 8/15 (4/8/15)	M/SD	4	CRL CRO
NPAA	p.o.	0, 350, 700, or 1200 mkd	4/ 8/15 (4/8/15)	M/SD	4	CRL CRO
D-Serine	i.p.	0, 75, 250, or 750 mk	4/ 8/15 (4/8/15)	M/SD	4	CRL CRO
Propylenimine	i.p. (2 ml/kg)	0, 11, or 22 ul/kg (day one)	7/21 (7/14/21)	M/HW	5	HW: Wistar Hanover
Adriamycin	i.v. (5 ml/kg)	0, 4 (2 doses during week 1 & 2) 7.5 mk (single dose)	7/14/28 (7/14/28)	M/HW	5	HW: Wistar Hanover
Isoproterenol	i.v. (5 ml/kg)	0, 0.064, 0.25, or 1 mkd	2/3/8 (3/8)	M/SD	8	Heart toxin
Furan	p.o. (10 ml/kg)	0, 4, 40, or 60 mkd	2/4/8 (2/4/8)	M/SD	4	Liver toxin
Genipin	i.p. (1.11-10 ml/kg)	0 or 75 mkd	3 (3)	M/SD	6	Liver toxin
Cerivastatin	p.o. (5 ml/kg)	0 or 1 mkd	9/15 (8/10/12/15)	F/SD	4	Skeletal muscle toxin
CCl_4 or $BrCCl_3$	p.o. (0.75 - 1.25 ml/kg)	0, 0.03, 0.1, 0.3, or 0.03/0.5 mL/kg (CCl_4) 0, 0.03, or 0.1 mL/kg ($BrCCl_3$)	2/4 (2/4)	M/SD	5	Liver toxins
Diuresis: H_20 or 2% NaCl	p.o.	4.5% body weight	4 (1/4)	M/SD	4	No toxicity
Diuresis: Control H_20 or 4 % sucrose	in water	N/A	4 (1/4)	M/SD	4	No toxicity

Novartis studies

Test Compound	Dose Levels Dose Route Dose regimen	Histopathology 3h after dosing on days	Urine collection times 2 – 8 pm, 1st day 8 pm -6 am, 2nd day	Blood/plasma 3 hrs after dosing on Days
		Nephrotoxicants		
Gentamicin sulfate	0, 35, 70, 140 mg/kg IP 1x daily 5 mL/kg	1, 3, 7, 14	2, 3 6, 7 13, 14	1, 3, 7, 14
Vancomycin hydrochloride	0, 70, 140, 210 mg/kg IP 1 daily injection 5 mL/kg	1, 3, 7, 14	2, 3 6, 7 13, 14	1, 3, 7, 14
Doxorubicin chlorhydrate	0, 2.5, 5.0, 7.5 mg/kg IV Once at day 1 5 mL/kg	1, 3, 7, 14	2, 3 6, 7 13, 14	1, 3, 7, 14
Furosemide	0, 45, 90, 180 mg/kg oral gavage 2x daily 5 mL/kg	1, 3, 7, 14	2, 3 6, 7 13, 14	1, 3, 7, 14
Lithium Carbonate	0, 1, 2, 3 mEQu/kg oral gavage 1x daily 5mL/kg	1, 3, 7, 14	2, 3 6, 7 13, 14	1, 3, 7, 14
Cisplatin	0.0.5, 1, 3 mg/kg IP Once at day 1 5mL/kg	1, 3, 7, 14	2, 3 6, 7 13, 14	1, 3, 7, 14
Puromycin dihydrochloride	0, 10, 20, 40 mg/kg IP 1x daily 10mL/kg	3, 7, 14, 22	2, 3 6, 7 13, 14 21, 22	3, 7, 14, 22
Tacrolimus/ FK506	0, 9, 12, 15, IP 1x daily 5mL/kg	3, 7, 14, 21	2, 3 6, 7 13, 14 20, 21	3, 7, 14, 21
		Hepatotoxicants		
ANIT (α-naphtyl-isothiocyanate)	0, 5, 15, 30 mg/kg Oral gavage daily 5mL/kg	1, 3, 7, 14	2, 3 6, 7 13, 14	1, 3, 7, 14
Methapyrilene hydrochloride	0, 15, 30, 60 mg/kg Oral gavage 1x daily 5mL/kg	1, 3, 7, 14	2, 3 6, 7 13, 14	1, 3, 7, 14

FDA studies

Data Reported to C-Path from Dose-Response Studies				
Test Compound	**Dose Levels** **Dose Route** **Dose regimen** **Dose volume**	**Histopathology after dosing**	**Urine collection times (relative to dosing at 0 hr)**	**Blood collected after dosing**
Gentamicin sulfate	0, 25, 50, 100, 150, 200, 400 mg gent/kg SC 1 daily injection for 3 days 2 ml/kg	48 hr after last dose	24-48 hr	48 hr after the last dose
Mercuric chloride, $HgCl_2$	0, 0.25, 0.5, 1 mg Hg/kg IV 1 injection 2 ml/kg	24 hr after single dose	0-24 hr	24 hr after single dose
Chromium, $Na_2Cr_2O_7$	1, 2.5, 5, 10, 20 mg Cr/kg SC 1 injection 2 ml/kg	24 hr after single dose	0-24 hr	24 hr after single dose
Data Reported to C-Path from Time-Course Studies				
Test Compound	**Dose Levels** **Dose Route** **Dose regimen** **Dose volume**	**Histopathology after dosing**	**Urine collection times (relative to dosing at 0 hr)**	**Blood collected after dosing**
Gentamicin sulfate	O, 100, 150 mg gent/kg SC 1 daily injection for 3 days 2 ml/kg	72 hr after last dose	48-72 hr (see Exptl Protocol and Note sections- only time interval with corresponding histopath score)	72 hr after last dose
Mercuric chloride, $HgCl_2$	0, 0.25 mg Hg/kg IV 1 injection 2 ml/kg	72 hr after single dose	48-72 hr (see Exptl Protocol and Note sections)	72 hr after single dose
Chromium, $Na_2Cr_2O_7$	O, 5 mg Cr/kg SC 1 injection 2 ml/kg	72 hr after single dose	48-72 hr (see Exptl Protocol and Note sections)	72 hr after single dose
Test Compound	Dose Levels Dose Route Dose regiment	Histopathology 72 hours after dosing on day	Urine collection times (relative to dosing at 0 hr)	Blood collected 72 hr after dosing on day
Cisplatin	0, 1, 3, or 6 mg/kg i.p. Single dose	1	-12 hr to 0 hr 0 to 8 hr 8 hr to 24h 24hr to 48 hr 48 hr to 72 hr	1

iii. Summary table of ROC analysis using an exclusion model

Initial ROC analysis by Merck – Merck versus SRI histopathology - Exclusion model – Merck VXDS03 – 9/20/07

Merck Histopathology

Marker	AUC	SE	log2 cutoff	fold cutoff	FPR	TPR	npos	nneg
Kim-1.ucr	1.000	NA	2.53	5.77	0.043	1.00	40	23
SCr(Kim-1) mg/dL	0.946	0.028	0.32	1.25	0.043	0.78	40	23
BUN(Kim-1) mg/dL	0.900	0.038	0.40	1.32	0.043	0.73	40	23
Albumin.ucr	0.987	0.013	0.96	1.94	0.042	0.93	44	24
S.Cr mg/dL	0.950	0.025	0.32	1.25	0.042	0.80	44	24
BUN mg/dL	0.901	0.037	0.40	1.32	0.042	0.73	44	24

Subsequent Merck analysis of SRI data– VXDS03 and Supplement 11/08/07

SRI Histopathology

Marker	AUC	SE	log2 cutoff	fold cutoff	FPR	TPR	npos	nneg
Kim-1.ucr	0.978	0.017	2.53	5.77	0.048	0.93	41	21
SCr(Kim-1) mg/dL	0.909	0.037	0.32	1.25	0.048	0.76	41	21
BUN(Kim-1) mg/dL	0.904	0.038	0.40	1.32	0.048	0.71	41	21
Albumin.ucr	0.958	0.023	0.90	1.87	0.045	0.87	46	22
S.Cr mg/dL	0.899	0.037	0.32	1.25	0.045	0.76	46	22
BUN mg/dL	0.906	0.035	0.40	1.32	0.045	0.72	46	22

BQRT Maximum Composite Injury Score (SRI assessment) (includes interstitial inflammation)

Marker	AUC	SE	log2 cutoff	fold cutoff	FPR	TPR	npos	nneg
Kim-1.ucr	0.975	0.017	0.84	1.79	0.050	0.91	46	20
SCr(Kim-1) mg/dL	0.865	0.044	0.32	1.25	0.050	0.67	46	20
BUN(Kim-1) mg/dL	0.853	0.046	0.40	1.32	0.050	0.63	46	20
Albumin.ucr	0.960	0.021	0.90	1.87	0.048	0.87	53	21
S.Cr mg/dL	0.844	0.045	0.32	1.25	0.048	0.66	53	21
BUN mg/dL	0.847	0.045	0.40	1.32	0.048	0.62	53	21

AUC: area under the curve, se: standard error, fold.cutoff: fold change, FPR: specificity (~0.05 error), TPR: sensitivity, npos: number positive samples by histopathology, nneg: number negative samples by histopathology. Maximum composite score was used for histopathology. Note there were 11 samples for which KIM-1 was not measured. BUN(KIM-1) and S.Cr(KIM-1) are values for an analysis of BUN and S.Cr with these 11 samples

iv. Statistical analysis using an exclusion model

Initial ROC statistical analysis by Merck - Exclusion model – Merck VXDS04 – 10/01/07

Comparison	npos	nneg	marker AUC	CTL AUC	Diff AUC	SE	pvalue	p.Holm	q.BH
Kim-1.ucr-SCr	77	46	0.997	0.850	0.147	0.032	**0.00001**		
Kim-1.ucr-BUN	77	46	0.997	0.902	0.095	0.027	**0.00038**		
Alb.Combined.ucr-SCr	246	224	0.901	0.766	0.136	0.022	**9.99E-10**		
Alb.Combined.ucr-BUN	246	224	0.901	0.822	0.079	0.021	**0.00015**		
nTFF3.ucr-SCr	134	105	0.900	0.891	0.009	0.024	0.70557	0.70557	0.70557
nTFF3.ucr-BUN	134	105	0.900	0.917	-0.017	0.029	0.56538	1.00000	0.56538
nTFF3.ex-SCr	111	106	0.917	0.894	0.023	0.021	0.27375	0.54750	0.41062
nTFF3.ex-BUN	111	106	0.917	0.899	0.019	0.032	0.54993	1.00000	0.56538
nTFF3.conc-SCr	135	117	0.931	0.896	0.035	0.020	0.07381	0.22143	0.22143
nTFF3.conc-BUN	135	117	0.931	0.901	0.030	0.027	0.26749	0.80246	0.56538

Comparison: biomarker compared to control (BUN or SCr), npos: number of samples with positive histomorphologic change (Maximum Composite), nneg: number of samples with negative histomóphologic change (Maximum Composite), biomarker AUC: AUC from ROC curve for putative biomarker, CTL AUC: AUC from ROC curve for SCr or BUN, Diff AUC: biomarker AUC - CTL AUC, SE: standard error of Diff AUC from DeLong analysis, pvalue: p-value from DeLong test, p.Holm: adjusted p-value using Holm procedure for multiplicity in testing TFF3 normalized or ng/mL three ways, q.BH: Benjamini and Hochberg false discovery rate for multiplicity in testing TFF3 normalized or ng/mL three ways. Note that p value fonts at p<.05 are bolded and underlined.

PSTC Statistical analysis – Novartis Exclusion (VXDS02)

Exclusion Model: Tubular Damage S1-S3

Comparison	npos	nneg	marker AUC	CTL AUC	Diff AUC	SE	pvalue	p.Holm	q.BH
Clusterin-SCr	132	289	0.877	0.732	0.145	0.038	**1.16E-04**		
Clusterin-BUN	132	289	0.877	0.788	0.089	0.033	**7.89E-03**		
Kim1-SCr	132	283	0.910	0.732	0.178	0.035	**3.02E-07**		
Kim1-BUN	132	283	0.910	0.789	0.121	0.032	**1.20E-04**		

Exclusion Model: Glomerular Alterations / Damage

Comparison	npos	nneg	marker AUC	CTL AUC	Diff AUC	SE	pvalue	p.Holm	q.BH
Cystatin-SCr	40	291	0.915	0.532	0.382	0.079	**1.47E-06**	**2.95E-06**	**2.95E-06**
Cystatin-BUN	40	291	0.915	0.800	0.115	0.053	**3.13E-02**	6.27E-02	6.27E-02
B2Microglobulin-SCr	40	291	0.889	0.532	0.357	0.083	**1.72E-05**	**3.44E-05**	**3.44E-05**
B2Microglobulin-BUN	40	291	0.889	0.800	0.089	0.055	1.09E-01	2.17E-01	2.17E-01
Total Protein-SCr	40	291	0.858	0.532	0.326	0.074	**1.12E-05**	**2.23E-05**	**2.23E-05**
Total Protein-BUN	40	291	0.858	0.800	0.058	0.068	3.98E-01	7.96E-01	7.96E-01

Testing results for statistically comparing the significance of differences of AUCs of the ROC analyses between markers and BUN or SCr for the exclusion analysis in the context of glomerular alterations / damage. The first column represents the marker and standard being compared, the second column the number of diseased samples, the third column the number of control samples, the fourth and fifth column the AUCs for the markers and the standards, the sixth column represents the difference of AUCs, the seventh column the standard error of Diff AUC from DeLong analysis, the eights column the p-value from DeLong test, the ninth column the adjusted p-value using Holm procedure for multiplicity in testing the glomerular markers also for tubular damage and the tenth column the Benjamini and Hochberg false discovery rate for multiplicity in testing glomerular markers also for tubular damage. Note: p<0.05 are in bolded font.

v. Statistical analysis to test whether a biomarker adds value to sCr and BUN

Merck statistical analysis: Logistic regression likelihood ratio test of whether marker adds value to sCr and BUN – Inclusion ROC analysis

Marker	npos	nneg	Like.Stat (w marker)	Like.Stat (w/o marker)	pvalue	p.Holm	q.BH
KIM1.ucr	77	101	175.9	116.9	**1.63E-14**		
Alb.Combined.ucr	251	449	301.5	199.8	**6.58E-24**		
nTFF3.ucr	135	175	219.0	200.0	**1.31E-05**	**2.63E-05**	**1.97E-05**
nTFF3.ex	111	176	192.5	176.9	**8.00E-05**	**8.00E-05**	**8.00E-05**
nTFF3.conc	136	187	241.4	203.6	**7.79E-10**	**2.34E-09**	**2.34E-09**

Marker putative marker, npos: number of samples with positive histopathology (Maximum Composite), nneg: number of samples with no observed (negative) histopathology (Maximum Composite), Like.Stat (w marker): Likelihood statistic (-2 *log nikelihood(intercept)/ likelihood(model) n for logistic model log(p/(l-p))= intercept + marker + SCr + BUN + SCr*BUN, Like.Stat (w/o marker): Likelihood statistic for logistic model log(p/(l-p))= intercept + SCr + BUN + SCr*BUN,i.e. without the marker, pvalue: pvalue for likelihood ratio test, p.Holm: adjusted p-value using Holm procedure for multiplicity in testing TFF3 normalized or ng/mL (three ways), q.BH: Benjamini and Hochberg false discovery rate for multiplicity in testing TFF3 normalized three ways. Note that pvalue, p.Holm, and q.BH are in bold and underlined. Albumin (Alb) Combined=ITA+EIA.

Novartis statistical analysis:

Table 12: Testing results (inclusion analysis) of logistic regression likelihood ratio test of whether a marker adds value to serum Creatinine and BUN in the context of tubular damage (S1-S3 or non-localizable).

Marker	npos	nneg	Like.Stat (w marker)	Like.Stat (w/o marker)	pvalue	p.Holm	q.BH
Urinary KIM1	135	595	362.4	173.4	**2.73E-43**		
Urinary Clusterin	135	604	310.9	175.2	**1.18E-31**		

Table 13: Testing results (inclusion analysis) of logistic regression likelihood ratio test of whether a marker adds value to serum Creatinine and BUN in the context of glomerular alterations / damage.

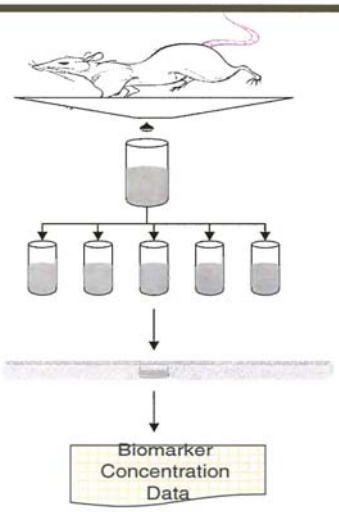
Marker	npos	nneg	Like.Stat (w marker)	Like.Stat (w/o marker)	pvalue	p.Holm	q.BH
Urinary Cystatin C	41	698	167.2	107.3	**4.97E-15**	**9.94E-15**	**9.94E-15**
Urinary B2 Microglobulin	41	698	174.9	107.3	**9.98E-17**	**2.00E-16**	**2.00E-16**
Urinary Total Protein	41	698	168.0	107.3	**3.29E-15**	**6.58E-15**	**6.58E-15**

The first column shows the putative marker, the second column the number of diseased samples, The third column the number of control samples, the fourth column the likelihood statistic (-2*log{likelihood[intercept]/ likelihood[model]}) for logistic model log(p/(1-p))= intercept + marker + SCr + BUN + SCr*BUN, Like.Stat the fifth column the likelihood statistic for logistic model log(p/(1-p))= intercept + SCr + BUN + SCr*BUN, i.e. without the marker, the sixth column the pvalue for likelihood ratio test, the sevenths column the adjusted p-value using Holm procedure for multiplicity in testing the glomerular markers also for tubular toxicity, and the eights column the Benjamini and Hochberg false discovery rate for multiplicity in testing the glomerular marker also for tubular toxicity. Note p<0.05 are in bolded font.

vi. PSTC Flow chart of sample handling

PSTC Flow-chart for sample handling

Urine Samples and Immunoassays



Urine samples were collected on dry ice during last 16 hours prior to sacrifice.

Urine samples were thawed, placed on ice, centrifuged to remove precipitate. Volume measured. Several aliquots made, stored at -70°C.

One urine aliquot was thawed and iced. A portion was added to immunoassay (ELISA or Luminex format). [1, 2]

Instrument-output raw data & standard curve archived. Biomarker concentration determined from standard curve.

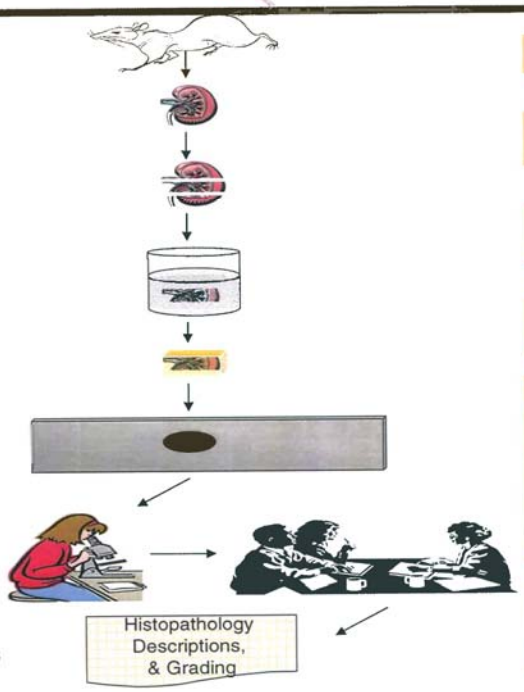
Novartis & FDA assays were blinded

Notes:

1 For several studies, samples were thawed an additional time and re-assayed.

2 For the TFF3 assay, 50uL of urine were cleaned up on a microcon ultrafiltration unit.

Histopathology



Kidney removed at necropsy

Poles removed

Central section fixed in neutral-buffered formalin

Fixed tissue processed, paraffin embedded

Sliced by microtome

Sections mounted on slides

Sections stained with Hematoxylin & Eosin

QC by Technical Staff

Slides assessed by Qualified Pathologists

Peer-review by senior Pathologists

vii. PSTC discussion of "prodromal" relative to exclusion and inclusion analysis

The following discussion is from the combined PSTC report submitted 09/05/07:

In the original VXDS submission, the term prodromal was used to describe biomarker changes that preceded observed histopathologic changes in time or at lower dose. Since these changes could be either false positives or prodromal events, the term was changed to "potentially prodromal", and data analysis was performed using two approaches. Firstly, the potentially prodromal data were removed from "exclusion analysis" to avoid potential bias from this interpretation. As the PSTC members consider this analysis as the most appropriate way to analyze the data, all claims emphasize this analysis. Secondly, the data were analyzed using "inclusion analysis" presuming that these data indicate false positive values. We feel that by excluding these data in one analysis, and presuming these data to be false positives in the second analysis, presents a conservative and balanced view.

The PSTC members Merck and Novartis were asked by FDA and EMEA to elaborate on their use of the term "prodromal" from the original June 2007 submission as a speculative interpretation of biomarker signals observed at some low doses and/or early time points in some animals in the absence of a observed histopathologic change~ The PSTC members Merck and Novartis would like to clarify that at this stage, do not seek to make any regulatory claims that the biomarkers are prodromal markers of injury. We were merely discussing trends seen in the data such as 1), significant treatment-related increases in biomarker concentrations are observed in the absence of histopathologic change; 2) and at later time points in animals with continued treatment, histopathologic change is seen with correspondingly higher fold-changes in biomarkers.

Alternatively, such results might be described as "false positive". However this raises the possibility of classifying a true prodromal signal as incorrect (i.e., preceding histopathologic change). Histologic examination of any additional remnant kidney section is not considered likely to conclusively resolve this issue. The experimental designs in the submitted VXDS studies used doses of nephrotoxic ants intended to produce mild lesions in order to test whether markers correlated with histopathologic change. A section of one kidney was inspected microscopically, while a renal focal lesion elsewhere might contribute to the appearance of biomarker in the urine from an overnight collection.

More experimentation using samples from additional specially-designed studies is envisioned to more fully understand whether biomarkers are actually prodromal, or whether certain samples might be considered false negatives for histopathologic change, or whether the biomarkers may be

yielding false positive values. The PSTC therefore makes no regulatory claims that the biomarkers showed prodromal properties in this submission.

Since it is not clear whether certain biomarker signals are prodromal or false positives; then, for purposes of the main type of ROC analysis, they were excluded since they might best be considered neither clearly positive nor clearly negative. The ROC exclusion analysis takes the conservative approach of excluding all samples from treated animals that did not present with histopathologic change. This avoids potentially incorrect conclusions in cases where the biomarker was positive and histopathology was not observed. In order to present fully balanced viewpoints we also ran the additional inclusion ROC analyses, which include all of these samples as well, treating them as false positives. Importantly, the statistical testing of the claims showed a high consistency regardless of whether the inclusion or exclusion ROC analyses were used.

Since the PSTC makes no regulatory claims that the biomarkers have prodromal properties, the term "prodromal" as a description of biomarker performance was revised to "potentially prodromal" in the current submission.

viii. Novartis ROC analysis of different histopathology lesions

Novartis ROC analysis of different histopathology lesions

Primary Histopathology Process	Secondary Lesion	Structural Element / Segment	Serum Creatinine	BUN	Urinary Creatinine	Urinary Protein	Ur. B2-Microglobulin	Urinary Cystatin C	Urinary Clusterin	Urinary Kim-1	N° Controls	N° Diseased
Tubular Cell Degeneration/Necrosis/Apoptosis	Necrosis	No precise localization possible	0.92	0.91	0.61	0.65	0.71	0.56	0.93	0.93	291	36
		Prox. convoluted tubule (PCT, s1-s2)	0.78	0.75	0.53	0.73	0.71	0.81	0.90	0.95	292	51
		Thick descending tubule (s3)	0.72	0.81	0.54	0.58	0.51	0.61	0.86	0.91	292	100
		Thick ascending tubule	0.74	0.83	0.68	0.51	0.53	0.67	0.96	0.93	292	41
		Distal convoluted tubule	0.55	0.61	0.75	0.88	0.99	0.99	0.83	0.70	292	2
		Collecting duct	0.98	0.85	0.59	0.70	0.68	0.54	0.93	0.88	292	18
	Apoptosis	No precise localization possible	0.52	0.59	0.61	0.67	0.65	0.74	0.85	0.71	290	6
		Prox. convoluted tubule (PCT, s1-s2)	0.79	0.67	0.61	0.71	0.84	0.85	0.91	0.86	291	17
		Thick descending tubule (s3)	0.92	0.95	0.63	0.81	0.56	0.60	0.98	0.88	291	5
		Thick ascending tubule	0.89	1.00	0.71	0.99	0.64	0.62	1.00	1.00	292	1
		Collecting duct	0.81	0.60	0.57	0.55	0.52	0.72	0.59	0.57	292	11
Tubular Cell Regeneration	Basophilia	No precise localization possible	0.68	0.65	0.56	0.56	0.58	0.54	0.72	0.72	258	108
		Prox. convoluted tubule (PCT, s1-s2)	0.65	0.73	0.53	0.68	0.67	0.75	0.81	0.86	283	86
		Thick descending tubule (s3)	0.74	0.82	0.54	0.58	0.50	0.66	0.89	0.93	290	97
		Thick ascending tubule	0.75	0.85	0.58	0.52	0.56	0.57	0.84	0.88	290	76
		Distal convoluted tubule	0.83	0.76	0.53	0.55	0.53	0.62	0.67	0.65	290	18
		Collecting duct	0.86	0.60	0.50	0.57	0.55	0.79	0.82	0.76	291	28
	Mitosis increase	No precise localization possible	0.99	1.00	0.67	0.70	0.57	0.68	1.00	1.00	292	2
		Prox. convoluted tubule (PCT, s1-s2)	0.72	0.79	0.51	0.89	0.84	0.94	0.97	0.95	292	28
		Thick descending tubule (s3)	0.66	0.80	0.52	0.78	0.68	0.80	0.90	0.98	292	34
		Thick ascending tubule	0.60	0.85	0.65	0.73	0.73	0.78	0.93	1.00	292	13
		Distal convoluted tubule	0.98	0.70	0.65	0.66	0.70	0.91	0.99	0.84	292	6
Tubular Cell Alterations	Hyaline droplet formation	No precise localization possible	0.82	0.83	0.75	0.78	0.77	0.54	0.92	0.75	292	1
		Prox. convoluted tubule (PCT, s1-s2)	0.55	0.65	0.59	0.76	0.72	0.72	0.79	0.74	283	57
		Thick descending tubule (s3)	0.69	0.71	0.68	0.97	0.93	0.92	0.90	0.80	292	23
		Thick ascending tubule	0.93	0.64	0.84	1.00	1.00	1.00	1.00	0.91	292	11
	Hypertrophy/Enlargement	No precise localization possible	0.99	1.00	0.77	0.95	0.91	0.70	1.00	1.00	292	2
		Prox. convoluted tubule (PCT, s1-s2)	0.59	0.81	0.71	0.77	0.71	0.79	0.97	0.87	292	14
		Thick descending tubule (s3)	0.69	0.74	0.53	0.79	0.72	0.82	0.95	0.98	292	36
		Loop of Henle	0.55	0.67	0.56	0.71	0.55	0.72	0.92	0.99	292	27
		Thick ascending tubule	0.66	0.98	0.78	0.81	0.84	0.90	1.00	1.00	292	5
		Distal convoluted tubule	0.99	0.97	0.59	0.75	0.70	0.62	1.00	0.98	292	7
		Collecting duct	0.79	0.69	0.56	0.66	0.59	0.78	0.59	0.53	292	28
	Cellular sloughing	No precise localization possible	0.94	0.93	0.64	0.65	0.69	0.58	1.00	1.00	292	19
		Prox. convoluted tubule (PCT, s1-s2)	0.95	0.85	0.56	0.92	0.86	0.97	0.99	1.00	292	22
		Thick descending tubule (s3)	0.83	0.82	0.57	0.74	0.68	0.79	0.91	0.97	292	31
		Thick ascending tubule	0.78	0.97	0.56	0.62	0.61	0.51	0.79	0.93	292	4
		Distal convoluted tubule	1.00	1.00	0.80	1.00	1.00	1.00	1.00	1.00	292	1
	Vacuolation	Prox. convoluted tubule (PCT, s1-s2)	0.63	0.71	0.54	0.69	0.67	0.62	0.81	0.77	289	33
		Thick descending tubule (s3)	1.00	0.61	0.85	1.00	1.00	1.00	1.00	0.96	292	8
		Thick ascending tubule	1.00	0.95	0.92	1.00	1.00	1.00	1.00	1.00	292	2
		Collecting duct	0.95	0.70	0.66	0.56	0.65	0.74	0.86	0.86	292	7
	Granular	Thick descending tubule (s3)	0.88	1.00	0.80	1.00	1.00	1.00	1.00	0.99	292	4
Intratubular Casts	Hyaline (proteinaceous, pigmented)	No precise localization possible	0.83	0.87	0.57	0.52	0.59	0.59	0.85	0.84	290	42
		Prox. convoluted tubule (PCT, s1-s2)	0.51	0.78	0.60	0.91	0.90	0.89	0.91	0.81	291	36
		Thick descending tubule (s3)	0.53	0.80	0.62	0.89	0.87	0.88	0.93	0.82	291	35
		Loop of Henle	0.58	0.77	0.56	0.88	0.83	0.87	0.94	0.84	292	46
		Thick ascending tubule	0.65	0.80	0.59	0.78	0.71	0.78	0.93	0.87	292	53
		Distal convoluted tubule	0.87	0.90	0.60	0.76	0.78	0.86	0.98	0.92	292	16
		Collecting duct	0.81	0.86	0.62	0.79	0.79	0.90	0.88	0.82	291	19
	Leukocytic	No precise localization possible	0.94	0.99	0.68	0.66	0.83	0.63	1.00	1.00	292	12
		Thick descending tubule (s3)	0.96	0.92	0.77	0.82	0.72	0.79	0.98	1.00	292	1
	Mineralization	Thick descending tubule (s3)	0.90	0.84	0.50	0.65	0.82	0.85	0.63	0.83	292	20
		Loop of Henle	0.85	0.79	0.84	0.95	0.90	0.78	0.64	0.54	292	2
		Thick ascending tubule	0.88	0.74	0.59	0.74	0.81	0.89	0.52	0.81	292	10
Tubular Dilatation	Tubular Dilatation	Cortex	0.82	0.78	0.54	0.64	0.59	0.77	0.89	0.90	288	80
		Medulla	0.81	0.76	0.54	0.67	0.61	0.79	0.88	0.88	288	79
		Papilla	0.71	0.60	0.55	0.76	0.66	0.87	0.77	0.80	292	21
	Tubular Cystic Dilatation / Tubular Cyst(s)	Cortex	0.77	0.60	0.58	0.54	0.95	0.88	0.63	0.97	292	1
		Medulla	0.68	0.60	0.57	0.66	0.68	0.53	0.84	0.64	291	3
Glomerular Alteration	Mesangial proliferation/expan	Glomerulus	0.53	0.80	0.56	0.86	0.89	0.91	0.83	0.76	292	40
	Glomerular Vacuolation	Glomerulus	0.82	0.96	0.60	0.83	0.95	0.95	0.83	0.74	292	14
Juxtaglomerular Apparatus Hypertrophy		Juxtaglomerular	0.75	0.62	0.56	0.58	0.76	0.86	0.57	0.54	292	30
Pelvis Dilatation		Pelvis	0.62	0.53	0.59	0.64	0.56	0.54	0.55	0.51	278	26
Inflammation	Interstitial,acute	Pelvis	0.70	0.76	0.51	0.55	0.72	0.74	0.88	0.73	291	1
	Interstitial,chronic	Cortex	0.71	0.67	0.52	0.55	0.51	0.58	0.72	0.76	270	78
		Medulla	0.63	0.65	0.59	0.84	0.70	0.84	0.89	0.89	292	9
		Pelvis	0.70	0.76	0.51	0.55	0.72	0.74	0.88	0.73	291	1
Fibrosis	Interstitial	Cortex	0.73	0.78	0.51	0.84	0.67	0.76	0.96	0.70	291	4
		Medulla	0.51	0.56	0.54	0.88	0.63	0.72	0.95	0.62	290	2
	Interstitial Bowman's capsule	Glomerulus	0.83	0.95	0.59	0.90	0.95	0.95	0.85	0.73	291	15
Concentric Lamellar Bodies		Cortex	0.54	0.55	0.54	0.57	0.55	0.66	0.62	0.57	290	10
		Medulla	0.86	0.68	0.74	0.58	0.65	0.68	0.67	0.75	291	1
		Papilla	0.82	0.83	0.75	0.78	0.77	0.54	0.92	0.75	292	1
Urothelial hypertrophy-hyperplasia		Papilla	0.96	0.71	0.55	0.98	0.61	0.55	0.81	0.95	292	1
Any kidney lesion			0.64	0.64	0.53	0.55	0.51	0.56	0.65	0.67	209	373
Any tubular lesion			0.65	0.65	0.53	0.56	0.51	0.58	0.70	0.70	234	313
Glomerular Alteration / Damage (Mesang. prolif., Interstit. Bowmans Capsule fibrosis, Glom. vacuolation)			0.52	0.80	0.56	0.86	0.89	0.91	0.83	0.75	291	41
Tubular Necrosis, S1-S2 + nl			0.83	0.81	0.57	0.60	0.55	0.73	0.93	0.95	291	78
Tubular Degeneration/Necrosis/Apoptosis/Cell Sloughing/Basophilia/Mitosis/Dilatation, S1-S3 + nl			0.67	0.67	0.54	0.55	0.50	0.59	0.74	0.76	246	249
Tubular Degeneration/Necrosis/Apoptosis/Cell Sloughing, S1-S2 + nl			0.81	0.80	0.56	0.61	0.58	0.74	0.93	0.93	289	90
Tubular Degeneration/Necrosis/Apoptosis/Cell Sloughing, S1-S3 + nl			0.73	0.78	0.55	0.59	0.53	0.66	0.87	0.90	289	135

Matrix of results of ROC analyses for different clinical parameters and biomarkers. The first 3 columns show the pathological processes and lesions. Columns 4 to 11 show the area under curves (AUCs), for the different markers, whereby orange represents AUCs>0.9, bright yellow AUCs>0.8 and light yellow AUCs>0.7. Column 12 shows the number of controls used for the corresponding analysis and column 13 shows the number of animals with the corresponding pathology, whereby pathologies with less than 20 cases are colored by gray.

ix. Additional PSTC information concerning histopathology practices

In June 2008 after multiple discussions between the PSTC and the BQRT, the PSTC provided the following additional descriptions of the histopathology practices.

Following identification of a test article-related change, Merck and Novartis pathologists performed blinded reevaluation to ensure consistent determination of incidence, severity, and dose relationship, as needed. Expected levels of background histomorphologic features (defined as subtle histomorphological variations commonly observed in age-matched control animals within the spectrum of morphologies defined as "not remarkable") were established from careful evaluation of concurrent controls and were generally not scored in either control or treated animals. The procedure with the Novartis studies was slightly modified because these were conducted at a CRO. A first histological evaluation was conducted by the CRO pathologists prior to sending slides to Novartis. CRO pathologists first conducted a fully unblinded analyses followed by a fully blinded analysis and peer evaluation. When the slides were received at Novartis, a fully blinded check on 30% of the slides was conducted. Any major discrepancies were resolved through communication between pathologists from the two sites.

Later, the additional clarification was provided.

When a targeted blinded reevaluation was performed [at Merck], *the pathologist and the peer pathologist and PWG when necessary were blinded to the individual animal treatment group assignment. Generally what they did when there were questionable findings at the 0 to 1 histopathology grade boundary, was to combine all the questionable treatment group slides with the control group slides and read them blinded to the dose group assignment. So at that point there is no knowledge of individual animal chemistry data linked to the slides. Throughout the entire process the pathologists were blinded to all new biomarker data.*

The initial fully unblinded analysis ……includes knowledge of treatment group, as well as access to other study data including traditional chemistry results, but not to any of the new biomarker data results. Whether or not pathologists actually looked at clinical chemistry data is unlikely…..., but it is unclear and we [Merck] *cannot state definitively whether or not any of them ever looked at the traditional chemistry markers.*

7. References

Amin RP, Vickers AE, Sistare F, Thompson KL, Roman RJ, Lawton M, Kramer J, Hamadeh HK, Collins J, Grissom S, Bennett L, Tucker CJ, Wild S, Kind C, Oreffo V, Davis JW 2nd, Curtiss S, Naciff JM, Cunningham M, Tennant R, Stevens J, Car B, Bertram TA, Afshari CA. 2004. Identification of putative gene based markers of renal toxicity. Environ Health Perspect. 112(4): 465-79.

Aoyagi T, Hayakawa K, Miyaji K, et al. 2003. Cadmium nephrotoxicity and evacuation from the body in a rat modeled subchronic intoxication. In J Urol; 10: 332-338.

Bakris GL. 2004. Implications of albuminuria on kidney disease progression. J Clin Hypertens (Greenwich). 11 (S3): 18-22.

Bernard A, Viau C, Ouled A, et al. 1986. Effect of Gentamicin on the Renal Uptake of Endogenous and Exogenous Proteins in Conscious Rats. Tox Appl Pharm; 84: 431-438.

Berndt WO. 1976. Renal function tests: what do they mean? A review of renal anatomy, biochemistry, and physiology.Environ Health Perspect. 15: 55-71.

Branten AJW, Buf-Vereeijken PW, Klasen Is et al. 2005. Urinary Excretion of β2 Microglobulin and IgG Predict Prognosis in Idiopathic Membranous Nephropathy: A Validation Study. J Am Soc Nephrol; 16: 169-174.

Christensen EI, Birn H. 2002. Megalin and cubilin: multifunctional endocytic receptors. Nat Rev Mol Cell Biol. 3(4): 256-66.

Compton WD, Osicka T. 2005. Detection of urinary albumin. Adv in Chronic Kid Dis; 12(2): 170-186.

Conti M, Moutereau S, Zater M et al. 2006. Urinary cystatin C as a specific marker of tubular dysfunction. Clin Chem Lab Med; 44(3): 288-291.

Correa-Rotter R, Ibarra-Rubio ME, Schwochau G, et al. 1998. Induction of Clusterin in Tubules of Nephrotic Rats. J Am Soc Nephrol; 9(1): 33-37.

Correa-Rotter R, Hostetter TH, Manivel JC. 1992. Intrarenal distribution of clusterin following reduction of renal mass. Kid Int; 41: 938-950.

D'Amico G, Bazzi C. 2003. Pathophysiology of proteinuria. Kidney International, 63: 809–825.

Darling IM, Morris, ME. 1991. Evaluation of "true" creatinine clearance in rats reveals extensive renal secretion, Pharm. Res., 8: 1318.

Davis JW 2nd, Goodsaid FM, Bral CM, Obert LA, Mandakas G, Garner CE 2nd, Collins ND, Smith RJ, Rosenblum IY. 2004. Quantitative gene expression analysis in a nonhuman primate model of antibiotic-induced nephrotoxicity.Toxicol Appl Pharmacol. 200(1): 16-26.

Debata PR, Panda H and Supakar PC. 2007. Altered expression of trefoil factor 3 and cathepsin L gene in rat kidney during aging, Biogerontology 8 (1): 25-30.

Eknoyan G. 2003. On testing for proteinuria: time for a methodical approach. Cleve Clin J Med.70(6): 493, 496-7, 501.

Figueroa DJ, Jiang, Y, Muniappa, N et al. 2007. Immunohistochemical localization of Trefoil Factor 3 (TFF3) in the kidney of normal and nephrotoxin-treated rats. Unpublished manuscript.

Gatanaga H, Tachikawa N, Kikuchi Y. 2006. Urinary β2 Microglobulin as a Possible Sensitive Marker for Renal Injury Caused by Tenofovir Disoproxil Fumarate. Aids Res Hum Retrovir; 22(8):744-748.

Goodsaid F, Frueh FW. 2007a. Implementing the U.S. FDA guidance on pharmacogenomic data submissions. Environ Mol Mutagen. 48(5): 354-8.

Goodsaid F, Frueh F. 2007b. Biomarker qualification pilot process at the US Food and Drug Administration. AAPS J. 9(1): E105-8.

Guder WG, Hofmann W. 1992. Markers for the diagnosis and monitoring of renal tubular lesions.Clin Nephrol. 38(S1): S3-7

Han WK, Bailly V, Abichandani R, Thadhani R, Bonventre JV. 2002. Kidney Injury Molecule-1 (KIM-1): a novel biomarker for human renal proximal tubule injury. Kidney Int. 62(1): 237-44

Hanley BP. 2007a. Variance in multiplex suspension array assays: intraplex method improves reliability. Theor Biol Med Model. 4: 32.

Hanley BP, Xing L, Cheng RH. 2007b. Variance in multiplex suspension array assays: microsphere size variation impact. Theor Biol Med Model. 4: 31.

Hanley BP. 2007c. Variance in multiplex suspension array assays: carryover of microspheres between sample wells. J Negat Results Biomed. 6:6.

Haas M, Kluppel AC, Moolenaar F, Meijer DK, de Jong PE, de Zeeuw D. 1997. Urine collection in the freely moving rat: reliability for measurement of short-term renal effects. J Pharmacol Toxicol Methods. 38(1):47-51.

Harvey AM, Malvin RL. 1965. Comparison of creatinine and inulin clearances in male and female rats. Am J Physiol. 209(4):849-52.

Iannone MA, Consler. 2006. TG.Effect of microsphere binding site density on the apparent affinity of an interaction partner.Cytometry A. 69(5):374-83.

Ichimura T, Bonventre JV, Bailly V, Wei H, Hession CH, Cate RL, Sanicola M. 1998. Kidney Injury Molecule-1 (KIM-1), a putative epithelial cell adhesion molecule containing a novel immunoglobulin domain, is up-regulated in renal cells after injury. J. Biol Chem. 273(7), 4135–4142.

Ichimura T, Hung CC,1 Yang SA, Stevens JL, Bonventre JV. 2004. Kidney injury molecule-1: a tissue and urinary biomarker for nephrotoxicant-induced renal injury. Am J Physiol Renal Physiol. 286: F552–F563.

ICH M3 [International Conference on Harmonization of Technical Requirements for Registration of Pharmaceuticals for Human Use]. 2000.

ICH S7A [International Conference on Harmonization of Safety Pharmacology Studies for Human Pharmaceuticals]. 2000.

Kaye WA, Griffiths WC, Camara PD, et al. 1981. The Significance of β-2 Microglobulinuria Associated with Gentamicin Therapy, Ann Clin Lab Sci; 11(6): 530-537.

Kharasch ED, Schroeder JL, Bammler T, et al. 2006. Gene Expression Profiling of Nephrotoxicity from the Sevoflurane Degradation Product Fluoromethyl-2,2-difluoro-1-(trifluoromethyl)vinyl Ether ("Compound A") in Rats. Toxicol Sci 90: 419-431.

Lane JT. 2004. Microalbuminuria as a marker of cardiovascular and renal risk in type 2 diabetes mellitus: a temporal perspective. Am J Physiol Renal Physiol 286: F442–F450.

Lee JW, Devanarayan V, Barrett YC, Weiner R, Allinson J, Fountain S, Keller S, Weinryb I, Green M, Duan L, Rogers JA, Millham R, O'Brien PJ, Sailstad J, Khan M, Ray C, Wagner JA. 2006. Fit-for-purpose method development and validation for successful biomarker measurement. Pharm Res. 23(2): 312-28.

Lee JW, Weiner RS, Sailstad JM, Bowsher RR, Knuth DW, O'Brien PJ, Fourcroy JL, Dixit R, Pandite L, Pietrusko RG, Soares HD, Quarmby V, Vesterqvist OL, Potter DM, Witliff JL, Fritche HA, O'Leary T, Perlee L, Kadam S, Wagner JA. 2005. Method validation and measurement of biomarkers in nonclinical and clinical samples in drug development: a conference report. Pharm Res. 22(4): 499-511.

Liangos O, Perianayagam MC, Vaidya VS, Han WK, Wald R, Tighiouart H, MacKinnon RW, Li L, Balakrishnan VS, Pereira BJG, Bonventre JV, Jaber BL. 2007. Urinary N-Acetyl-β-(D)-Glucosaminidase Activity and Kidney Injury Molecule-1 Level Are Associated with Adverse Outcomes in Acute Renal Failure. J Am Soc Nephrol 18: 904–912.

Löfberg H, Gubb AO. 1979. Quantitation of γ-trace in human biological fluids:indications for production in the central nervous system. Scad J Clin Lab Invest; 39: 619-626.

Lydakis C, Lip GY. 1998. Microalbuminuria and cardiovascular risk. QJM. 91(6): 381-91.

Madero M, Sarnak MJ, Stevens LA. 2006. Serum cystatin C as a marker of glomerular filtration rate. Cur Opin Nephrol Hypertens; 15: 610-616.

Metz CE. 1978. Basic principles of ROC analysis. Semin Nucl Med.8(4): 283-98.

Miyata T, Jadoul M, Kurokaw K, et al. 1998. β2-Microgolbulin in Renal Disease. J Am Soc Neprhol; 9: 1723-1735.

Mussap M, Plebani M. 2004. Biochemistry and Clinical Role of Human Cystatin C. Crit Rev Clin Lab Sci; 41(5-6): 467-550.

Namnum P, Insogna K, Baggish D, Hayslett JP. 1983. Evidence for bidirectional net movement of creatinine in the rat kidney. Am J Physiol. 244(6):F719-23.

Ngai HH, Sit WH, Jiang PP, Xu RJ, Wan JM, Thongboonkerd V. 2006. Serial changes in urinary proteome profile of membranous nephropathy: implications for pathophysiology and biomarker discovery. J Proteome Res. 5(11): 3038-47.

Oken DE, Flamenbaum W. 1971. Micropuncture studies of proximal tubule albumin concentrations in normal and nephrotic rats J Clin Invest. 50(7): 1498–1505.

Ohse T, Inagi R, Tanaka T, Ota T, Miyata T, Kojima I, Ingelfinger JR, Ogawa S, Fujita T, Nangaku M. 2006. Albumin induces endoplasmic reticulum stress and apoptosis in renal proximal tubular cells. Kidney Int. 70(8): 1447-55.

Orr MS, Goodsaid F, Amur S, Rudman A, Frueh FW. 2007. The experience with voluntary genomic data submissions at the FDA and a vision for the future of the voluntary data submission program. Clin Pharmacol Ther. 81(2): 294-7.

Pedrinelli R. 2002. Microalbuminuria, an integrated marker of cardiovascular risk in essential hypertension. Journal of Human Hypertension 16, 79–89.

Ransohoff DF. 2005. Bias as a threat to the validity of cancer molecular-marker research. Nature Reviews Cancer 5(2): 142-149.

Rodicio JL, Ruilope LM. 1995. Assessing renal effects and renal protection. J Hypertens Suppl. 13(4): S19-25.

Rosenberg ME, Silkensen J. 1995. Clusterin and the Kidney. Exp Neprhol; 3: 9-14.

Ruggenenti P, Remuzzi G. 2006. Time to abandon microalbuminuria? Kidney Int. 70(7): 1214-22.

Russo LM, Sandoval RM, McKee M, Osicka TM, Collins AB, Brown D, Molitoris BA, Comper WD. 2007. The normal kidney filters nephrotic levels of albumin retrieved by proximal tubule cells: retrieval is disrupted in nephrotic states.Kidney Int. 71(6): 504-13.

Rybak MJ, Frankowski JJ, Edwards DJ et al. 1987. Alanine Aminopeptidase and β2-Microgolbulin Excretion in Patinets Receiving Vancomycin and Gentamicin. Antimicrob Agents Chemother; 31(10): 1461-1464.

Schieppati A, Remuzzi G. 2003. Proteinuria and its consequences in renal disease. Acta Pediatrica Suppl. 443: 9-13.

Schmid H, Henger A, Cohen CD, Frach K, Gröne HJ, Schlöndorff D, Kretzler M. 2003. Gene expression profiles of podocyte-associated molecules as diagnostic markers in acquired proteinuric diseases. J Am Soc Nephrol.14(11): 2958-66.

Silkensen JR, Agarwal A, Nath KA, Manivel JC, Rosenberg ME. 1997. Temporal induction of clusterin in cisplatin nephrotoxicity. J Am Soc Nephrol. 8(2): 302-5.

Shankland SJ. 2006. The podocyte's response to injury: Role in proteinura andglomerulosclerosis. Kidney Int. 69: 2131-2147

Soejima A, Ishizuka S, Suzuki M, Miyake N, Fukuoka K, Nagasawa T. 1998. Biochemical renal manifestations induced by consecutive administration of gentamicin in rats. Nephron. 80(3): 331-9.

Sorensen PG, Nissen MH, Groth S, et al. 1985. Beta-2-microglobulin exretion: An indicator of long term nephrotoxicity during cis-platinum treatment? Cancer Chemother Pharmacol; 14: 247-249.

Suemori S, Lynch-Devaney K, and Podolsky, DK. 1991. Identification and characterization of rat intestinal trefoil factor: tissue- and cell-specific member of the trefoil protein family. Proc Natl Acad Sci USA 88(24): 11017-11021.

Tartoff J.B. 1996. "Clearance Methods in Renal Toxicology" in Methods in Renal Toxicology, edited by R.K. Zalups and L.H. Lash. CRC Press Inc., Boca Raton, FL, pp. 1-18

Tenstad O, Roald AB, Grubb A et al. 1996. Renal handling of radiolabelled human cystatin C in the rat. Scand J Clin Lab Invest; 56: 409-414.

Thielemans N, Lauwerys R, Bernard A. 1994. Competition between Albumin and Low-Molecular-Weight Proteins for Renal Tubular Uptake in Experimental Nephropathies. Nephron; 66: 453-458.

Thompson CS, Mikhailidis DP, Jeremy JY, Bell JL, Dandona P. 1987. Effect of starvation on biochemical indices of renal function in the rat. Br J Exp Pathol. 68(6): 767-75.

Thompson KL, Afshari CA, Amin RP, Bertram TA, Car B, Cunningham M, Kind C, Kramer JA, Lawton M, Mirsky M, Naciff JM, Oreffo V, Pine PS, Sistare FD. 2004. Identification of platform-independent gene expression markers of cisplatin nephrotoxicity. Environ Health Perspect. 112(4): 488-94.

Tkaczyk M, Nowicki M, Lukamowicz J. 2004. Increased Cystatin C concentration in urine of nephrotic children. Pediatr Nephrol; 19: 1278-1280.

Trof RJ, Di Maggio F, Leemreis J, et al. 2006. Biomarkers of Acute Renal Injury and Renal Failure. Shock; 26(3): 245-253.

Trollfors B, Bergmark J, Hiesche K, Jagenburg R. 1984. Urinary Alanine Aminopepticase and β2-microglobulin as Measurements of Aminoglycoside-associated Renal Impariment. Infection; 12: 20-22.

Tsuchiya Y, Tominaga Y, Matsubayashi K et al. 2005. Investigation on urinary proteins and renal mRNA expression in canine renal papillary necrosis induced by nefiracetam. Arch Toxicol; 79: 500-507.

Tucker SM, Mason RL, Beauchene RE.1976. Influence of diet and feed restriction on kidney function of aging male rats. J Gerontol. 31(3): 264-70.

Uchida K, Gotoh A. 2002. Measurement of cystatin-C and creatinine in urine. Clin Chim Acta; 323: 121-128.

van den Berg JG, van den Bergh Weerman MA, Assmann KJM et al. 2004. Podocyte foot process effacement is not correlated with the level of proteinuria in human glomerulopathies. Kidney Int. 66: 1901-1906

Van Liew JB, Zamlauski-Tucker MJ, Feld LG. 1993. Endogenous creatinine clearance in the rat: strain variation. Life Sci. 53(12): 1015-21.

Vaidya VS, Ramirez V, Ichimura T, Bobadilla NA, Bonventre JV. 2006. Urinary kidney injury molecule-1: a sensitive quantitative biomarker for early detection of kidney tubular injury. Am J Physiol Renal Physiol. 290(2): F517-29.

Verbaeys A, Ringoir S, Van Maele G, Lameire N. 1995. Influence of feeding, blood sampling method and type of anaesthesia on renal function parameters in the normal laboratory rat. Urol Res. 22(6): 377-82.

Verstrepen WA, Persy VP, Verhulst A, Dauwe S, De Broe ME. 2001. Renal osteopontin protein and mRNA upregulation during acute nephrotoxicity in the rat. Nephrol Dial Transplant. 16(4): 712-24.

Wade CE and Morey-Holton 1998. Alteration of renal function of rats following spaceflight. Am J Physiol Regulatory Integrative Comp Physiol 275: 1058-1065.

Wagner J. 2002. Nuclear (receptor) power: retinoids in rat mesangioproliferative disease. Nephrol Dial Transplant. 17 [S9]: 81–83

Waterboer T, Sehr P, Pawlita M. 2006. Suppression of non-specific binding in serological Luminex assays.J Immunol Methods. 309(1-2): 200-4.

Weichert-Jacobsen KJ, Bannowski A, Küppers F, Loch T, Stöckle M. 1999. Direct amifostine effect on renal tubule cells in rats. Cancer Res. 59(14): 3451-3.

Weir MR. 2007. Microalbuminuria and cardiovascular disease.Clin J Am Soc Nephrol. 2(3): 581-90.

Yoshida T, Kurella M, Beato F, et al. 2002. Monitoring changes in gene expression inrenal ischemia-reperfusion in the rat. Kid Int; 61: 1646-1654.

Zager RA.1987. Exogenous creatinine clearance accurately assesses filtration failure in rat experimental nephropathies. Am J Kidney Dis. 10(6): 427-30.

Zhou Y, Vaidya VS, Brown RP, Zhang J, Rosenzweig BA, Thompson KL, Miller TJ, Bonventre JV, Goering PL. 2008. Comparison of kidney injury molecule-1 and other nephrotoxicity biomarkers in urine and kidney following acute exposure to gentamicin, mercury, and chromium. Toxicol Sci. 101(1): 159-70.