21 October 2010
EMA/CHMP/SAWP/627644/2010
Committee for Medicinal Products for Human Use (CHMP)

Overview of comments received on 'Consultation on qualification opinion ILSI/HESI submission of novel renal biomarkers for toxicity' (EMA/283298/2010)

Interested parties (organisations or individuals) that commented on the draft document as released for consultation.

Stakeholder no.	Name of organisation or individual
1	Merck
2	ILSI HESI Nephrotoxicity Technical Committee
3	EFPIA

7 Westferry Circus • Canary Wharf • London E14 4HB • United Kingdom
Telephone +44 (0)20 7418 8400 **Facsimile** +44 (0)20 7418 8416
E-mail info@ema.europa.eu **Website** www.ema.europa.eu

1. General comments – overview

Stakeholder no.	General comment (if any)	Outcome (if applicable)
1	There is general agreement with the contents of the paper regarding the biomarkers. There is only one concern which is in regards to the statements in the paper pertaining to the histopathology.	
2	The HESI Nephrotoxicity Technical Committee (NTC) recognises the comprehensive nature of the review of its renal biomarker qualification submission conducted by the EMA Qualification Team and greatly appreciates the work involved. The NTC acknowledges and accepts the qualification opinion of the EMA. The NTC will continue to work to provide additional data to address many of the gaps identified during the course of the qualification exercise. However, the NTC feels it is useful to provide clarification of its position on some of the gaps identified by the CHMP with which it does not concur.	
3	EFPIA welcomes the EMA opinion of the ILSI/HESI submission of novel renal biomarkers for toxicity. The recommendations made on the basis of a critical assessment of the science base required to demonstrate the potential utility of promising biomarkers to assess kidney toxicity in rodent toxicology studies in general appears reasonable. However, EFPIA has major concerns about the identified histopathology gaps in the qualification process. Regarding reading slides in a not fully blinded fashion EFPIA considers that blinding of pathologists to dose group is not recommended. We do however agree that blinding of pathologists to novel and/or traditional biomarker changes is appropriate, as was done in the HESI submission. Regarding the use of multiple sections to increase reliability of the histological assessment, provided all relevant topographic regions	

Stakeholder no.	General comment (if any)	Outcome (if applicable)
	unique to each organ are evaluated for each animal (as is the current practice), no increase in diagnostic power would be achieved by review of additional sections. Thus, as further detailed under specific comments, EFPIA proposes to delete/modify these statements.	

2. Specific comments on text

Line no.	Stakeholder no.	Comment and rationale; proposed changes	Outcome
manuscript page 16. Section = Gaps identified by CHMP in the current qualification exercise, subsection = Histopathology Sentence = "Histopathology reading was not fully blinded Knowledge of treatment assignment can bias the results."	1	Comments: We believe that histopathologic evaluation with knowledge of dose group (e.g. unblinded evaluation) creates a high degree of fidelity and consistency in the ability to identify animal toxicities. This approach is aligned with established methods endorsed by the Society of Toxicologic Pathology ("Best Practices Guideline: Toxicologic Pathology" Crissman et al., 2004) for evaluation of nonclinical safety studies and concurs with many previous reports (Burkhardt, 2010, Dodd 1988, Iatropoulos 1984, Prasse 1986, Roe 1988, Weinburger 1979) that unblinded slide evaluation in non clinical safety studies, supplemented by "targeted masked evaluation", is the most sensitive way to accurately discriminate "true" ("toxicity-related") signals from "false" ("normal biological variation") signals. As this is also a critical objective of histopathology evaluation in biomarker qualification studies, we believe that an analogous unblinded approach is appropriate. In addition, biomarker studies must account for significant spontaneous changes that could affect biomarker values. Histopathology evaluation with knowledge of dose group does not detract from this, as the impact of spontaneous background changes will be captured in the variance of biomarker values of the control group. Given the strong historical performance of these Best Practices ("unblinded" and "targeted masked" histopathology evaluation) in non clinical safety studies, it is our opinion that these same Best Practices are appropriate to the conduct of non-clinical safety biomarker qualification studies. We believe that the	Partly accepted. Sentence will be removed but a statement will be included in the recommendations for the future experiments. Fully blinded histopathology would be an ideal requirement for full qualification of a BM without any limitations on its use, but in the proposed context full blinding is not required for biomarkers qualification. Justification: The objective of histopathology in biomarker (BM) qualification studies is not to discriminate toxicity-related signals from normal biological variation, but to assess the biomarkers ability to report histopathology changes whether they are the result of drug-induced toxicity, normal biological variation, or development of disease. A biomarker qualification study is not equivalent to a safety assessment study. However in the proposed context of use (detection of drug induced nephrotoxicity in addition to histopathology and current biomarkers) full blinding is not required.

		benefits of the unblinded approach far outweigh any concern of bias, and therefore that this not be identified as a "gap in the ...qualification exercise." This topic is also further discussed in the articles on the review of the PSTC submission of renal safety biomarkers that have recently been published in Nature (NATURE Biotechnology, Vol. 28 No. 5, May 2010). References attached at end of this comment document.[i] Definition: Targeted masked evaluation - involves re-evaluation of selected dose groups randomly combined with controls without knowledge of animal or group identity to determine whether a subtle or equivocal histomorphologic finding can be identified consistently from control tissues. Proposed change (if any): Suggest that the sentences in the manuscript are removed or that a clear and balanced paragraph is written that includes literature referenced in our comment. References are shown below.	
Page 16 Analytical methods	2	Comments: Regarding the statement that the impact of the criteria for repeatability, intermediate precision and reproducibility on the diagnostic performance of the biomarkers was not evaluated: The diagnostic utility of a biomarker cannot be evaluated independently of the corresponding assay repeatability, intermediate precision, and reproducibility. The assays utilized in the HESI program were adequate to yield data which demonstrated the diagnostic utility of novel biomarkers clusterin, α-GST, and RPA-1. Thus, the assays were adequate for purposes of the HESI program, though more specific and/or stringent criteria for repeatability, intermediate precision, and reproducibility may be valuable to optimize the diagnostic utility of the markers in future routine use.	**Not accepted.** Different assay performance could have an impact on the diagnostic performance of the biomarkers and this impact ideally should be evaluated (e.g. through modelling and simulation).

		Proposed change (if any): Remove the statement that the impact of the criteria for repeatability, intermediate precision and reproducibility on the diagnostic performance of the biomarkers was not evaluated.	
Page 16 Histopathology	2	Comments: Regarding the statement that the histopathology reading was not fully blinded and that knowledge of treatment assignment could bias the results: The histopathology was conducted in a manner consistent with the Society of Toxicologic Pathology Best Practices Guideline (Crissman et al, 2004, Toxicologic Pathology, 32, 126-131). Thus the reading was done as is standard for preclinical safety studies and knowledge of the treatment groups should not be considered to bias the results. Moreover, in the renal biomarker qualification studies, the pathologists were "blinded" to the biomarker data during the slide evaluation, and therefore there was no bias regarding these correlations. The practice of 'unblinded' followed by 'targeted masked' histopathology evaluation has been recently endorsed as appropriate for the conduct of nonclinical safety biomarker qualification studies (Burkhardt et al, 2010, Toxicologic Pathology, 38, 666-667) Proposed change (if any): Delete this as a gap.	Partly accepted See above.
Page 16 Histopatholo	2	Comments:	Partly accepted. Gap will be removed. Statement will be made in the recommendations for the future experiments.

gy		Regarding the statement that the histopathology would be more reliable if multiple sections from each kidney were evaluated: Reliability of the data should not be in question as long as all the important anatomic locations are present in the sections evaluated (e.g. cortex, medulla, papilla, pelvis) as was the case for the renal biomarker qualification studies. Since exposure to the kidney occurs via the blood, the distribution of the changes is expected to be similar throughout the kidney for a given anatomic structure. One section per kidney is considered adequate for preclinical safety assessment studies, and therefore this should also be regarded as adequate for the renal biomarker qualification studies. Proposed change (if any): Delete this as a gap.	For this context of use a single section is adequate. However in future submissions in order to elucidate the cases of positive BM response with no histopathology signal and for prodromal claims sponsors will need to address the number of sections needed.
Page 16 Biomarkers Normalisation	2	Comments: Regarding the statement that the biomarkers should be normalised to the individual baseline biomarker values: Baseline or pre-treatment data are typically used for the purpose of reducing extraneous sources of variability and thereby increasing the precision of estimated differences between experimental groups. For all novel markers considered in the HESI program, available repeat urinary measurements taken on control animals exhibited intra-animal variability of similar or greater magnitude than inter-animal variability. This suggests the use of baseline data may be of limited value with respect to variance reduction, and care should be taken to ensure that such use of baseline data does not result in variance inflation. However, collection of baseline data in future studies may be beneficial to characterize the dynamic range for each marker and the effects of age, gender, diet and	Accepted.

		circadian rhythm. Proposed change (if any): Refer to this as a suggestion for consideration in future programmes rather than a gap.	
Page 17 Reproducibility of experiments	2	Comments: Regarding the statement that the inconsistency between dose-finding and definitive studies for gentamicin and NPAA makes the interpretation difficult: The use of a selected dosage regimen is to produce the desired range of pathology in the target organ. The purpose of the dose-finding studies was to aid the dose selection of the definitive studies. The interpretation of the results was based solely on the results from the definitive studies. The diagnostic utility of a biomarker is assessed by evaluation of the marker's association with a defined pathology and is, therefore, independent of the dose used. Proposed change (if any): Delete this as a gap.	Not accepted. Consistent results between dose finding and definite studies would have provided additional reassurance for the validity of the qualification exercise.
Page 17 Difference between strains and inference	2	Comments: Regarding the statement that pooling together the results from the two strains is not considered optimal and complicates inference: Prior to conducting the pooled analyses, AUC_{ROC} values (with standard error) were calculated for each strain separately. For the novel biomarkers and relevant pathologies, the resulting AUC_{ROC} values were in close agreement. No inconsistencies in AUC_{ROC} values between strains were observed with respect to any qualification claims proposed in the HESI summary report.	Not accepted. A side by side comparison of Biomarkers AUCROC values show differences between strains. Confidence in a biomarker's performance is increased when both rat strains show higher sensitivity and specificity than sCr and BUN as was observed for clusterin for cortical tubular regeneration/basophilia and RPA-1 for collecting duct degeneration/necrosis.

		Proposed change (if any): Delete this as a gap.	
Page 17 Replication of evidence	2	Comments: For each novel biomarker considered in the HESI program, formal statistical procedures were utilized to adjust for the testing of multiple pathologies and reference markers. It is agreed that replicated evidence, if available, would serve to further bolster the program's findings. Proposed change (if any):	Accepted. It will be placed under recommendations for future experiments.
pp 16, lines 15-16: 'Histopathological reading .... can bias the results'	3	Comments: Best-practices for post-mortem assessment involve an understanding of the experimental conditions, including, but not limited to, dosage and duration of treatment. The need for an "unblinded" slide evaluation to obtain maximal fidelity and consistency is endorsed by the Society of Toxicologic Pathology ("Best Practices Guideline: Toxicologic Pathology" Crissman et al., 2004) for evaluation of nonclinical safety studies and concurs with many previous reports (Burkhardt, 2010, Dodd 1988, Iatropoulos 1984, Prasse 1986, Roe 1988, Weinburger 1979). Since the proper identification of histo-pathological alterations (considered to be the gold standard) is also a critical objective of histopathology evaluation in biomarker qualification studies, we believe that an analogous unblinded approach is appropriate. Moreover, in the renal biomarker qualification studies, the pathologists were "blinded" to the biomarker data during the slide evaluation, and therefore there was no bias regarding these correlations. EFPIA believes that the benefits of the unblinded	Partly accepted. See above.

		approach far outweigh any concern of bias, and therefore that this should not be identified as a "gap in the ...qualification exercise." Proposed change: It is recommended that the indicated lines are omitted.	
pp 16, lines 17-18: 'Assessment of .....was standardised across studies'	3	Comments: Regarding the use of multiple sections to increase reliability of the histological assessment, EFPIA considers, that provided all relevant topographic regions unique to each organ are evaluated for each animal (as is the current practice), no increase in diagnostic power would be achieved by review of additional tissue sections. Further, standardization of sectioning or blocking of tissues is not imperative provided that all relevant regions are captured and examined for each animal. In the HESI studies, different labs used blocking patterns that in their practice optimized the ability to comprehensively review renal topography, and serial sections were evaluated where needed to ensure that this occurred. Proposed change: It is recommended that the above mentioned text is omitted or reflects a more balanced view.	Partly Accepted for multiple sections. See above Accepted for the standardisation of sections or blocking.
pp 16, lines 19-30: 'Limitations of the studies ... a particular site'	3	Comments: CHMP suggests investigations on the specificity of the biomarkers for kidney injury; however, they make no comment on specificity of these BM for kidney injury being linked to their presence in the urine.	Partly accepted. The presence of BM in the urine limits some potential sources for false positive results but does not address all the factors that could affect specificity. E.g. excretion of creatinine into the urine is also a function of muscle mass.

		Proposed change: Suggest adding the comment that these BM are measured the urine, therefore, having the most potential to be kidney specific.	
pp 17, lines 10-18: 'Difference between strains and inference'	3	Comments: Regarding differences between strains and inference on marker performance. Although strain differences in susceptibility to toxicant-mediated injury are well recognized, the pooling of biomarker data from multiple strains in the ROC analysis performed by HESI is entirely justified. The strain differences in novel biomarker responses in the HESI studies reflected the differences in severity of histopathology changes between the 2 strains, however in a ROC analysis, the marker change for each individual animal is evaluated against its own respective histopathology finding across all possible diagnostic thresholds. In addition, anticipating this question on the part of the review team, HESI ran separate ROC analyses for the 2 strains as indicated in tables 9 and 10.	Not accepted. See above

[i] Burkhardt, J.E., et al. (2010). Topic of Histopathology Blinding in Nonclinical Safety Biomarker Qualification Studies. Toxicol Pathol 38, 666-667.
Crissman, J.W., Goodman, D.G., Hildebrandt, P.K., Maronpot, R. R., Prater, D. A., Riley, J. H., Seaman, W. J., and Thake, D. C. (2004). "Best Practices Guideline: Toxicologic Pathology." Toxicol Pathol 32, 126-131.
Dodd, D. A. (1988). Blind Slide Reading or The Uninformed versus The Informed Pathologist. Comm Toxicol 2, 81-91.
Iatropoulos, M. J. (1984). Appropriateness of Methods for Slide Evaluation in the Practice of Toxicologic Pathology. Toxicol Pathol 12, 305-6.
Prasse, K., et al. (1986). Letter to the Editor. Toxicol Appl Pharmacol 83, 184-5.
Roe, F. J. C. (1988). Toxicity Testing: Some Principles and Some Pitfalls in Histopathologic Evaluation. Human Toxicol 7, 405-410.
Weinberger, M. A. (1979). How valuable is Blind Evaluation in Histopathologic Examinations in Conjunction with Animal Toxicity Studies? Toxicol Pathol 7, 14-17.