PEER REVIEW HISTORY

BMJ Open publishes all reviews undertaken for accepted manuscripts. Reviewers are asked to complete a checklist review form (see an example) and are provided with free text boxes to elaborate on their assessment. These free text comments are reproduced below. Some articles will have been accepted based in part or entirely on reviews undertaken for other BMJ Group journals. These will be reproduced where possible.

ARTICLE DETAILS

TITLE (PROVISIONAL)	Association of Serum Alkaline Phosphatase with Mortality in non-selected European CKD 5D Patients. An Observational, three-center Survival Analysis
AUTHORS	Beige, Joachim; Wendt, Ralph; Girndt, Matthias; Fiedler, Roman; Queck, Karl-Heinz; Jehle, Peter

VERSION 1 - REVIEW

REVIEWER	Pasch, Andreas Dept. of Nephrology, Hypertension and Clinical Pharmacology University Hospital Bern, Inselspital Bern, Switzerland
REVIEW RETURNED	04-Nov-2013

GENERAL COMMENTS	This is an interesting manuscript showing with a relatively small number of patients that total alkaline phosphatase (but less so bone AP) levels are associated with all-cause mortality of CKD5D patients. While this is astonishing at first glance, the teleological reasoning that this might be related to the degradation of pyrophosphate, a calcification inhibitor, appears sound. In this regard, I would be grateful to see more detailed data on the alkaline phosphatase for which reportedly the mean value of several measurements were used. What was the variability and the intraclass correlation? How stable is total AP in a given individual over time? What does a single measurement tell us? Given the lack of association with mortality in this cohort, do the authors recommend that total alkaline phosphatase should be measured instead of calcium, phosphate and PTH in the future? Could the authors speculate on the reason for the enhanced activity of total AP, but not bone AP, in CKD5D patients? May there be a link between total AP, chronic inflammation and oxidative stress (e.g. Lee HL, et al.; Biochem Biophys Res Commun 2010), and may AP in this regard be a marker of inflammation/oxidative stress? Can the authors provide data on inflammation markers in this cohort (e.g. CRP)?

REVIEWER	Abramowitz, Matthew Albert Einstein College of Medicine, USA
REVIEW RETURNED	12-Nov-2013

GENERAL COMMENTS	This is a study by Beige and colleagues examining the association of CKD-MBD markers with mortality in dialysis patients, with the stated focus of a comparison of total AP with bone AP (or skeletal

	AP). More detail should be provided about the selection of the study cohort. There is insufficient detail provided regarding the statistical analyses and too little data resulting from those analyses is provided. This makes it difficult to evaluate the import of their findings. In addition, the Discussion does not provide a clear interpretation of their findings. My specific comments are below. 1. If total AP is more informative than bAP and in fact neither bAP nor other MBD markers were associated with mortality, then what is the meaning of total AP? The Discussion does not adequately address this question. This also makes their mention of a possible clinical intervention difficult to understand. Is the association with mortality reflective of MBD and/or vascular calcification, or is it perhaps something distinct, like a marker of subclinical liver disease (e.g. fatty liver)? Also, their lack of confirmation of others' findings related to MBD markers is most likely a result of the modest sample size of this study. 2. More information is needed about the cohort. Of 719 patients, 407 were included. Is this due to their stated eligibility criteria, or were there other factors? Were patients excluded because of an insufficient number of laboratory samples? How many lab measures did people have? How many missing timepoints were there? How many patients were HD versus PD? 3. The rationale behind averaging 3 years of laboratory data into one number is unclear. There are other methods to deal with longitudinal data that would better preserve the variability over time in these measures and would likely be preferable. 4. The 2 definitions used to define the start of follow-up time are both troubling because of their potential to introduce bias. Follow-up time should likely start with the entry into the study, i.e. the date of the first lab measurement. Using the start of dialysis in a mixed cohort of incident and prevalent patients means the prevalent patients' start time begins years before the predictors of interest were measured. Using the date of median lab analysis means the duration of follow-up will vary by the number of measurements. 5. The authors should consider excluding patients with known liver disease (not only biliary tract disease as they did). They could also consider this as a sensitivity analysis, and should also consider hepatitis B and C status. 6. What were the parameters used for the forward stepwise analysis? Was there consideration given to confounding? 7. The univariate survival analyses should be reported before the multivariable analyses. The logistic regression analysis of mortality should be removed as it does not account for the factors that are accounted for by the Cox regression analyses. How were the p-values in Table 2 calculated? Hazard ratios and 95% CIs should be reported for the Cox regressions. 8. Separate models for total AP and bAP should be analyzed. Given their high degree of correlation, it is inappropriate to include them in the same model. 9. Data should be provided for the quartile analyses discussed on page 8. 10. P-values should be reported instead of n.s. 11. Table 1: vintage, PTH, and bAP should probably be reported as median (IQR) as they do not appear to be normally distributed. 12. The first paragraph of the Methods should clarify that entry into the cohort was from 2008-2010 (assuming that this is correct).

VERSION 1 – AUTHOR RESPONSE

Reviewer 1

- *"… more detailed data on the alkaline phosphatase"*: Tab. 1 and Tab. 3 and the text section of methods and discussion were modified with these new data.
- *"Given the lack of association with mortality in this cohort, do the authors recommend that total alkaline phosphatase should be measured instead of calcium, phosphate and PTH in the future?"*: We had already commented on this effect by a hind on our population exhibiting lower values compared to the classical *Block* study. We would not recommend skipping phosphate just now but would suggest analyze this along with bone parameters in further studies (p10 l14).
- *"Could the authors speculate on the reason for the enhanced activity of total AP, but not bone AP, in CKD5D patients? May there be a link between total AP, chronic inflammation and oxidative stress (e.g. Lee HL, et al.; Biochem Biophys Res Commun 2010), and may AP in this regard be a marker of inflammation/oxidative stress? Can the authors provide data on inflammation markers in this cohort (e.g. CRP)?"*: We strengthened the already available discussion on pyrophosphate and inflammation (p11, l15), but have unfortunately only a non-comprehensive dataset on CRP.

Reviewer 2:

1. With concern to our new finding on the association of low extreme sAP and better survival we modified the discussion on the differential meaning of AB and sAP with a focus on repeated values, extremes and trends. *"Also, their lack of confirmation of others' findings related to MBD markers is most likely a result of the modest sample size of this study."* We commented this point already in our text on p10, l13.
2. We now give one more figure (n=493 maintenance patients) to the already available data to evaluate to inclusion efficacy. *"How many missing timepoints were there?"* – Number of data time points is now available in Table 1. *"How many patients were HD versus PD?"* – see Table 1.
3. We now included both range and extreme analyses in the survival studies.
4. We have now included the time of the first laboratory data as start time into analyses
5. "Know liver diseases", e.g. fatty liver are not easy to characterize in a quantitative fashion and could therefore not be separated. Concerning hepatitis, there was only one patient with known hepatitis C antibody positivity, but no apparent liver disease included in the dataset.
6. Forward stepwise analysis factors are given on p8, l15. Confounding was considered by adjusting for diabetes, age and vintage as described.

7. The univariate survival analyses have now been switched to the beginning of the result section. *"The logistic regression analysis of mortality should be removed as it does not account for the factors that are accounted for by the Cox regression analyses"*. I do not entirely understand what is meant by logistic regression analyses in opposite to Cox analyses. The Odd´ s ratio for mortality of 2.7 (p7 l25) was calculated by a contingency table and not by logistic regression. Because this raw, unadjusted estimate is a central measure of risk analysis we suggest to leave it within the manuscript. We added all p-values to Tab. 3, including the non-significant ones and the hazard ratios (exp. B) as requested.
8. We repeated all analyses with separate models for AP and sAP. Concerning total AP, no differences to the global model were noted. Mean skeletal AP entered the models after AP exclusion and was indicated in such way in Tab. 3.
9. The quartile analyses were now deleted because they do not inherit additional information. Together with the requested detailed information for 4 but not 2 strata (comparable to Tab. 1) they would make the manuscript difficult to follow because of changing sorting principles.
10. P-values are now available in Tab. 3.
11. Table 1: vintage, PTH, and bAP have now been included also as median values.
12. Within methods section, both cross-sectional data-set entry, follow-up and laboratory retrieval was defined in more detail (p4 l15).

VERSION 2 – REVIEW

REVIEWER	Pasch, Andreas University Hospital Bern, Inselspital Dept. Nephrology, Hypertension and Clinical Pharmacology Freiburgstrasse 15 3012 Bern Switzerland
REVIEW RETURNED	04-Jan-2014

GENERAL COMMENTS	My questions have been adequately answered in the revised version of the manuscript. There are minor points which need to be corrected though: p 22: "...no significant population differences..." -> this is not true as diabetes is different; please correct. p 27: "signal threshold signal" -> I guess is one "signal" too many. p 28: Figure 1a and 1b show histograms of sAP and AP, not of men and women. please use sAP, not bAP in Figure 1a. p 28: Figure 2 "," instead of "." p 28: Figure 3 "survival after laboratory test" -> does that mean

	"survival after first laboratory test"?

REVIEWER	Abramowitz, Matthew Albert Einstein College of Medicine, Medicine
REVIEW RETURNED	08-Jan-2014

GENERAL COMMENTS	There are still several issues that should be addressed. 1. While an interesting finding, the analyses using the minimum sAP value seem somewhat arbitrary, especially knowing they were not part of the original analysis. For example, why is an analogous analysis for total AP not presented? Why was the minimum chosen as opposed to the maximum? The authors should better explain the motivation for this analysis. 2. Regarding the conditional stepwise analyses, what were the parameters used (i.e. p value) to determine inclusion/exclusion of covariates? Was the magnitude of change in the coefficient of the predictor of interest (e.g. sAP) also used to determine inclusion/exclusion of other covariates? 3. Table 3: the HRs for the averaged sAP should be provided. 4. How do the authors explain the disparate results between the time-averaged sAP and minimum sAP in the model without total AP? Given the multicollinearity between sAP and total AP, these results are likely the most valid, as opposed to the model including both measures. 5. Page 26, line 4: Again, given the multicollinearity between sAP and total AP, it is very likely not meaningful that the association of mean sAP is no longer significant after inclusion of total AP in the model. Similarly, I do not think it is valid to use this to suggest that the mean of one is important whereas the minimum of other is more useful.

VERSION 2 – AUTHOR RESPONSE

Reviewer 1
Language advice: The manuscript was cross-checked both by a German-native translation professional and by an English native non-medical scientist.
Vintage: This is a well defined wording for the time period under disease condition but before study recruitment.
p 22: We added a half sentence within text section to address this issue
p 27: "signal threshold signal …" Thank you, we corrected this typo.
p 28: Figure 1a and 1b show histograms of sAP and AP, not of men and women.
We modified the figure description accordingly.
p 28: Figure 2 "," instead of "." We corrected this, thanks again.
p 28: Figure 3 "survival after laboratory test" -> does that mean "survival after first laboratory test"?
Yes
Reviewer 2
1. While an interesting finding, the analyses using the minimum sAP value seem somewhat arbitrary, especially knowing they were not part of the original analysis. For example, why is an analogous analysis for total AP not presented? Why was the minimum chosen as opposed to the maximum? The authors should better explain the motivation for this analysis.
We fully agree with your concern. However, this kind of post-hoc analysis was suggested by Reviewer 1 and we do also understand his reasoning to explore the value distribution. We now indicated the

minimum – maximum data as “secondary analysis” in “Methods”, “Results” and “Discussion” and hope this might help to evaluate our data. We do not presented AP minimum – maximum data since this analysis yielded no significant association with mortality.

2. Regarding the conditional stepwise analyses, what were the parameters used (i.e. p value) to determine inclusion/exclusion of covariates? Was the magnitude of change in the coefficient of the predictor of interest (e.g. sAP) also used to determine inclusion/exclusion of other covariates?

For analysis in the conditional models, variables were selected based on p-values and assumptions coming from literature. This procedure was added on p5 l16.

3. Table 3: the HRs for the averaged sAP should be provided.

Unfortunately, the SPSS package does not provide HR for not significant values. I would be thankful for an advice how to produce these HR´s.

4. How do the authors explain the disparate results between the time-averaged sAP and minimum sAP in the model without total AP? Given the multicollinearity between sAP and total AP, these results are likely the most valid, as opposed to the model including both measures.

To address this point we changed the Discussion section on p10 and payed attention to the correlation between AP and sAP and the possible bias.

5. Page 26, line 4: Again, given the multicollinearity between sAP and total AP, it is very likely not meaningful that the association of mean sAP is no longer significant after inclusion of total AP in the model. Similarly, I do not think it is valid to use this to suggest that the mean of one is important whereas the minimum of other is more useful.

We added your perceptions to the Discussion section at p11 and describe the different associations in the new version from a more descriptive viewpoint and do not make recommendations concerning a particular marker.