RANK-ligand and osteoprotegerin as biomarker in the differentiation between periprosthetic joint infection and aseptic prosthesis loosening - Response to Reviewers

Reviewer 02699758:

1. *Accuracy or coefficients of variation for serum RANKL and OPG is recommended to be described.*

Here we have calculated the column statistics for OPG and RANKL:

RANKL	PJI	Aseptic Loosening	Control
Mean	111,8	177,6	214,3
Std. Deviation	51,49	189,7	274,5
Std. Error of Mean	8,964	26,82	52,84
Lower 95% CI of mean	93,5	123,7	105,7
Upper 95% CI of mean	130	231,5	322,9
Coefficient of variation	**46,08%**	**106,81%**	**128,10%**
Geometric mean	99,99	116,7	148,8
Geometric SD factor	1,633	2,424	2,209
Skewness	**0,3623**	**2,187**	**4,195**
D'Agostino & Pearson normality test			
K2	6,569	35,61	55,28
P value	0,0375	<0,0001	<0,0001
Passed normality test (alpha=0.05)?	No	No	No

OPG	PJI	Aseptic Loosening	Control
Mean	8,057	6,777	7,538
Std. Deviation	4,658	2,627	4,056
Std. Error of Mean	0,8234	0,3791	0,7805
Lower 95% CI of mean	6,378	6,014	5,933
Upper 95% CI of mean	9,736	7,539	9,142
Coefficient of variation	**57,81%**	**38,76%**	**53,80%**
Geometric mean	7,274	6,33	6,946
Geometric SD factor	1,529	1,452	1,452
Skewness	**3,261**	**1,271**	**3,7**
Kurtosis	13,76	2,732	16,63
D'Agostino & Pearson normality test			
K2	47,27	17,39	49,42
P value	<0,0001	0,0002	<0,0001
Passed normality test (alpha=0.05)?	No	No	No

We did not assume normal distribution in testing for significance, as described in the statistical methods section. As it can be seen from the log-scale graphs and the numbers above, our data has no normal distribution, therefore the %CV is difficult to apply.

Considerable Skewness of the data suggested log-normal distribution; We have therefore log-transformed our data, and re-did the column statistics:

LOG10(RANKL)	PJI	Aseptic Loosening	Control
Mean	2	2,067	2,173
Std. Deviation	0,213	0,3845	0,3443
Std. Error of Mean	0,03708	0,05437	0,06626
Lower 95% CI of mean	1,924	1,958	2,036
Upper 95% CI of mean	2,076	2,176	2,309
Coefficient of variation	**10,65%**	**18,60%**	**15,85%**
Geometric mean	1,989	2,034	2,147
Geometric SD factor	1,114	1,198	1,169
Skewness	-0,1586	0,6045	0,6557
Kurtosis	-1,287	-0,8193	1,325
D'Agostino & Pearson normality test			
K2	8,429	6,097	4,318
P value	0,0148	0,0474	0,1154
Passed normality test (alpha=0.05)?	No	No	Yes

LOG10(OPG)	PJI	AL	Kontrolle
Mean	0,8661	0,7995	0,8399
Std. Deviation	0,1858	0,1545	0,165
Std. Error of Mean	0,03338	0,02278	0,03237
Lower 95% CI of mean	0,7979	0,7536	0,7733
Upper 95% CI of mean	0,9342	0,8454	0,9066
Coefficient of variation	**21,46%**	**19,33%**	**19,65%**
Geometric mean	0,848	0,784	0,8259
Geometric SD factor	1,23	1,226	1,202
Skewness	1,014	-0,05679	1,511
Kurtosis	2,39	0,2444	4,797
D'Agostino & Pearson normality test			
K2	9,905	0,3609	17,39
P value	0,0071	0,8349	0,0002
Passed normality test (alpha=0.05)?	No	Yes	No

Again, the tests for normal distribution showed a significant non-normal distribution in three out of 6 columns, but Skewness was well reduced and data distribution smoothened. The values for coefficients of variation are now within an acceptable range for a diagnostic procedure, given the limited sample size.

We agree that accuracy is a matter of importance in diagnostic testing, however, the data set limits the use of the statistical test. In our opinion, the full presentation of these column statistics would confuse most the readers. Even among medical professionals, the number of those with statistical knowledge sound enough to correctly interpret the data is very limited. We therefore suggest presenting the “Geometric CV after logarithmic transformation” in the paper, which was included into the methods and results sections of the manuscript.

2. *Whether implant is stable or loose is according to the operation record. This judgment is crucial in this paper. There are acetabular component and femoral component in the case with THA and also there are 2 or 3 components in the TKA case. How did operator make judgments? If at least one component is loose, is the case grouped into the loose case?*

Initially, we differentiated if one or more component was loosened, and analyzed in separate groups, but no difference was seen here. In the final evaluation, we regarded any implant as "loose" where at least one component with contact to bone was loosened, and still could not show a difference in RANKL, OPG or Ca/Ph levels between stable and loose implants.

Assessment of implant stability is crucial in many surgical interventions and day-to-day routine for the experienced arthroplasty surgeon. The intra-operative decision between loose and stable is also included in the acknowledged treatment algorithms for arthroplasty revision. After soft tissue debridement, the surgeons assessed implant stability intraoperatively by attempting implant removal with a proper removal instrument; If the implant would move within the bony interface or could be removed without further effort, especially without further disruption of the interface, the implant was considered loose; We have included this briefly in the materials and methods part page 7

3. *I suggest that preoperative radiography is taken into consideration. Did all patients with PJI or AL have clear zone or loosening sign around the prosthesis on pre-operative radiography? If not, I recommend authors reevaluate RANKL and OPG according to the X-ray findings.*

We regarded the surgeon's intraoperative judgement of stability as the gold standard. Preoperative radiographs were taken from all patients and were assessed previous to surgery. The findings on conventional radiographs are highly variable between different types of implants, and have a moderate inter-observer reliability, especially without the possibility to compare to previous radiographs. Radiographic signs of loosening were taken into account where implant migration or dislocation could be seen; The mere presence of a lysis margin adjacent to the implant or the cement was no valid criteria; We have clarified the materials & methods section accordingly on page 7

Reviewer 01408945:

1. *There are many grammatical errors in English.*

Our manuscript is well written and in readable, understandable, fluent English. We acknowledge that as non-native speakers, further language editing might be necessary, which was completed through a professional text editing service.

2. *The introduction section is very long. Authors should inventory it.*

To the informed orthopedic expert, the introduction might seem lengthy, while experts from other areas will be glad for the additional information. The introduction is focused on the role of RANK/OPG/RANKL, and the pathophysiologic differences between aseptic loosening and periprosthetic infection, which is crucial to the paper. We see no need to shorten or inventory this section any further.

3. *Authors described that between 2010 and 2011 we included 120 consecutive patients. Why dose [sic!] authors use these old data?*

Indeed, the data was collected in 2010 and 2011, and the analysis was carried out in this time; We did not publish the data immediately in 2011, but had focused on other projects and other factors, as you can see from the group's publication record. A recent literature review showed us that very little has since been published on the matter of RANKL and PJI, motivating us to revisit our data, and assemble them for publication.

4. *Authors repeated their hypothesis frequently.*

Hypothesis is stated once at the end of the introduction and again at the beginning of the discussion, as it is common in scientific writing.

5. *Authors described that we found no significant differences in the mean values of circulating RANKL and OPG in PJI vs. AL or control group, but with a certain trend of lower RANKL concentrations and higher OPG concentrations in the PJI group. These results suggested that there are no worthy to measure RANKL and OPG.*

Negative or non-significant results still contribute to scientific evidence and should be considered for publication. The trend towards publication bias of highly significant results is not favorable; We comment on the limitations of our study in the discussion section, and do point out that our results are not significant, therefore we do not see exactly what the reviewer is trying to argue here.

6. *Abbreviation must be cited when they appeared at the first time.*

The manuscript was text-edited to make sure abbreviations are cited correctly.

Reviewer 02699644:

1. *My question around the methodology would be were the PJI group all diagnosed with positive cultures and if so this should be mentioned. if not a reason given and explanation as to whether these may be aseptic loosening.*

PJI was considered proven in accordance with the MSIS consensus paper on the diagnosis of PJI by Parvizi et al; Positive cultures are a minor (1 positive) or major (2 positive) factor in this definition. But also with no bacterial growth, patients with a positive finding in histology and cell count/differentiation of the aspirate will be grouped into the PJI group without positive microbiology. We have clarified this in the materials & methods section of the manuscript.

In our collective, 31 out of 48 patients (64%) in the PJI group had consistent finding of two or more positive microbiology cultures, matching the "major" MSIS criterium for microbiology; Another 5 had one positive culture, the remaining 12 patients were "culture negative" PJIs. We have included this into the results section of the manuscript.

2. *Also several researchers have shown a variable positive organism growth in 'aseptic' loosening and some mention should be made to this in the discussion. Diagnosing PJI can be very difficult and there may have been some low virulence organisms in the AL group that weren't cultured...comment about this should be made as well.*

This is a valuable comment, as it touches one of the core problems in diagnostic studies in revision arthroplasty. The MSIS consensus classification is the current "gold standard", but MSIS acknowledges in their paper that "infection may be present" even if with less than three minor criteria are fulfilled; However, we have to agree on one gold standard definition. We cannot guarantee that patients with low-grade infections and low virulence may be misclassified into the "aseptic loosening" group; We have added this consideration to the limitations of our study in the discussion section of the manuscript.

A long time follow up can help to see determine if primary diagnosis was correct and if the treatment path chosen was successful; Five-year follow up data is just now coming in and will be subject to another future publication.