PEER REVIEW HISTORY

ARTICLE DETAILS

TITLE (PROVISIONAL)	Kidney failure related to broad-spectrum antibiotics in critically ill patients: secondary end point results from a 1200 patient randomized trial
AUTHORS	Jens Ulrik Jensen, Lars Hein, Bettina Lundgren, Morten Heiberg, Thomas Mohr, Mads Holmen Andersen, Klaus Julius Thornberg, Jesper Løken, Morten Steensen, Zoe Fox, Hamid Tousi, Peter Søe-Jensen, Anne Øberg, Ditte Gry Strange, Nanna Reiter, Katrin Thormar, Paul Christian, Kim Michael Larsen, Niels-Erik Drenck, Maria Egede Johansen, Lene Ryom, Christian Østergaard, Jesper Kjær, Jesper Grarup, Jens D. Lundgren and The Procalcitonin And Survival Study (PASS) Group

VERSION 1 - REVIEW

REVIEWER	Natalie Ives Senior Research Fellow University of Birmingham, UK No conflicts of interest
REVIEW RETURNED	22/11/2011

THE STUDY	In the statistical analysis section the authors give various renal endpoints, but which is considered the primary endpoint is not specifically stated. Summary: For a secondary analysis, the use of the words 'To determine...' seem a little strong. The original study was never powered to assess this, and they do state in the strengths and weaknesses that the observed numbers call for a slightly higher sample size. Abstract: Here they actually give the results from the multiple effects model, which is lacking from the main results.
RESULTS & CONCLUSIONS	Although I have answered yes to the first two questions above, this is based on the use of appropriate methods to address the question. However, I do think the results have been over-emphasised based on one highly significant result for the % of days within day 1-28 with eGFR<=60 ml/min/m2 ($p<0.0001$). It would be of interest to know how many patients this related to, which is how the other two renal endpoints are reported. Further, the other renal endpoints were less conclusive with $p=0.02$ and $p=0.04$. It would also have been useful to present some point estimates (odds ratios, relative risks) alongside the 95%CI to assess the results. Similarly in the discussion stating that this is "the first substantive evidence" seems a little strong considering the p-values, and with

	the lack of data presented on the regression analyses. Although in the conclusions the authors do state the need for further research. The statistics relating to Figure 3 are missing, what were the results from the statistical analysis. The plot is useful but it doesn't inform the reader fully. I also don't find Figure 4 vey helpful, it would be more useful to present the results from the multiple effects model in a Table which is included in the main paper, this was not even provided as supplementary information. More statistics are presented for the sensitivity analyses than the main analyses, which makes it all a little confusing.
GENERAL COMMENTS	Include information on the regression analyses (table of the results would be useful to the reader), and reduce the amount of information provided on the various sensitivity analyses.

REVIEWER	Mical Paul Senior physician and Consultant Unit of Infectious Diseases, Rabin Medical Center, Beilson Hospital Israel
REVIEW RETURNED	26/11/2011

THE STUDY	The current analysis does not have a primary outcome. Since the fact that the "high exposure" arm had a higher rate of renal failure has been previously reported, the new information in the current manuscript is on the association between type of antibiotic and renal failure. The methods for assessment of this question are not well described. See details in the comments to the authors below. Careful attention is needed to the English language in the manuscript. For example, in abstract: “extreme bilirubin or triglycerides” and “persons held by force”.
RESULTS & CONCLUSIONS	Results regarding the association between type of antibiotic and renal failure are less credible because the analyses presented are not clear enough. See details in the comments to the authors below.
GENERAL COMMENTS	The current analysis is based on a well-performed and interesting study published in CCM and referred to in the current manuscript. The assessment of the outcome of renal failure is worthy. Comparisons of ICU days with estimated glomerular filtration rate <60 mL/1.73 m2; Patients with estimated glomerular filtration rate <60 mL/1.73 m2; ICU days spent with dialysis treatment; and patients in treatment with dialysis at discharge/ death were reported in the original report of the trial in CCM. Even though this is a secondary analysis of a RCT, I believe that a primary outcome measure for the current analysis should have been defined and used. Currently, conclusions are based on the variable with most significant results and not necessarily the most important variable for the patient. Considering patient-relevant outcomes I would regard as most important a dichotomous variable of death or discharge alive with RIFLE criteria of loss or ESRD, which was not assessed in this study. Looking at the original trial report it seems that there is no difference between the groups with regard to this outcome measure.

	Study outcomes: what about RIFLE categories loss and ESRD? Only risk, injury and failure are defined in methods (these should be spelled out). The comparison between antibiotics is not based on randomization and is thus an observational finding. As such, the analysis assessing the association between antibiotics and renal failure is not presented in enough details in methods and the results presented questionable. An appropriate analysis of this association would be complex because patients received concurrently more than one antibiotic, as per the antibiotic guidelines presented in the main manuscript. Furthermore, this association should have been adjusted to other risk factors for renal injury and this adjustment although probably performed is not well presented in methods and the adjusted analysis is not presented in tables. The authors should be commended for considering death or renal failure as an adverse outcome (and not renal failure alone). The handling of deaths should be presented more clearly in the methods section. Tables 1 and 2 and figures 1 and 2 are similar to or the same as those presented in the main report of the trial in CCM and are probably redundant here. Rather than the "all" column in table 1, p-values would have been more helpful

REVIEWER	I declare that my institution has been the recipient of an unrestricted grant from Brahms AG, the manufacturer of the Procalcitonin assay used in this study. Yours sincerely, Assoc. Prof. Michelle S Chew Departments of Intensive Care Medicine and Clinical Sciences Malmö Malmö University Hospital Lund University Sweden
REVIEW RETURNED	04/12/2011

GENERAL COMMENTS	**Summary** The present study is an extension of the original PASS study by Jensen et al published recently (1). This study showed that an antibiotic escalation algorithm based on repeated Procalcitonin (PCT) measurements did not improve 28 day survival. Further, despite substantially higher use of broad-spectrum antibiotics, the study identified several unexpected effects of this strategy: increased time spent on mechanical ventilation, increase length of ICU stay and increased number of days with renal failure defined as eGFR 60ml/min/1.73m2. Based on these findings, the aim of the present study was to further delineate the effect of the PCT strategy (high exposure) on various markers of renal function, compared to the current 'standard of care' (standard exposure). This secondary analysis of the PASS data is relevant because renal failure occurs commonly in the critically ill and is associated with adverse outcomes. The study appears to have been carefully designed and the escalation protocol for the 'high exposure' group is relevant for the specified critically ill population.

	Kidney injury end-points were defined as 1) Risk (R), Injury (I) or Failure (F) according to the RIFLE criteria; 2) eGFR<30 or 60 ml/min/1.73m2 corresponding to moderately severe and severely impaired renal function respectively; or 3) BUN>20mmol/L at any time during the 28-day follow-up period. Further, the duration of renal failure based on eGFR measurements was also analyzed. The main results were that a 'high exposure' strategy of antibiotic escalation based on PCT measurement was associated with adverse renal outcomes. Specifically, a multivariate analysis identified that this effect was due Piperacillin-Tazobactam (Pip-Tazo). Supporting the suggestion of a nephrotoxic effect of Pip-Tazo was the rate of improvement of eGFR when the drug was discontinued. The paper is generally well written (only minor language revisions required), the hypotheses clearly stated, and the methods specified in detail in the accompanying supplement. The authors have obviously considered the study protocol carefully, and use appropriate statistical methods to analyze the data. I do however have a number of concerns that I think the authors should address. **Choice of end points** One may question the utility of BUN as a renal end-point in this population. BUN is often increased in the critically ill due to parenteral feeding and catabolism and does not accurately reflect renal function. Another important limitation is using eGFR in this population since this measurement is only valid in steady state. Since eGFR measurement forms the bulk of evidence behind the conclusions in this study, I suggest that the authors clearly state and discuss its limitations. Finally, the choice of the RIFLE criteria instead of the AKIN criteria may be questioned since the AKIN criteria reflect even smaller changes in renal function, which are associated with poorer outcomes (2,3), and classifies patients with RRT into stage 3 (≈F in RIFLE). The authors should include these limitations in their discussion since it using the AKIN or other criteria than eGFR may potentially change their results and conclusions. Except for renal injury defined as the 'risk (R)' group, the number of patients with AKI was not significant between groups according to the RIFLE criteria. This is inconsistent with the eGFR findings. The authors should discuss this further, taking into account the limitations of the chosen end-points. **Higher exposure, sicker patients, or both?** The high exposure group indeed had significantly increased Defined Daily Doses (DDDs) of broad spectrum antibiotics administered and significantly more days spent in ICU with at least three antimicrobials. There were significant increases in the DDDs of Piperacillin-tazobactam in the high exposure group (2925 vs 1893), as well Ciprofloxin although the difference with the latter antibiotic was much less marked (8382 vs 6210 DDDs). While the data is clear that increased exposure to antibiotics were associate with adverse outcomes, it is not clear whether or not this was due to a greater severity of illness in patients at the time of antibiotic escalation in the high exposure group. Was Pip-Tazo prescribed to sicker patients and with already higher degrees of renal impairment? Is it possible the kidney injury occurred more commonly in the 'high exposure' group due to more severe underlying disease? While similar 28-day mortalities in the two groups speak against a difference in illness severity, I note that the 'high exposure' group spent more time in ICU, and in their original study the authors noted a tendency for

	increased vasopressor and inotrope requirements and longer periods with severe sepsis and septic shock (1). eGFR multivariate analysis was adjusted for APACHE and degree of host response at baseline, but no adjustments for continuing organ failure were made. Do the different antibiotic groups differ in terms of SOFA, need for RRT and other important diagnoses such as sepsis for example, at thetime of antibiotic administration/+ escalation? Notably eGFR at the initiation of the investigated antibiotic was included in the multivariate analysis. The authors should be commended for their foresight in including this variable, however, my previous concerns regarding using eGFR are valid even here. A substantial number of days (slightly less than 30% in both groups) for the assessment of renal failure were not included. The authors have mentioned this as a limitation of their study. Whilst the authors have noted that the rate of renal failure on the last day of follow-up in the two groups were comparable, I am uncertain how it reflects a 'temporary extension of the duration of renal failure in the high exposure group'. Given the magnitude of the findings in the 28-day follow up, I would have expected the rate of renal failure to be higher even at this timepoint. Ca the authors explain this, as well as how the data reflect a 'temporary extension of renal failure'? Another query relates to the use of a 10-day eGFR to indicate recovery of renal function. If 28-days of follow-up data are available, why limit the recovery period to 10 days? Kidney injury is known to adversely affect survival in critically ill patients, therefore it is surprising that despite the striking results seen in the present study, no mortality effect was observed. Do the authors have a possible explanation for this? The authors write that this large clinical study suggests that Pip-Tazo in itself is nephrotoxic, and that this finding is supported by other studies in experimental settings. This is not accurate. Previous studies show that Piperacillin reduces elimination of other antibiotics and increase the toxicity of other antibiotics ostensibly by affecting renal clearance (4-7). The text should be changed to reflect this. As an extension to this, have the authors considered that co-administration of Pip-Tazo with other drugs may have caused a nephrotoxic effect due to the decreased elimination of these other drugs? Finally, the study design does not allow for a conclusion to be made regarding cause and effect. The authors should state this clearly. **Strengths of this study** Notwithstanding the above concerns, this study has some important, clinically relevant strengths. Firstly, the conclusion that Pip-Tazo adversely affects renal outcome is supported by the analysis that eGFR increases after discontinuation of the drug. Secondly, excluding patients with early deaths (within 7 days), fungal infections and combined □- lactam+fluorquinolone exposure did not change the conclusions. Finally, the adverse effect on renal function was seen only with Pip-Tazo but not Cefuroxime and Meropenem. Whilst a cause-effect relationship cannot be proven due to the design of this study (and this should be highlighted by the authors), the findings ring a warning bell for the administration of Pip-Tazo on an escalation regime based on PCT levels. This study should catalyse future studies investigating the possible nephrotoxic effect of Pip-Tazo, including the use of various dosing regimes (eg.

	extended infusions), and coadministration with other antibiotics. **References** 1. Jensen JU, Hein H, Lundgren B et al. for the Procalcitonin and Survival Study (PASS). Procalcitonin-guided interventions against infections to increase early appropriate antibiotics and improve survival in the intensive care unit: A randomized trial. Crit Care Med 2011;**39**(9):2048-58 2.Barrantes F, Tian J, Vazquez R, Amoateng-Adjepong Y, Manthous CA. Acute kidney injury criteria predict outcomes of critically ill patients. Crit Care Med. 2008;**36**(5):1397-403. 3. Chertow GM, Burdick E, Honour M, Bonventre JV, Bates DW. Acute kidney injury, mortality, length of stay, and costs in hospitalized patients. J Am Soc Nephrol. 2005;**16**(11):3365-70. 4. Landersdorfer CB, Kirkpatrick CM, Kinzig M, et al. Inhibition of flucloxacillin tubular renalsecretion by piperacillin. Br J Clin Pharmacol. 2008; **66**(5): 648-59. 5. Saitoh H, Oda M, Gyotoku T, et al. A beneficial interaction between imipenem and piperacillin possibly through their renal excretory process. Biol Pharm Bull. 2006; **29**(12): 2519-22. 6. Komuro M, Maeda T, Kakuo H, et al. Inhibition of the renal excretion of tazobactam by piperacillin. J Antimicrob Chemother. 1994; **34**(4): 555-64. 7. Santos Navarro M, Zarzuelo Castaneda A, Lopez FG, Sanchez Navarro A, Arévalo M, Lanao JM. Pharmacokinetic parameters of netilmicin and protective effect of piperacillin regarding nephrotoxicity caused by netilmicin. Eur J Drug Metab Pharmacokinet 1998;**23**(2): 143-7

VERSION 1 – AUTHOR RESPONSE

Response to reviewers
Thank you for taking your time for reviewing our paper.
Below we have displayed your comments with numbers and straight below, our response (R1 --Rn)

Response-general:
Response to all reviewers regarding endpoints (specific responses below):
We agree that dichotomous endpoints like RIFLE endpoints or AKIN can be very useful in ICU populations for detecting decreasing renal function. However, RIFLE endpoints do not capture “lack of recovery”, but rather a further decrease in renal function from the baseline level.
In our study (and often in severely infected ICU patients), a bacterial “hit” has resulted in acute onset renal failure, and this bacterial hit (and the related organ failure) is often the reason for ICU admittance in these patients. Renal function is thus lowest on average at ICU admission/baseline (see figure 1). With the correct treatment of the underlying infection, we expect renal function to recover. In such situations “lack of recovery” is a non-desirable situation, which may be very serious for the patient. We wanted to explore this, and realising, RIFLE (at least RIF) could not capture this (since they capture further deterioration from baseline and forth), we have used eGFR<60 ml/min/1.73 m2 as the primary endpoint and examined this from different angles (eGFR<60 ml/min/1.73 m2 at day 7, days with ml/min/1.73 m2 . RIFLE L and E could not be analysed, since they are relatively seldom and additionally since we followed patients regarding renal function for 28 days.
The multiple effects model was built to capture actual estimates of renal function improvement using different antibiotics and adjusting for other known or suspected. This has now been explained in the introduction.
We have additionally documented, that eGFR<60 ml/min/1.73 m2 is an independent predictor of 28 day mortality, when adjusted for other known and suspected predictors (table T1, supplementary

online material). Apart from high age, this variable is the strongest predictor of mortality in the model.

Reviewer: Natalie Ives
Senior Research Fellow
University of Birmingham, UK

No conflicts of interest

1) In the statistical analysis section the authors give various renal endpoints, but which is considered the primary endpoint is not specifically stated.

R1: Thank you for pointing this out. The primary endpoint in these analyses was 'estimated GFR<60 ml/min/1.73 m2.(please see above for discussion of this). This was looked at in several analyses: "Days with estimated GFR<60 ml/min/1.73 m2", "Risk of estimated GFR<60 ml/min/1.73 m2 within the first 7 days". This is now stated in the abstract and the methods:
"Statistical analysis
The primary endpoint was 'estimated GFR<60 ml/min/1.73 m2' and several analyses were made to explore this: 'days with estimated GFR<60 ml/min/1.73 m2', 'risk of estimated GFR<60 ml/min/1.73 m2 on day 1-7'. Secondary endpoints were a) delta eGFR after starting/stopping a drug, b) RIFLE-criteria Risk 'R', Injury 'I' and Failure 'F' www.adqi.net.Other endpoints explored were 'ever' blood-urea level ≥20 mmol/L and eGFR<30."

2) Summary: For a secondary analysis, the use of the words 'To determine...' seem a little strong. The original study was never powered to assess this, and they do state in the strengths and weaknesses that the observed numbers call for a slightly higher sample size.

R2: The phrasing has now been changed:
"To explore whether a strategy of more intensive antibiotic therapy with antibiotics not normally considered to be nephrotoxic leads to adverse renal outcomes in intensive care patients."

3) Abstract: Here they actually give the results from the multiple effects model, which is lacking from the main results.

R3: Thank you for pointing this out. A new table (table 3) has been made to report these results thoroughly. Previously, the main signal was communicated in figure 4, but we realize, that a table with these results gives a more nuanced insight into these results.

4a) Although I have answered yes to the first two questions above, this is based on the use of appropriate methods to address the question. However, I do think the results have been over-emphasised based on one highly significant result for the % of days within day 1-28 with eGFR<=60 ml/min/m2 ($p<0.0001$). It would be of interest to know how many patients this related to, which is how the other two renal endpoints are reported. Further, the other renal endpoints were less conclusive with $p=0.02$ and $p=0.04$.

R4a: The total number of patients meeting this endpoint is now reported so it is easier to get an overview over the magnitude of this effect. :
"The frequency of eGFR<60 ml/min/1.73 m2 on day 7 (or at death or last follow-up day) in the trial was 523/1200 = 43.6%."
Apart from this, see 4b)

4b: It would also have been useful to present some point estimates (odds ratios, relative risks)

alongside the 95%CI to assess the results.

R4b: Point estimates from the logistic regression have now been presented in a new table (table 4). Additionally, more details have been presented regarding the multiple effects model (table 3+other places).

5) Similarly in the discussion stating that this is "the first substantive evidence" seems a little strong considering the p-values, and with the lack of data presented on the regression analyses. Although in the conclusions the authors do state the need for further research.

R5: We have now presented the data from the regression analyses for both univariate estimates and multivariate, adjusted estimates (new table 3). Further, we have changed the phrasing to:
“To our knowledge, this study provides the first clinical report to inform this critical issue within ICU medicine.”

6) The statistics relating to Figure 3 are missing, what were the results from the statistical analysis. The plot is useful but it doesn't inform the reader fully.

R6: Statistics have now been added to figure 3 in the legend:
“Differences between eGFR in patients receiving piperacillin/tazobactam vs. meropenem: day 1 (p=0.78), day 2 (p=0.18), day 3 (p=0.09), day 4 (p=0.008), day 5 (p=0.001), day 6 (p=0.001), day 7 (p=0.0004), day 8 (p=0.005), day 9 (p=0.006), day 10 (p=0.02).”

7) I also don't find Figure 4 vey helpful, it would be more useful to present the results from the multiple effects model in a Table which is included in the main paper, this was not even provided as supplementary information.

R7: Figure 4 has been deleted. A table with these results is now presented (New table 3)

8) More statistics are presented for the sensitivity analyses than the main analyses, which makes it all a little confusing.

R8: The description of the sensitivity analyses has been reduced in the text.

9) Include information on the regression analyses (table of the results would be useful to the reader), and reduce the amount of information provided on the various sensitivity analyses.

R9: The data are now presented as suggested.

Reviewer: Mical Paul
Senior physician and Consultant
Unit of Infectious Diseases, Rabin Medical Center, Beilson Hospital
Israel

1) The current analysis does not have a primary outcome. Since the fact that the "high exposure" arm had a higher rate of renal failure has been previously reported, the new information in the current manuscript is on the association between type of antibiotic and renal failure. The methods for assessment of this question are not well described. See details in the comments to the authors below.

R1: Thank you for pointing this out. The primary endpoint in these analyses was ‘estimated GFR<60

ml/min/1.73 m2'. A discussion of this is placed above. Additionally, the choice of endpoint is argued in the introduction. The primary endpoint is now stated in the abstract and the methods:
"Statistical analysis
The primary endpoint was 'estimated GFR<60 ml/min/1.73 m2' and several analyses were made to explore this: 'days with estimated GFR<60 ml/min/1.73 m2', 'risk of estimated GFR<60 ml/min/1.73 m2 on day 1-7'. Secondary endpoints were a) delta eGFR after starting/stopping a drug, b) RIFLE-criteria Risk 'R', Injury 'I' and Failure 'F' www.adqi.net.Other endpoints explored were 'ever' blood-urea level ≥20 mmol/L and eGFR<30."

2) Careful attention is needed to the English language in the manuscript. For example, in abstract: "extreme bilirubin or triglycerides" and "persons held by force". Results regarding the association between type of antibiotic and renal failure are less credible because the analyses presented are not clear enough. See details in the comments to the authors below.
The current analysis is based on a well-performed and interesting study published in CCM and referred to in the current manuscript. The assessment of the outcome of renal failure is worthy.

R2: The manuscript has now been commented by a native English speaker, who has suggested several amendments. The mentioned phrases have been omitted.

3) Comparisons of ICU days with estimated glomerular filtration rate <60 mL/1.73 m2; Patients with estimated glomerular filtration rate <60 mL/1.73 m2; ICU days spent with dialysis treatment; and patients in treatment with dialysis at discharge/ death were reported in the original report of the trial in CCM.
Even though this is a secondary analysis of a RCT, I believe that a primary outcome measure for the current analysis should have been defined and used. Currently, conclusions are based on the variable with most significant results and not necessarily the most important variable for the patient.
Considering patient-relevant outcomes I would regard as most important a dichotomous variable of death or discharge alive with RIFLE criteria of loss or ESRD, which was not assessed in this study.
Looking at the original trial report it seems that there is no difference between the groups with regard to this outcome measure.

R3: Please see above in the general discussion of endpoint.
As correctly stated, the results on the dichotomous endpoints are much less convincing (table 3). However, there is a consistent tendency towards more renal failure in the 'high exposure group', in some analyses reaching significant levels. We want to report this accurately to make clear for the reader that we cannot at present make a statement of whether 'high exposure' to the used antibiotics does cause de novo renal failure.(We cannot neither reject nor confirm this within the number of RIFLE endpoints that were observed).
To account for the nuance between emergence of renal failure vs. prolongation of existing renal failure, we have further emphasized in the concluding remarks in the discussion, that this seems to be a reversible nephrotoxicity, and that data on the dichotomous endpoints are not clear.
Additionally, after analysing the eGFR´s from day-to-day after starting different drugs and adjusting these results in the multiple effects models, we observed that renal function did, in fact, recover rapidly when piperacillin/tazobactam was discontinued (now reported in table 3), so we would not expect ESRD to be more frequent in the 'high exposure group' or in general, when treating with piperacillin/tazobactam. However, we do not have sufficient follow up to conclude on the matter of ESRD.
"Conclusion
In conclusion, the use of piperacillin/tazobactam caused a delayed renal recovery in critically ill patients, and renal function improved after discontinuation of the drug. However, the study is not designed to investigate de novo emergence of renal failure, since the lowest renal function is at baseline in most patients. We cannot within the sample size and follow-up time of this trial establish

whether the use of piperacillin/tazobactam, in some cases causes persistent renal failure, and thus, further research to explore this is warranted. We think this impact on renal function is more likely caused by a toxic effect on the renal tubule than by a lack of effect towards the infection, since this drug is independently associated with a high chance of survival in other infected populations8, and we must emphasize that our findings are strictly confined to critically ill patients..

4) Study outcomes: what about RIFLE categories loss and ESRD? Only risk, injury and failure are defined in methods (these should be spelled out).

R4: We agree, it would be interesting to follow these patients for a longer period of time. For the current study, however, we did only follow the patients for 28 days. Thus, it was not possible to assess RIFLE criteria Loss and ESRD. RIFLE L and E demand follow up for more than 4 weeks after the renal dysfunction has emerged.

5) The comparison between antibiotics is not based on randomization and is thus an observational finding. As such, the analysis assessing the association between antibiotics and renal failure is not presented in enough details in methods and the results presented questionable. An appropriate analysis of this association would be complex because patients received concurrently more than one antibiotic, as per the antibiotic guidelines presented in the main manuscript. Furthermore, this association should have been adjusted to other risk factors for renal injury and this adjustment although probably performed is not well presented in methods and the adjusted analysis is not presented in tables.

R5: We did not communicate this clear enough:
This is now stated in the new table 3 , legend:
“The multiple effects models were adjusted for treatment arm (‘low exposure’ vs. ‘high exposure’), gender, age (>=65 vs. <65 years), APACHE II score (>=20 vs. <20), Clinically judged infection (severe sepsis/septic shock vs. milder or no infection), patient category (surgical vs. medical) and eGFR level at administration of the antibiotic, (1: <30 ml/min/1,73 m2, 2: 31-60 ml/min/1,73 m2, 3: >60 ml/min/1,73 m2) as well as in the statistics part of methods.”
Additionally, the results from the multivariate logistic regression analysis have been presented more clearly in the text:
“The frequency of eGFR<60 ml/min/1.73 m2 on day 7 (or at death or last follow-up day) in the trial was 523/1200 = 43.6%. This endpoint was investigated in a forward censored (p<0.2) logistic regression. Use of piperacillin/tazobactam and other frequently used beta-lactam drugs for at least three days within these first seven days, as well as known and suspected predictors of renal failure were explored in a multivariable logistic regression analysis. Five independent predictors of renal failure on day 7 were identified: Age above 65 years, APACHE II score >20, Charlson´s co-morbidity score >=2, estimated GFR at baseline and use of piperacillin/tazobactam for at least 3 days within the first 7 days (table 4)”
And these results are presented with point estimates in a new table (table 4).
The multivariable model was conducted in a forward censored manner (p<0.2) with the following known and suspected predictors of renal failure: Age (>=65 y), APACHE II (>=20), Severe sepsis/shock vs. milder infection), Auto-immune disease (Y/N), Cancer (Y/N), Charlson´s comorbidity score (>=2 vs. <2), Surgical patient (Y/N), Body Mass Index >=25 vs. <25, Gender (M/F), eGFR level at baseline (<30 ml/min/1.73 m2, 30-60 ml/min/1.73 m2, >60 ml/min/1,73 m2), Use of the following beta-lactam antibiotics for >=3 days within the first 7 days after baseline: piperacillin/tazobactam, meropenem, cefuroxim.
In the final model (after forward censoring), the following were included: 1) Age (+65 y), 2) APACHE II (>=20), 3) Severe sepsis/shock vs. milder infection), 4) Charlson´s comorbidity score (>=2 vs. <2), 5) Body Mass Index >=25 vs. <25, 6) Gender (M/F), 7) eGFR level at baseline (<30 ml/min/1.73 m2, 30-60 ml/min/1.73 m2, >60 ml/min/1,73 m2), Use of the following beta-lactam antibiotics for >=3 days

within the first 7 days after baseline: 8) piperacillin/tazobactam, 9) cefuroxim.

6) The authors should be commended for considering death or renal failure as an adverse outcome (and not renal failure alone). The handling of deaths should be presented more clearly in the methods section.

R6: This has now been included in the methods section:
“Mortality was followed via the National Patient Register in which all deaths in Denmark are registered within 14 days.”

7) Tables 1 and 2 and figures 1 and 2 are similar to or the same as those presented in the main report of the trial in CCM and are probably redundant here. Rather than the “all” column in table 1, p-values would have been more helpful

R7: We agree, this has been reported earlier. However, we think these data are important for the understanding of the current paper. Therefore, tables 1 and 2 and figures 1 and 2 have been removed and are now supplied as “supplementary digital material”.

Reviewer: Assoc. Prof. Michelle S Chew
Departments of Intensive Care Medicine and Clinical Sciences Malmö
Malmö University Hospital
Lund University
Sweden

It was a pleasure to review this paper by Jensen J-U et al. Please find my detailed comments in the attached file.
I declare that my institution has been the recipient of an unrestricted grant from Brahms AG, the manufacturer of the Procalcitonin assay used in this study.

The present study is an extension of the original PASS study by Jensen et al published recently (1). This study showed that an antibiotic escalation algorithm based on repeated Procalcitonin (PCT) measurements did not improve 28 day survival. Further, despite substantially higher use of broad-spectrum antibiotics, the study identified several unexpected effects of this strategy: increased time spent on mechanical ventilation, increase length of ICU stay and increased number of days with renal failure defined as eGFR <60ml/min/1.73m2.
Based on these findings, the aim of the present study was to further delineate the effect of the PCT strategy (high exposure) on various markers of renal function, compared to the current ‘standard of care’ (standard exposure).

This secondary analysis of the PASS data is relevant because renal failure occurs commonly in the critically ill and is associated with adverse outcomes.
The study appears to have been carefully designed and the escalation protocol for the ‘high exposure’ group is relevant for the specified critically ill population.

Kidney injury end-points were defined as 1) Risk (R), Injury (I) or Failure (F) according to the RIFLE criteria; 2) eGFR<30 or 60 ml/min/1.73m2 corresponding to moderately-severe and severely impaired renal function respectively; or 3) BUN>20mmol/L at any time during the 28-day follow-up period. Further, the duration of renal failure based on eGFR measurements was also analyzed.

The main results were that a ‘high exposure’ strategy of antibiotic escalation based on PCT

measurement was associated with adverse renal outcomes. Specifically, a multivariate analysis identified that this effect was due Piperacillin-Tazobactam (Pip-Tazo). Supporting the suggestion of a nephrotoxic effect of Pip-Tazo was the rate of improvement of eGFR when the drug was discontinued.

The paper is generally well written (only minor language revisions required), the hypotheses clearly stated, and the methods specified in detail in the accompanying supplement. The authors have obviously considered the study protocol carefully, and use appropriate statistical methods to analyze the data. I do however have a number of concerns that I think the authors should address.

Choice of end points
1) One may question the utility of BUN as a renal end-point in this population. BUN is often increased in the critically ill due to parenteral feeding and catabolism and does not accurately reflect renal function.

R1: We agree. This is why we only used this in the analysis of differences between the randomized arms, since confounding due to nutritional strategies, catabolism etc. can be ruled out (because we are looking at a randomized comparison). These results are not given much attention, however, we do think it is important to explore the renal function from different angles.

2) Another important limitation is using eGFR in this population since this measurement is only valid in steady state. Since eGFR measurement forms the bulk of evidence behind the conclusions in this study, I suggest that the authors clearly state and discuss its limitations.

R2: eGFR as defined by Cochcroft & Gault does not always reflect a precise estimate of creatinine clearance in ICU patients. However, we do think, as documented by others(1), that changes in eGFR reflect changes in renal function, at least on a population basis. We did not have the opportunity to make daily assessments of creatinine clearance in these patients. Additionally, see above in the general response regarding endpoints.
Second, eGFR may not be an accurate measure of creatinine clearance, as recently documented by Martin et al. (22). However, even though this measure is not accurate to describe the creatinine clearance, changes in eGFR reflect changes in renal function, as validated, and is closely correlated to outcome(23). Additionally, we found that eGFR<60 ml/min/1.73 m2 on day 7 is a strong independent predictor of mortality (table T1).

3) Finally, the choice of the RIFLE criteria instead of the AKIN criteria may be questioned since the AKIN criteria reflect even smaller changes in renal function, which are associated with poorer outcomes (2,3), and classifies patients with RRT into stage 3 (≈F in RIFLE). The authors should include these limitations in their discussion since it using the AKIN or other criteria than eGFR may potentially change their results and conclusions.

R3: This is a secondary endpoint. We chose the RIFLE criteria, since we have not been able to find convincing evidence, that AKIN criteria offer crucial improvement, as also presented in a recent meta-analysis by Bentley et al.3

4) Except for renal injury defined as the ‘risk (R)’ group, the number of patients with AKI was not significant between groups according to the RIFLE criteria. This is inconsistent with the eGFR findings. The authors should discuss this further, taking into account the limitations of the chosen end-points.

R4: Rifle R is a secondary endpoint. See discussion of the choice of primary endpoint above in the general discussion of this. This discussion also captures the reasons not to expect substantial

differences in RIFLE endpoints.

5) Higher exposure, sicker patients, or both?
The high exposure group indeed had significantly increased Defined Daily Doses (DDDs) of broad spectrum antibiotics administered and significantly more days spent in ICU with at least three antimicrobials. There were significant increases in the DDDs of Piperacillin-tazobactam in the high exposure group (2925 vs 1893), as well Ciprofloxin although the difference with the latter antibiotic was much less marked (8382 vs 6210 DDDs).

While the data is clear that increased exposure to antibiotics were associated with adverse outcomes, it is not clear whether or not this was due to a greater severity of illness in patients at the time of antibiotic escalation in the high exposure group. Was Pip-Tazo prescribed to sicker patients and with already higher degrees of renal impairment? Is it possible the kidney injury occurred more commonly in the 'high exposure' group due to more severe underlying disease? While similar 28-day mortalities in the two groups speak against a difference in illness severity, I note that the 'high exposure' group spent more time in ICU, and in their original study the authors noted a tendency for increased vasopressor and inotrope requirements and longer periods with severe sepsis and septic shock (1). eGFR multivariate analysis was adjusted for APACHE and degree of host response at baseline, but no adjustments for continuing organ failure were made. Do the different antibiotic groups differ in terms of SOFA, need for RRT and other important diagnoses such as sepsis for example, at the time of antibiotic administration/+ escalation? Notably eGFR at the initiation of the investigated antibiotic was included in the multivariate analysis. The authors should be commended for their foresight in including this variable, however, my previous concerns regarding using eGFR are valid even here.

R5: Patients were randomised to these two different antibiotic strategies. Baseline characteristics were comparable. If patients thus at a later point in follow-up come to suffer from different organ failures, this is caused by the interventions.
Point estimates of the logistic regression have been added (table 4). Estimates have been adjusted for influence from other factors, that are considered to possibly influence renal function.

6) A substantial number of days (slightly less than 30% in both groups) for the assessment of renal failure were not included. The authors have mentioned this as a limitation of their study. Whilst the authors have noted that the rate of renal failure on the last day of follow-up in the two groups were comparable, I am uncertain how it reflects a 'temporary extension of the duration of renal failure in the high exposure group'. Given the magnitude of the findings in the 28-day follow up, I would have expected the rate of renal failure to be higher even at this timepoint. Can the authors explain this, as well as how the data reflect a 'temporary extension of renal failure'?

R6: Please see the general discussion of choice of endpoint.

7) Another query relates to the use of a 10-day eGFR to indicate recovery of renal function. If 28-days of follow-up data are available, why limit the recovery period to 10 days?

R7: The main objective of these analyses was to find out whether there was a difference in the rate of renal function improvement on the first days after a) administration of the different drugs, b) discontinuation of the drugs. If one drug showed a certain pattern, this would argue, that this drug had a special impact on renal function improvement. Thus, the first days after administration resp. discontinuation were our main interest.

8) Kidney injury is known to adversely affect survival in critically ill patients, therefore it is surprising that despite the striking results seen in the present study, no mortality effect was observed. Do the authors have a possible explanation for this?

R8: We have discussed this in the study group. Our main explanations are:
1) Kidney failure can be either caused by a) generalized breakdown of organ functions caused by universal damaging pathofysiology like bacterial toxins, systemic inflammation, endothelial dysfunction, mitochondrial dysfunction etc, eventually leading to tissue apoptosis. In this situation, renal failure may be the first manifest organ dysfunction detected, although this should not be perceived as an isolated effect on the kidney. In such a situation, prognosis is naturally influenced heavily in negative direction;) or b) An isolated toxic (reversible/ non-reversible) effect on the kidney. In such a situation, the prognosis may not be influenced to the same degree, since other organs, crucial cell processes and thereby tissues may function normally. If b is the case, this could be the explanation, of why the prognosis between the two groups is not different.
2) The procalcitonin-guided pro-active antibiotic strategy may have improved the prognosis for some patients, providing adequate antibiotics at an earlier time point and opposite have caused renal failure in others, this having a negative effect on the prognosis. These two oppositely directed effects may have resulted in an overall “no effect”

9) The authors write that this large clinical study suggests that Pip-Tazo in itself is nephrotoxic, and that this finding is supported by other studies in experimental settings. This is not accurate. Previous studies show that Piperacillin reduces elimination of other antibiotics and increase the toxicity of other antibiotics ostensibly by affecting renal clearance (4-7).
The text should be changed to reflect this. As an extension to this, have the authors considered that co-administration of Pip-Tazo with other drugs may have caused a nephrotoxic effect due to the decreased elimination of these other drugs?

Finally, the study design does not allow for a conclusion to be made regarding cause and effect. The authors should state this clearly.

R9: Since the present results are based on a randomized design, where more days with renal failure occur in the ‘high exposure’ group, it can initially be concluded that this is caused by some part of the intervention. Exploring this further in multiple effects models, the lowest renal recovery was observed in patients receiving piperacillin/tazobactam (compared to the four other most used antibiotics).In these analyses, adjustment was made for other factors suspected to influence renal function, and the baseline renal function. The results were confirmed in logistic regression analysis. Entering combinations of drugs (all the beta-lactam drugs as mono-therapy and additionally in combination with ciprofloxacin) did not change the estimate or statistics. Starting pip/tazo reduced the rate of renal recovery and discontinuing the drug resulted in a high rate of eGFR improvement.
So we find reason to conclude, the observed effect is attributable to piperacillin/tazobactam. The experimental studies with healthy volunteers suggest a plausible mechanism for this, but of course, we did not perform these ourselves, so they only give input for the discussion. Our conclusions are based on our own data.

Strengths of this study
Notwithstanding the above concerns, this study has some important, clinically relevant strengths. Firstly, the conclusion that Pip-Tazo adversely affects renal outcome is supported by the analysis that eGFR increases after discontinuation of the drug. Secondly, excluding patients with early deaths (within 7 days), fungal infections and combined β-lactam+fluorquinolone exposure did not change the conclusions. Finally, the adverse effect on renal function was seen only with Pip-Tazo but not Cefuroxime and Meropenem.

Whilst a cause-effect relationship cannot be proven due to the design of this study (and this should be highlighted by the authors), the findings ring a warning bell for the administration of Pip-Tazo on an

escalation regime based on PCT levels. This study should catalyse future studies investigating the possible nephrotoxic effect of Pip-Tazo, including the use of various dosing regimes (eg. extended infusions), and co-administration with other antibiotics.

References

1. Jensen JU, Hein H, Lundgren B et al. for the Procalcitonin and Survival Study (PASS). Procalcitonin-guided interventions against infections to increase early appropriate antibiotics and improve survival in the intensive care unit: A randomized trial. Crit Care Med 2011;39(9):2048-58
2.Barrantes F, Tian J, Vazquez R, Amoateng-Adjepong Y, Manthous CA. Acute kidney injury criteria predict outcomes of critically ill patients. Crit Care Med. 2008;36(5):1397-403.
3. Chertow GM, Burdick E, Honour M, Bonventre JV, Bates DW.
Acute kidney injury, mortality, length of stay, and costs in hospitalized patients. J Am Soc Nephrol. 2005;16(11):3365-70.
4. Landersdorfer CB, Kirkpatrick CM, Kinzig M, et al. Inhibition of flucloxacillin tubular renalsecretion by piperacillin. Br J Clin Pharmacol. 2008; 66(5): 648-59.
5. Saitoh H, Oda M, Gyotoku T, et al. A beneficial interaction between imipenem and piperacillin possibly through their renal excretory process. Biol Pharm Bull. 2006; 29(12): 2519-22.
6. Komuro M, Maeda T, Kakuo H, et al. Inhibition of the renal excretion of tazobactam by piperacillin. J Antimicrob Chemother. 1994; 34(4): 555-64.
7. Santos Navarro M, Zarzuelo Castaneda A, Lopez FG, Sanchez Navarro A, Arévalo M, Lanao JM. Pharmacokinetic parameters of netilmicin and protective effect of piperacillin regarding nephrotoxicity caused by netilmicin. Eur J Drug Metab Pharmacokinet 1998;23(2): 143-7

Response references:

1. Martin JH, Fay MF, Udy A, Roberts J, Kirkpatrick C, Ungerer J, et al. Pitfalls of using estimations of glomerular filtration rate in an intensive care population. Intern Med J. 2011.
2. Bagshaw SM, George C, Dinu I, Bellomo R. A multi-centre evaluation of the RIFLE criteria for early acute kidney injury in critically ill patients. Nephrol Dial Transplant. 2008; 23(4): 1203-10.
3. Bentley W. Towards evidence-based emergency medicine: Best BETs from the Manchester Royal Infirmary. BET 3: RIFLE criteria versus Acute Kidney Injury Network (AKIN) criteria for prognosis of acute renal failure. Emerg Med J. 2011; 28(10): 900-1.

VERSION 2 – REVIEW

REVIEWER	I declare that my institution has been the recipient of an unrestricted grant from Brahms AG, the manufacturer of the Procalcitonin assay used in this study. An ICMJE Conflict of Interest Form has previously been submitted. Michelle Chew Assoc. Prof. Department of Intensive Care Medicine Malmoe University Hospital, Sweden
REVIEW RETURNED	28/01/2012

GENERAL COMMENTS	General: The end-points are now clearly defined. I appreciate the new data on the univariate estimates, and the point estimates from the logistic regression analysis as well as the detailed information on the multiple effects model. Endpoints:

	The fact that RIFLE criteria was used as a secondary endpoint does not diminish the potential importance of its limitations as a measure of renal function. RIFLE is heavily dependent on baseline Creatinine values which are difficult if not impossible to obtain in this population. This is not a trivial issue since the reference Creatinine value can change the mortality estimate (1). Further, studies show that even absolute increases as small as 0.3 mg/dl are associated with adverse outcomes (2). These are factors that could possibly affect the results of the present study, and may have accounted for the finding that only renal injury defined as 'risk' but not 'injury' or 'failure' was statistically different between groups. Although I agree that there is at present a lack of evidence of the superiority of AKIN vs RIFLE criteria, AKIN is now the accepted standard for classification of kidney injury in critically ill patients. Added to the limitations of using eGFR as the primary endpoint in this population, I believe that it would be naïve not to mention these limitations. In the abstract 'Risk' is now given as a secondary endpoint, as opposed to R,I,F stated in the original manuscript and in the methods section. Which is the case? In the case of the former, was 'Risk' was chosen post-hoc given that it was the only one of the RIFLE criteria that was statistically significant? Comparison with other studies: Although it is attractive to conclude that Pip-Tazo in itself causes nephrotoxicity, there are a multitude of other factors that have not (and cannot) be analyzed in the present study. Previous evidence show that Piperacillin reduces the elimination of other antibiotics, and increases their toxicity ostensibly by affecting renal clearance (see comment 9) in my previous review). Therefore the use of Piperacillin may not be neprotoxic per se, but increases the likelihood of nephrotoxicity caused by other drugs. This study generates the interesting hypothesis that Piperacillin-Tazobactam may in itself be nephrotoxic in this population. However, I think that the conclusion '…plausible that piperacillin specifically causes nephrotoxicity' is an overstatement in the context of the present findings. 1) Cruz DN, Ricci Z, Ronco C. Clinical Review: RIFLE and AKIN-time for reappraisal? Crit Care 2009;13:211 2) Chertow GM, Burdick E, Honour M, Bonventre JV, Bates DW: Acute kidney injury, mortality, length of stay, and costs in hospitalized patients. J Am Soc Nephrol 2005, 16:3365-3370. Michelle Chew Assoc. Prof. Department of Intensive Care Medicine Malmoe University Hospital, Sweden

VERSION 2 – AUTHOR RESPONSE

General: OK.

Endpoints: The limitation of the RIFLE criteria has now been mentioned in the discssion:
"Third, the RIFLE criteria used as secondary endpoint measures are not suitable to detect renal failure from baseline and forth, since the reference is defined as the pre-morbid creatinine. Hence, renal failure caused by exposure to antibiotics beginning at baseline, will not necessarily be captured using these criteria. This was the reason for not using these as primary endpoints."

Abstract:
Risk, Injury and Failure are all analysed and reported. This has been corrected in the abstract:

"Main outcome measures: Primary endpoint: estimated GFR<60 ml/min/1.73 m2. Secondary endpoints: a) delta eGFR after starting/stopping a drug, b) RIFLE criterion Risk “R”, Injury ‘I’ and Failure ‘F’. Analysis was by intention to treat. "

Conclusion
This has been modified:
"The study was not designed to establish whether the use of piperacillin/tazobactam or other of the interventional drugs, in some cases cause persistent renal failure, and thus, further research to explore this is warranted. We think this impact on renal function is more likely caused by a – at least partially reversible - toxic effect on the renal tubule than by a lack of effect towards the infection, since this drug is independently associated with a high chance of survival in other infected populations8, and we must emphasize that our findings are strictly confined to critically ill patients. "