Dear Editor,

Thank you for having evaluated our manuscript. We have appreciated the Reviewers' comments that were very useful to improve the quality of the paper. We have addressed them all and you can find our answers to the Reviewers' comments at the bottom of this letter.

Statistical analysis was corrected according to Biostatistiscian comments. English language was improved and corrected by a professional English speaker.

Modifications in the manuscript were indicated in the response to reviewers and highlighted in yellow color for the ease of your review.

Moreover, we have provided certifications for quality both of the English language (grammar and style) and of the statistical analysis. A qualified statistician and a mother tongue translator have revised the manuscript and suggested some amendments.

Sincerely,

Carlo Torti, MD

Reviewer #1

- Comment #1
  - The motivation for the study can be improved. It's unclear why clinicians or policy makers should be interested in NGAL as a diagnostic tool? Is NGAL less expensive than other tests? Does it predict severe outcomes more accurately? If extreme values NGAL indicate toxicity, should extreme values be interpreted as a reason to discontinue treatment? Since NGAL is traditionally a measure of kidney function, what is the current clinical or theoretical understanding of kidney function and HCV infection? If there is a relationship between eGFR and NGAL, what is the relevance of eGFR to HCV infection status, diagnosis, liver disease, or HCV-related prognoses?
  - We agree in general that it is important to consider utility, cost and cost-effectiveness of the proposed test but literature data are scarce and not sufficient to provide clear recommendations. However, in principle, NGAL may help diagnose tubular dysfunctions that may be associated with HCV and/or with HCV therapy, more efficiently that simple creatinine can do. In line with these considerations, we have modified the text at page #8, lines #117-124 as follows:
    - ***Moreover, in the context of liver disease, kidney function is one of the key predictors of death and serum creatinine is a component of both King's college criteria and model for end-stage liver disease scoring systems that are used for prognostic stratification in patients with acute and chronic liver failure.***
    - ***Virus-related kidney manifestations mainly show up as a glomerular impairment, predominantly a membrano-proliferative glomerulonephritis (MPGN) with type 2 crioglobulinaemia and sub-endothelial or intra luminal deposits of IgG, IgM and complement components. Moreover, HCV core proteins were isolated in both glomerular and tubular tissues, suggesting the presence of a parallel tubular-interstitial damage.***
  - Moreover, at pages #19 and 20, lines #375-381 we have added the following paragraph:
    - ***However, many question points still remain to be answered. Indeed, it has to be established whether a clinical cut-off of NGAL may guide clinical decisions (e.g., dosage modification or stopping of the offending drug). Also, it has to be evaluated whether NGAL could predict AKI during HCV treatment, especially in most-at-risk patients such as those with advanced cirrhosis and a high risk of renal complications (e.g., hepato-renal syndrome). Lastly, cost-effectiveness studies need to be conducted to verify the hypothesis that NGAL should be routinary used to monitor kidney function during HCV treatment instead of (or in addition to) creatinine.***
- Comment #2
  - This empirical support for the study's primary conclusion is weak. The only stated conclusion, apart from that the clinical use of NGAL is not clear, is that NGAL could provide complementary information to eGFR. But this conclusion seems contradicted by the data presented in Figure 4, which shows minimal resemblance between NGAL and eGFR. In addition, the primary piece of evidence that seems to support this one conclusion seems to be found in panel B of Figure 2, which shows a very subtle downward slope, significant at 4.88%, between NGAL and eGFR. Unfortunately, the limited number of observations in this study does not warrant much confidence in this measurement. While statistically significant, this relationship could be due random chance or a mixture of random chance and some other small effects, such as age, that are being mediated by eGFR. Stronger evidence for this conclusion would include a much more coherent conceptual model of inflammation paired with an empirical model that estimates the relationship between

NGAL and eGFR while controlling for potential confounders identified in the conceptual model such as age, gender, hcv genotype, and treatment status.

- We agree with the limitations underlined by the Reviewer. In fact, we have deleted the conclusions that "creatinine clearance and NGAL could provide complementary information". We only have retained the observation that the two methods are discordant in the results provided, which appears to be supported by data when analyzed categorically. With reference to this issue, please see at page #17, lines #309-310:
  - ***These results suggest that a discordance between the two methods exist when interpretation is "categorical".***

- Comment #3
  - Title. The title does not represent the content of the paper. The title should be something like "Is neutrophil gelatinase useful in assessing the health of patients with hepatitis C infection?" In particular, I don't think paper's title should make reference to "before and after therapy containing DAAs". There's just too few observations of patients under treatment to justify stating treatment as a prominent component of the study. There's only 8 observations of patients under treatment, so at most, a discussion of the associations before and after treatment should be a small part of the results.
  - We agree and have modified the title that now reads as follows:
    - ***Is neutrophil gelatinase associated lipocalin useful in hepatitis C virus infection?***

- Comment #4
  - Abstract. Remove or revise the following statement from the abstract "Not statistically significant differences were demonstrated after 1 year." This statement doesn't reference any variables, so its impossible to tell what the "differences" are between or in reference to.
  - We have clarified the meaning of the sentence as follows:
    - **Not statistically significant predictions of NGAL at baseline were demonstrated for eGFR evolution 1 year later.**

- Comment #5
  - Please justify the choice of 65 years as a threshold age. Is the justification based on the median or mean value of the age variable? If not, the justification should include some biological or conceptual reasoning for choosing 65, instead of any other arbitrary value such as 60, or 50, or 70. If the choice of 65 is arbitrary, then additional sensitivity analyses should be conducted comparing NGAL values across different age thresholds.
  - We have provided justification of the choice of 65 yro for age (see page #13, lines #225-226) based on the reference provided [*Delanaye P, Glassock RJ, Pottel H, Rule AD: An Age-Calibrated Definition of Chronic Kidney Disease: Rationale and Benefits. Clin Biochem Rev. 2016 Feb;37(1):17-26*]:
    - ***Cut-off for age was set at 65 years because that is the threshold discriminating adulthood and elderly life in many western countries.***

- Comment #6
  - "…an univariate" should be "a univariate"
  - This typo has been corrected

- Comment #7
  - Can remove "at linear regression analysis"

- OK, we have removed it.

- Comment #8
  - Figure 2. Replace Figure 2 with a table that presents the univariate model results, coefficients, standard errors, p-values, R2, and observation counts. In particular, why does it look like there are a lot fewer FIB4 observations than the others? Figure 3.
  - We'd rather keep Figure 2 because we feel that a graph is more immediate to read than a table. You were right since we made a mistake in depicting fewer FIB-4 observations. However, in the present version of the paper we have represented correlations between plasmatic NGAL and eGFR at baseline (panel A), and plasmatic NGAL and age (panel B). Figure 2 was modified according to Spearman's test results.

Reviewer #2

- Comment #1
  - I think the title is misleading, because the majority of patients included in the study were not evaluated before/after DAA therapy. Additionally, it appears from the manuscript that NGAL levels were obtained during therapy, but not after therapy
  - The title has been modified according with the Reviewer's suggestions as follows:
    - ***Is neutrophil gelatinase associated lipocalin useful in hepatitis C virus infection?***

- Comment #2
  - I think a brief explanation of the differences between plasma and urinary NGAL would be helpful, and then maybe later they should indicate why they chose plasma NGAL instead of urinary NGAL
  - We have added the requested specifications at page #9, lines #149-152 as follows:
    - **NGAL urinary levels are increased after a tubular injury and may reveal a kidney damage earlier than an increase of creatinine. On the other hand, plasmatic or serum NGAL are more extensively adopted in AKI contexts, because their measurement is less limited by samples availability when patients are anuric.**

- Comment #3
  - Is there something that makes NGAL particularly useful in cirrhotic patients?
  - Although the data are scarce, there is a line of evidence suggesting that NGAL is useful in cirrhotics as dicussed in the paper by Firu and Collaborators (*J Med Life*. 2015; 8 Spec Issue: 15-20). So we have added the following sentence with the aforementioned reference at page #9, lines #156-157:
    - ***In cirrhotic patients, NGAL is a marker of AKI and urinary NGAL can help distinguish among different causes of renal impairment.***

- Comment #4
  - Why was an HCV VL of 1,000,000 used as a cut-off?
  - We have now provided a reference to support the choice of 1,000,000 copies/ml as a cut-off for "high" HCV RNA and a reference for example (see page #13, lines #224-225):
    - ***Viral load cut-off was set at 1,000,000 HCV RNA copies/ml because this value is been commonly considered to be high [Eur J Gastroenterol Hepatol.* 2001 Feb;13(2):149-55.]**

- Comment #5
    - Why only perform a univariate analysis? The factors that were felt to be significant should be included in a multivariable model
    - We agree that a multivariable model should be performed but, unfortunately, in our opinion the sample was not large enough to perform this analysis.
- Comment #6
    - Need to define worsening renal failure over time. Did you look at any change in GFR, or a specific cut-off?
    - We have defined worsening of renal failure over time as any change in eGFR (see page #13, line #222-223).
- Comment #7
    - Paragraph 2 is difficult to read. If the findings were non-significant, I think you can just say that all p-values were >0.05 and not include each p-value for each comparison
    - You are right. For the sake of clarity we have modified the paragraph which now reads as follows (see page #14, lines #253-260):
        - ***Differences in NGAL were not significant among patients ranked by HCV viral load, FIB-4 score and APRI. Quantitative HCV RNA was available in 36 patients, 19 (53%) of them having HCV RNA<1,000,000 copies/ml (i.e. low HCV RNA group). Median NGAL at baseline was 70 ng/dl (range: 132-27) versus 63 ng/dl (121-28), respectively. For FIB-4, 8 (17%) patients with FIB-4 ≤1.45 had median NGAL of 60.5 ng/dl (111-27), 16 (33%) with FIB-4 from 1.45 to 3.25 had median NGAL of 82 ng/dl (136-36) and 24 (50%) with FIB-4 ≥ 3.25 had median NGAL of 74.5 ng/dl (132-28). Regarding APRI, 10 (21%) patients were ≤0.5, 22 (46%) were between 0.5 and 1.5, and 16 (33%) were ≥1.5. Median NGAL values were 60.5 ng/dl (range: 136-27), 74.5 ng/dl (124-36) ng/dl and 80 ng/dl (132-28) in the three APRI groups, respectively.***
- Comment #8
    - I am not sure what the linear regression adds to the manuscript, the p-values are really not significant if rounded to conventional two decimal places. I would not make much of these conclusions. To that effect, I don't think Figure 2 is informative.
    - You are right that results are not significant from a statistical point of view (although with a more appropriate statistical test significance has been reached, especially for the correlation between NGAL and eGFR). However, we prefer to convey to the readers these results as also indicated by the Reviewer #1 who suggested to insert a table with the detailed results (but we'd rather keep a Figure because we feel that a graph is more immediate to read than a table). So, we propose to represent in the form of graphs only significant correlations between plasmatic NGAL and eGFR at baseline (Figure #2, panel A), and plasmatic NGAL and age (Figure #2, panel B).
- Comment #9
    - Is there any clinical significance to having an NGAL greater than limit of detection? This should be explained. Authors could consider adding a column to their first table for those patients with baseline NGAL >118 and in the text just state that all comparisons were non-significant with p-values > 0.05, rather than list everything in the text

- We could not state whether there is a clinical significance in having an NGAL greater than limit of detection, we hypothesize that an NGAL greater than limit of detection means that a tubular impairment is present. We did not answer a third column to table 1 because the number of patients with an NGAL greater than limit of detection was to small and we did not perform any analysis. Thus, we considered that a third column could have been confounding. We modified the text as you suggested (see page #15, lines #271-280)
    - ***In 6 (12.5%) patients, NGAL exceeded the upper limit of the reference interval (NGAL>118.11 ng/ml). When patients with NGAL>118.11 ng/dl were compared with patients with NGAL≤118.11 ng/dl, not statistically significant differences were present for age, gender, CKD classification and liver fibrosis (p >0.05). In fact: 6/6 (100%) patients with NGAL>118.11 ng/ml versus 25/42 (60%) patients with NGAL ≤118.11 ng/ml were ≥65 years-old ; 3/6 (50%) patients with NGAL >118.11 ng/ml versus 20/42 (48%) patients with NGAL≤118.11 ng/ml were males; 1/6 (17%) patients with NGAL>118.11 ng/ml versus 7/42 (17%) patients with NGAL≤118.11 ng/ml had a KDIGO CKD≥G3a; 4/6 (67%) patients with NGAL>118.11 ng/ml versus 26/42 (62%) patients with NGAL≤118.11 ng/ml had a KDIGO CKD classification ≥G2; 4/6 (67%) patients with NGAL>118.11 ng/ml versus 19/42 (45%) patients with NGAL≤118.11 ng/ml had a FIB-4≥3.25; 3/6 (50%) patients with NGAL>118.11 ng/ml versus 13/42 (31%) patients with NGAL≤118.11 ng/ml had an APRI≥1.5.***

- Comment #10
    - In the section looking at renal parameters after one year, I don't think the outcome of "worsening eGFR" and "stable/improved eGFR" over 1 year was defined. Did they use any change, or a degree of change / category change? Why not perform a linear analysis for this?
    - You are right. We defined the outcome at page #13, lines #222-223.
        - ***Patients were ranked by eGFR worsening versus stable/improved with respect to baseline values after 1 year. Any reduction of eGFR was considered as worsening.***

- Comment #11
    - Although technically telaprevir is a DAA, the term DAA has really come to represent second/third generation agents and not first generation PIs. I think the use of the term DAA is misleading.
    - We did not remove the term DAA when it was used to define telaprevir, because telaprevir is (together with boceprevir) the first antiviral drug for the treatment of HCV with a direct antiviral action.

- Comment #12
    - Did patients achieve SVR, or are they still on therapy? I think it would be more interesting to assess over longer term, after completion of therapy. It appears that at the time of this analysis, patients were still on therapy?
    - You are right but, unfortunately, we perform NGAL measurement only during the first 12 weeks of treatment. We specified in the text that patients were still on treatment (see page #13, lines #231-233)
        - ***Evolution of NGAL at different time points (baseline, week 4and week 12) was evaluated separately in patients who started antiviral treatment with DAA containing regimens during the first twelve weeks of therapy.***

- Comment #13

- In the last paragraph and in figure 4 authors discuss two patients who developed an NGAL above the reference level. Does this have any clinical meaning? If they think this may have been related to telaprevir I think they should perform statistical comparisons between NGAL levels on patients not on teleprevir, however it doesn't seem as if they have an appropriate comparison group. I think numbers are two small to make any real conslusions. I don't think they should include figure 4.
- We agree and we removed the last paragraph of results. Also we replace the old figure 4 with another one showing trends of both NGAL and eGFR in all 8 patients who started an antiviral therapy

- Comment #14
  - I think the first sentence is misleading as only eight patients were looked at after therapy and most therapies were first generation "DAAs" and not regimens that will be currently used.
  - We agree and we modified the first sentence of the discussion, as you can see at page #17, lines #304-305
    - ***This is the first study which investigated NGAL, both before and during the first twelve weeks of therapy with DAA including regimens***.

- Comment #15
  - The second paragraph is confusing. I think the idea of "complementary information" is unclear – the authors need to provide an argument for how NGAL could be useful, and I think they need more longitudinal data to do so. The last sentence does not make sense: do you mean "both cirrhotic and non-cirrhotic patients" (not "either cirrhotic or non-cirrhotic patients")?
  - We removed the idea of "complementary information" because we considered that our data were not enough strong to affirm that NGAL and eGFR are complementary, as you and Ref #1 kindly suggested. Paragraph was modified, as you can see at page #17, lines #307-319
    - ***We found that, in patients not exposed to DAA, eGFR was below normality in around 50%, while NGAL was in the range of normality in most individuals. Moreover, among the six patients with increased NGAL, in two cases eGFR was normal. These results suggest that a discordance between the two methods exist when interpretation is "categorical". Despite these findings, in the overall population plasmatic NGAL was statistically correlated with eGFR. Particularly, NGAL was significantly higher in patients with eGFR<60 ml/min than in patients with eGFR≥90 ml/min. It is difficult to explain apparent discrepancies with the current literature data because Alhaddad et al. had previously demonstrated that HCV positive cirrhotic patients with eGFR<60 ml/min had a significantly lower plasmatic NGAL than HCV positive cirrhotic patients with eGFR ≥60 ml/min [23]. The inclusion in our study of both cirrhotic and non-cirrhotic patients could be an explanation. In conclusion, we think that further studies should evaluate the rate of concordance between the two methods in diverse stages of liver disease. Also, it has to be seen whether these markers provide useful insights on the glomerular (especially detected by e-GFR) or tubular (especially detected by NGAL) damage in these conditions.***

- Comment #16
  - The third paragraph is also confusing. I do not fully understand the connection between all of the studies they are citing. Why cite the Azevedo and Sansonno studies, what information does that add to the current argument?

- You are right. We removed the reference to the studies by *Azevedo* and *Sansonno.*

- Comment #17
  - I think the focus of the results should be on paragraph 4 and 5
  - We agree but recognize that the number of patients is too small to provide strong conclusions. However, we tried to highlight the results of paragraph 4 and 5 in the discussion, being the main conclusion that NGAL could anticipate eGFR in revealing the development of a kidney damage (see page #19, lines #373-374). We have added the following sentence:
    - ***In particular, we hypothesize that NGAL reveals kidney damage earlier than eGFR during DAA containing regimens.***

- Comment #18
  - Again, I think they need to explain why NGAL is complementary information. What does it add? How might it help clinicians in the future?
  - Ideally, a scenario of how NGAL could be used should be provided but presently our data can only suggest a use of this biomarker as early predictor of kidney function deterioration, anticipating significant decreases of eGFR in analogy to what happens in other conditions. Please note how we have modified the discussion at pages #19 and 20, lines #375-381, in addition to what has been stated in response to the previous point (see answer to your comment #17):
    - ***However, many question points still remain to be answered. Indeed, it has to be established whether a clinical cut-off of NGAL may guide clinical decisions (e.g., dosage modification or stopping of the offending drug). Also, it has to be evaluated whether NGAL could predict AKI during HCV treatment, especially in most-at-risk patients such as those with advanced cirrhosis and a high risk of renal complications (e.g., hepato-renal syndrome). Lastly, cost-effectiveness studies need to be conducted to verify the hypothesis that NGAL should be routinary used to monitor kidney function during HCV treatment instead of (or in addition to) creatinine.***