# Answer to Reviewer's Comment

**Dear Editor:**

Thank you for reviewing the manuscript and suggesting your valuable comments.

We answer the comments below, and revisions or additions are indicated in blue letters in the manuscript.

**Reviewer #1:**

- Although it is obvious, please write the full names of LA, IP, AIDS and LL for the first time in the manuscript. Then the authors can freely use the abbreviations.

  **Answer:** According to the reviewer's comment, the full names of abbreviations were written throughout the manuscript.

- It is better not to use drug prescription orders like "QD" in scientific paper. Please change it to once daily or other similar terms.

  **Answer:** According to the reviewer's comment, we revised QD with 'once daily'.

- Please write the full name of abbreviations in figure 1.

  **Answer:** According to the reviewer's comment, the full names of abbreviations were written below Figure 1.

- In figure 2, please put the exclusion reasons from FAS in parenthesis. In this current format, it is misleading.

  **Answer:** According to the reviewer's comment, we revised the Figure 2.

- In table 1, please report P-values with similar decimals. The last row P-value has four digits, but the others have three digits. On the other hand, other P-values in other tables and texts are reported with four digits of decimals.

  **Answer:** According to the reviewer's comment, we revised the number of decimals in Table 1.

- The healing rate difference reported at 8 weeks in abstract and figure 3 are not the same as results

section. of 0.89% (95% CI, -0.86 to 2.64) vs. 0.9% (95% CI, -0.9−2.6).

- It is better to report erosive esophagitis healing rate difference at 4 weeks in abstract section.

  **Answer:** We would like to explain the answers to the above two comments together. According to reviewer's comment, we revised 'Abstract' by inserting the healing rate at week 4 and deleting the healing rate difference value (0.89%). Moreover, we revised the number of decimals as a whole.

- The authors found Fexuprazan effects rapidly observed after 4 weeks and the difference was similar compared to Esomeprazole. Thus, it might be possible to suggest a relative shorter treatment period for erosive esophagitis instead of 8 weeks of therapy. Although it needs further support, it could be promising. Please consider this point and expand it further in discussion.

  **Answer:** According to the reviewer's comment, we added the possibility of shorter treatment of fexuprazan than PPI for erosive esophagitis in the 'DISCUSSION' section.

- Please write full names of "fexu" and "esome" in results section.

  **Answer:** According to the reviewer's comment, we revised them in 'Healing rate of EE' of 'RESULTS' section.

- "Hp +" and "hp –" are quite misleading, especially when no abbreviations is previously determined. Please kindly edit this point.

  **Answer:** According to the reviewer's comment, we revised them in 'Healing rate of EE' of 'RESULTS' section.

- Total H.Pylori positive patients was 51 subjects (table 1), but in results it was reported as 43 (17 + 26). These controversial calculations have been also seen in H.Pylori negative as well as EM and PM subjects. Please clarify this issue and correct it accordingly.

  **Answer:** We answered this comment in the below [Answer to the discrepancy in the number of population]. Briefly, there were three analysis sets (FAS, PP and SS) in this study. Efficacy was evaluated by both the FAS and PPS, but PPS findings were interpreted as the main results. For the safety assessment, statistical analysis was performed on the SS. Therefore, there was the discrepancy of the number of participants in Table 1 (demographic data) and other Tables on efficacy and safety. Please see the answer below for more details.

- In table 1, EM and PM differed significantly according to intervention and control groups (P=

0.007), but there is not any explanation in results section. Please report it.

**Answer:** According to the reviewer's comment, we revised 'Baseline characteristics of the participants' of 'RESULTS' section by inserting "*There were no significant differences in the baseline characteristics between both groups, except CYP2C19 genotypes (EM or PM). A statistically significant difference was seen in the classification of CYP2C19 genotype (p=0.007), but the result was obtained from only some of the subjects who agreed to genotyping (n=51 and 56 in the fexuprazan and esomeprazole groups, respectively).*"

- Please briefly report the results of proportions of symptom-free days in the first 7 days and through the 8 weeks (supplementary tables 3 and 4) in results section.

**Answer:** According to the reviewer's comment, we added the results of proportions of symptom-free days in the first 7 days and through the 8 weeks as Supplementary table 3 and 4.

- It is not any comparison with reported P-values in table 2. In this current manner, a deduction cannot be made regarding the presence/absence of any difference between intervention and control groups or even within each group from the baseline. mod

**Answer:** According to the reviewer's comment, we revised the Table 2 by inserting P-values for the difference between treatment groups.

- It seems evaluation of GERD-HRQL questionnaire (table 3) was not done on all participants. It is better to report the probable reasons in results. Also if all participants were not assessed through RDQ, it is better to consider this issue and report the subjects being evaluated in table 2.

- In the last paragraph of "symptom response" section, the authors assessed symptoms relief in patients with moderate-to-severe symptoms and reported relevant percentages in the text. Also, they mentioned the supplementary tables. However, it should be noted that in supplementary table 5, the variables were assessed for total population (n= 218), NOT just for patients with moderate-to severe symptoms (n= 128 based on calculation in table 1). It is better to add relevant tables for these groups of patients and also report these variables based on total population.

**Answer:** We would like to explain the answers to the above two comments together. In this study, there were three analysis sets (full analysis set [FAS], per-protocol set [PP], and safety set [SS]): the FAS, based on the intention-to treat priniciple, included patients who received at least one dose of the study drug after randomization and had at least one primary efficacy assessment. The PPS included patients in the FAS who completed the study without any major protocol deviation. The SS group included all patients who received the study drug at least once after randomization. Efficacy was evaluated by both the FAS and PPS) and PPS findings were interpreted as the main results. For the safety assessment, statistical analysis was performed on the SS. Therefore, there was the discrepancy of the number of participants in Table 1 (demographic data) and other

Tables on efficacy and safety.

We described this in 'Statistical analysis of MATERIALS AND METHODS' and inserted the analysis set (FAS or PPS) in the Title of Tables.

- Which one is corrects? TEAE or TEAR? Both of these terms were used in the text and tables. Please select one of them. Otherwise, usage of these two indices is confusing. Also, please explain the definition of TEAE/TEAR.

  **Answer**: We corrected 'TEAR' to 'TEAE' in 'Safety' of 'RESULTS' section.

- In just supplementary tables 5 and 6, Fexuprazan is written as its generic name (DWP 14012). Please use one common name.

  **Answer:** According to the reviewer's comment, we wrote the 'Fexuprazan' instead of its generic name in the Supplementary tables 5.

- Although the manuscript has been edited by editorial agencies, it has still some minor issues needed to be corrected. For instance, "moderate events" has been written two times consecutively in results (safety section).

  **Answer**: We corrected it and revised 'Safety' of 'RESULTS' section.

- The authors used RDQ to assess severity of symptoms, but there is not any relevant references cited to this questionnaire. Also, there is not any explanation regarding the scoring system of this questionnaire. Please add them.

  **Answer:** According to the reviewer's comment, we inserted the explanation regarding RDQ in 'Protocol of MATERIALS AND METHODS' section instead of 'RESULTS' by inserting "*The RDQ is a self-administered questionnaires comprising of 12 items to assess the frequency and severity of heartburn, acid regurgitation, and dyspepsia. Each item for frequency and severity was scored from 0 to 5; the higher score, the more severe or frequent symptoms (Shaw et al., 2008). The RDQ demonstrated the validity and reliability for diagnosis of GERD in primary care and community settings (Shaw et al., 2001).*"

  [Added references]

Shaw M, Dent J, Beebe T, Junghard O, Wiklund I, Lind T, Johnsson F. The Reflux Disease Questionnaire: a measure for assessment of treatment response in clinical trials. Health Qual Life Outcomes 2008;6:31. DOI: 10.1186/1477-7525-6-31.

Shaw MJ, Talley NJ, Beebe TJ, Rockwood T, Carlsson R, Adlis S, Fendrick AM, Jones

R, Dent J, Bytzer P. Initial validation of a diagnostic questionnaire for gastroesophageal reflux disease. Am J Gastroenterol 2001;96(1):52-7. DOI: 10.1111/j.1572-0241.2001.03451.x.

- Moreover, the authors stated "Symptom severity in the daytime and at night were classified as none, mild, moderate, severe, or very severe". However, it seems none of the patients were in "none", "mild" and "very severe" groups (based on table 1). It is better to report this explanation in results section to avoid confusion.

  **Answer:** According to the reviewer's comment, we revised the scale for symptom severity in the daytime and at night in 'Protocol of MATERIALS AND METHODS' section instead of 'RESULTS' by inserting "*Symptom severity in the daytime and at night were measured according to the five-point scale (0: none, 1: mild, 2: moderate, 3: severe, 4: very severe).*".

- In table 1, the summation of H.Pylori, CYP2C19 (EM and PM) are not compatible with the total number of intervention and control groups. Please clarify this inconsistency.

  **Answer:** We answered the discrepancy in the participant number in Tables and Results above answer. Please see the above answer for more details.

- Despite the authors assessed ADR, no definition was provided in the main text. Also, there is not any information regarding this outcome in table 4.

  **Answer:** According to the reviewer's comment, we provided the definition of ADRs 'Protocol of MATERIALS AND METHODS' section by inserting "*Treatment-emergent adverse event (TEAE) was defined as an AEs newly occurred after the randomization and the first administration of study medication, and adverse drug reaction (ADR) was defined as any untoward and unintended response to the study medication of which causal relationship cannot be excluded.*".

- For appropriate comparison of TEAEs between groups, P-values are necessary. Please add them in table 4 or just mention those with significant difference between intervention and control groups, if applicable.

  **Answer**: According to reviewer's comment, we revised the Table 4 by inserting 'P-values' and 'Most frequently occurring TEAEs'.

- The authors stated "some statistically significant changes were observed in laboratory tests and vital signs". Please briefly mention them in the text.

  **Answer:** Compared to baseline, there were significant differences in the hematology

results (including RBC, hemoglobin and hematocrit), the results of serum chemistry (including uric acid and cholesterol) and the result of vital sign (including diastolic blood pressure and pulse rate) in the treatment groups at weeks 4 and 8. However, these changes were within the normal range with no clinical significance. Therefore, there were no clinically significant changes in the laboratory test, vital signs, physical examination and ECG findings. So, we revised 'Safety' of 'RESULTS' section by replacing with following sentence.

"*There were no clinically significant changes in the laboratory test, vital signs, physical examination and ECG findings.*"

- The item "n" is supplementary tables 3, 4 and 6 are redundant, because the main outcome in these tables is days, not participant numbers.

  **Answer:** According to reviewer's comment, we deleted the item 'n' in Supplementary tables 3 and 4.

- There is a major statistical issue in supplementary table 6 in "Proportion of Symptom-Free of Chronic Cough" variable. The median is equal to maximum which is statistically impossible. Re-analysis of the outcomes is highly recommended.

  **Answer:** We ask reviewers to review the results in the revised Supplementary Table 6 (hereinafter "S.Table 6") along with the Supplementary Table 5 (hereinafter "S.Table 5") of initial submission manuscript. Please review the reason why S.Table 6 results are not errors referring to the following.
  S.Table 5: 'Patient without Symptom of Chronic Cough in the first 3 days/7days/8 weeks' was defined as patient whose all symptoms (chronic cough) were assessed as 'None' during the first 3 days/7days/8 weeks. Each individual patient has a value of 'Yes' or 'No'.
  S.Table 6: 'Proportion of Symptom-Free Days of Chronic Cough in the first 3 days/7days/8 weeks' was derived as number of days which symptoms (chronic cough) were assessed 'None' during the first 3 days/7days/8 weeks divided by 3/7/study days. Each individual patient has a value from 0 to 100 (%). A patient whose is a value of '100(%)' in S.Table 6 has a value of 'Yes' in S.Table 5.
  For example, in the First 3 Days, the proportion of patients without symptom of chronic cough was 80.4% (86/107) in the fexuprazan 40 mg group in the S.Table 5. This means that 86 out of 107 patients had a value of 100(%) in the S.Table 6. If more than 50% of 107 patients have a value of 100(%), then both the median and maximum values can be presented as 100(%) in the S.Table 6.

**Reviewer #2:**

- The following points should be addressed First it is not correct approach to randomize and odd numbers of people in two groups it is highly recommended groups with equal number at randomization stage!

  **Answer:** We agree with your comment. Although odd numbers of participants were randomized in this study, the participants were randomized according to the appropriate Randomization method. We described the method of Randomization in 'MATERIALS AND METHODS'.

  In another previous study (Ashida K et al.,2016), odd numbers of participants were randomized. Please consider this.

  Ref: Ashida K, Sakurai Y, Hori T, Kudou K, Nishimura A, Hiramatsu N, Umegaki E, Iwakiri K. Randomised clinical trial: vonoprazan, a novel potassium-competitive acid blocker, vs. lansoprazole for the healing of erosive oesophagitis. Aliment Pharmacol Ther 2016; 43(2): 240-251 DOI: 10.1111/apt.13461

- The following results are not correct, when the efficacy rate is equal unto two groups the difference 0.89% is not correct. (99.1% (106/107) vs 99.1% (110/111)) with a difference of 0.89% (95% confidence interval, -0.86 to 2.64). How about the previous studies about the efficacy and safety of fexuprazan, please explain with more details.

  Answer: Please review the primary endpoint (healing rate of EE) statistical analysis method in [Statistical analysis] subsection in the revised manuscript. The difference of healing rates between treatment groups (common risk difference) and corresponding two-sided 95% confidence interval were calculated using the Cochran-Mantel-Haeszel method adjusted by a stratification factor (baseline LA grade). If difference of healing rates between treatment groups was calculated using simple proportion method without adjustment of stratification factor, we could derive the difference value of 0%. However, we considered the planned statistical analysis method to reflect the influence of baseline LA grade in this study. Therefore, it is not an error to calculate that common risk difference of healing rate was calculated 0.9%

- Which method of randomization (simple, block…) was used, please declare it. How about the random or regime assignment concealment?

  **Answer:** According to the reviewer's comment, we separately explained the randomization of this study in 'Randomization of MATERIALS AND METHODS'.

- Which rate of compliance was considered in current study?

  **Answer:** According to the reviewer's comment, we inserted following sentence in

'Baseline characteristics of the participants' of 'RESULTS' section.

**"***The mean compliance rates were 98.6±8.1% and 99.0±2.6 at weeks 4 and 8, and the overall compliance rate with study medication exceeded 95% in all treatment groups without between-group differences.*"

- Matters such as "EE (LA Classification Grades A to D) " and so on needs relevant references. More important: who about the inclusion and exclusion criteria? All variables you have reported them in results section should be introduced in methods section appropriately. My mean is those variable you reported them in table 1,….

  **Answer:** According to the reviewer's comment, we inserted footnotes explaining the LA Classification Grades and other variables below Table 1.

- The noninferiority margin 0.1 (your mean is 10% you wrote in it as 10% in results section) and type one error rate 2.5% and power 90% needs more sample size than 130 per group?!

  **Answer:** Our biostatistician estimated the sample size assuming that the complete healing rate of mucosal breaks was 94.8 at week 8 after treatment based on previous studies. And, the sample size was 104 patients per group using the conditions: non-inferiority margin of 10%, a one-sided significance level of 2.5%, 90% statistical power, and 1:1 randomization. We described this in the 'Sample size' of MATERIALS AND METHODS'.

- Please present the data about the validity and reliability RDQ and GERD-HRQL with relevant reference particularly in your country.

  **Answer:** According to the reviewer's comment, we inserted the explanation regarding the validity of RDQ and GERD-HRQL in 'Protocol of MATERIALS AND METHODS' section by inserting "*The RDQ is a self-administered questionnaires comprising of 12 items to assess the frequency and severity of heartburn, acid regurgitation, and dyspepsia. Each item for frequency and severity was scored from 0 to 5; the higher score, the more severe or frequent symptoms (Shaw et al., 2008). The RDQ demonstrated the validity and reliability for diagnosis of GERD in primary care and community settings (Shaw et al., 2001). … The GERD-HRQL was validated and considered as an appropriate instrument to evaluate typical GERD symptoms (Velanovich, 2007). Therefore, previous clinical studies performed in South Kore have used the RDQ and GERD-HRQL to evaluate the therapeutic effect in patients with GERD (Lee et al., 2019; Jeon et al., 2022)..*".

[Added references]

Shaw M, Dent J, Beebe T, Junghard O, Wiklund I, Lind T, Johnsson F. The Reflux Disease Questionnaire: a measure for assessment of treatment response in clinical

trials. Health Qual Life Outcomes 2008;6:31. DOI: 10.1186/1477-7525-6-31.

Shaw MJ, Talley NJ, Beebe TJ, Rockwood T, Carlsson R, Adlis S, Fendrick AM, Jones R, Dent J, Bytzer P. Initial validation of a diagnostic questionnaire for gastroesophageal reflux disease. Am J Gastroenterol 2001;96(1):52-7. DOI: 10.1111/j.1572-0241.2001.03451.x.

Velanovich V. The development of the GERD-HRQL symptom severity instrument. Dis Esophagus 2007;20(2):130-134. DOI: 10.1111/j.1442-2050.2007.00658.x.

Lee KJ, Son BK, Kim GH, Jung HK, Jung HY, Chung IK, Sung IK, Kim JI, Kim JH, Lee JS, Kwon JG, Park JH, Huh KC, Park KS, Park MI, Kim N, Lee OY, Jee SR, Lee SK, Youn SJ, Kim SK, Lee ST, Hong SJ, Choi SC, Kim TN, Youn YH, Park HJ, Kang MJ, Park CH, Kim BT, Youn S, Song GS, Rhee PL. Randomised phase 3 trial: tegoprazan, a novel potassium-competitive acid blocker, vs. esomeprazole in patients with erosive oesophagitis. Aliment Pharmacol Ther 2019;49(7):864-872. DOI: 10.1111/apt.15185.

Jeon HK, Kim GH, Lee MW, Joo DC, Lee BE. Randomized Controlled Trial Comparing the Efficacy of Sustained-Release Formula of Mosapride-Plus-Esomeprazole Combination Therapy to Esomeprazole Monotherapy in Patients with Gastroesophageal Reflux Disease. J Clin Med 2022;11(7):1965. DOI: 10.3390/jcm11071965.

- Results and Statistical analysis needs major revisions and your results should be based on new relevant statistical analyses; please get sophisticated consults from Biostatistician.

- How about the data presentation for continuous and categorical data?

- How about the normality evaluation data for continuous data?

**Answer:** We would like to explain the answers to the above three comments together. According to reviewer's comments, we revised the 'Results' and 'Statistical analysis' following the consults from our Biostatisticians.

- Which statistical tests you used for comparing variables in table 1? Please declare them in this section and footnote in below table.

**Answer:** According to reviewer's comment, we revised the footnotes of Table 1.

- Which statistical test you used for evaluating changes in serum gastrin levels in each group and between groups (Repeated measures ANOVA is needed with relevant and sound presentation of results). please refer the matters under heading Healing rate of EE (first paragraph) to figure 3.

**Answer:** In this study, we measured serum gastrin levels at weeks 4 and 8 to measure the gastrin level at each point. The serum gastrin levels were analyzed using an analysis

of covariance (ANCOVA) model, including treatment group as treatment effect and baseline gastrin levels and baseline LA grades as covariates. So, we inserted the statistical method of serum gastrin. We inserted the statistical test below Figure 4.

- What is LL? (Lower limit), also present relevant p-value for the both and second paragraph.

  **Answer:** According to reviewer's comment, we revised the abbreviations as a whole, and presented the p-values in 'RESULTS' section.

- Data presented in Supplementary tables 1-4 need p-value and you should clarify which tests are they based. Table 2 should be based on repeated measures ANOVA or GEE with relevant presentation (p for time, intervention and interaction of time and intervention. Table 4 needs p-value for compared data.

  **Answer:** According to reviewer's comment, we revised the whole Tables in the manuscript.