Title of Manuscript: Comparison of laparoscopic Nissen fundoplication with PPIs for laryngopharyngeal reflux with hiatal hernia based on oropharyngeal pH-monitoring and symptomatic-scale diagnosis.

ESPS Manuscript NO: 31201

Comments to the authors:

This is an interesting single-centre, comparative study comparing outcomes (short, mid-term) in patients who underwent Nissen fundoplication or were managed conservatively for symptoms of laryngopharyngeal reflux in the presence of type 1 hiatal hernia. The primary outcomes of interest are well described and defined in the manuscript. The authors acknowledge the limitations of the findings they report, arising from the design of the study.

Thank you for reviewing our manuscript and the support. The following commets and constructive suggestions will help us improve it to a better scientific level, we revised it and quite a lot of changes have taken place. At the following, the points mentioned by the reviewer will be discussed.

1) Why was the study not conducted in randomised manner or even a case-controlled manner so that is more reflective of true clinical practice?

Thank you for the comments. Patients were allocated to of proton pump inhibitors (PPIs) therapy or laparoscopic Nissen fundoplication (LNF) according to their own preference and physical condition after instructing them as following: PPI medication focused on the anti-acid, which need life-long medication but could not cause other damage or complication on upper gastrointestinal, whereas, LNF was an invasive operation, aiming to make a one-way flap by fundus for anti-reflux, with more possibility of injury and complications but lower recurrence rate. As suggested by the reviewer, we added it in the methods (Page 9 line 17-22).

2) Did the authors perform a power calculation test to confirm that the study's cohort of 70 patients would be sufficient to produce a robust result?

Thank you for the targeted comments. The nonparametric test was chosen when the shape of the underlying distribution is not known. However without making an explicit assumption about the distribution, detailed sample size calculations are impossible. Depending on the nature of the distribution, the nonparametric tests might require either more or fewer subjects. But they never require more than 15% additional subjects, if the following two assumptions are true (Type I error: α, Type II error: β). Indeed, The leaset sample size of a nonparametric test could compute for a parametric test and add 15%[1].

Comperhensive symptom score (RSI) and single symptom (typical symptom: globus) after treatment were used to calculate the sample size will be statistically analyzed using a parametric test (t test, two tailed). So we use the following formula to estimate the minimal mouse number for each group with α of 0.05, and b of 0.20.

$$n_{t-test} = 2(\mu_\alpha + \mu_\beta)^2(1 + 1/k)\sigma^2/\delta^2$$

$n_{\text{nonpar.}} = (1 + 15\%) \times n_{t-test}$

μ_α: means u value when α=0.05

μ_β: means u value when β=0.20

σ: means deviation between two groups

δ: means average standard deviation of two groups

k: means the ratio of sample size of control group in experiment group(k=PPI group/LNF group)

Indeed, after calculation the minor sample size of RSI score is 71, and the minor sample size of single symptom is 50, that size of 70 could be sufficient to draw the conclusion as the study was designed. In addition, we also mention the limitation of small sample size, enlarge the sample size could draw a more robust result as reviwer suggestion.

3) Why did the authors use 2 different non-parametric tests (Mann Whitney/Wilcoxon rank test) to analyse their results?

Thank you for the comments. The data of measurements do not obey a normal distribution. Thus we performed non-parametric test, which includes several statistical methods such as Mann-Whitney rank and Wilcoxon rank test. According to the suggestion from the experts, the Mann-Whitney rank test was often used in the different or independent sample; and Wilcoxon test was the most commonly used same (paired) or dependent sample. Therefore, we used Mann-Whitney rank test to compare the data between PPI and LNF groups, and Wilcoxon are more appropriate in the within group, like pre- and post-treatment in Table 4, we revised the statistical analysis in "MATERIALS AND METHODS" part (Page 11 line 13-18), and specified to explain difference in the implement of both non-parametric methods.

4) *LPR was defined as a major laryngopharyngeal symptom (hoarseness, globus, throat clearing/pain, mucus, and chronic cough) that occurred at least once a week, and was suspected by an ear, nose, and throat examination*- should be moved to methods section definitely.

Thank you for the constructive suggestion. I have moved it to the "MATERIALS AND METHODS" part (Patients enrolment) (Page 7 line 6-9).

5) *A total of 66.7-74.2% patients also suffered from typical GERD symptoms (regurgitation and/or heartburn) in the LNF and PPI groups.* Can the authors clarify what do they mean by the figure 66.7-74.2%?

Thank you for the comment. The pathology of GERD is mainly located in esophagus, and the esophageal symptoms include heartburn and regurgitation with high occurance rates. Thus, esophageal symptoms were defined as the typical symptoms of GERD, whereas, LPR symptoms were defined as atypical symptom. The mechanism of esophageal and LPR symptom generation is

complicated, and it does not follow "all-or-none" law. And this cohort was consisted of LPR patients no matter with or without typical esophageal symptoms. In this study, the data showed that 66.7-74.2% patients also suffered from typical heartburn and regurgitation (Table 1), and each symptom occurrence between LNF and PPI group has no significant difference. Figure 2 only showed the frequency and severity of LPR patients with heartburn and regurgitation and the relief following PPI and LNF treatment, which excluded the LPR patients without these typical esophageal symptoms.

6) The authors aimed to report the 2 year outcome, however only 53/70 patients completed the 2 year follow-up.

Thank you for the comment. This study was designed for two time-point follow-up (6 month and 2 year). Although only 53/70 patients completed the 2 year follow-up, and the whole loss of follow-up was over 20%, the most loss of follow-up came from the PPI group (PPI group: 11/39 vesus LNF group: 6/31). PPI administration was considered as classical treatment for LPR, and prescribed as the empirical medication when patients were suspected as LPR[2,3]. Most PPI patients are short-term in hospital or outpatients, which was more likely to be loss of follow-up. However, the less loss rate of follow-up in LNF group guaranteed the confidence of this study.

In addition, because this study was performed in the Xuanwu hospital and Second artillery general hospital of Chinese People's Liberation Army in Beijing, both of which are Grade-three Class-A hospitals, the top level in China. So, most of patients were admitted from other province or city, and it is hard to follow up all the patients by phone call, mail or E-mail, especially the outpatients with PPI medication within two years, although we got the informed consent of study from all patients.

7) It is unclear in the manuscript whether any of the patients undergoing Nissen's fundoplication were treated with a PPI at any point during the study.

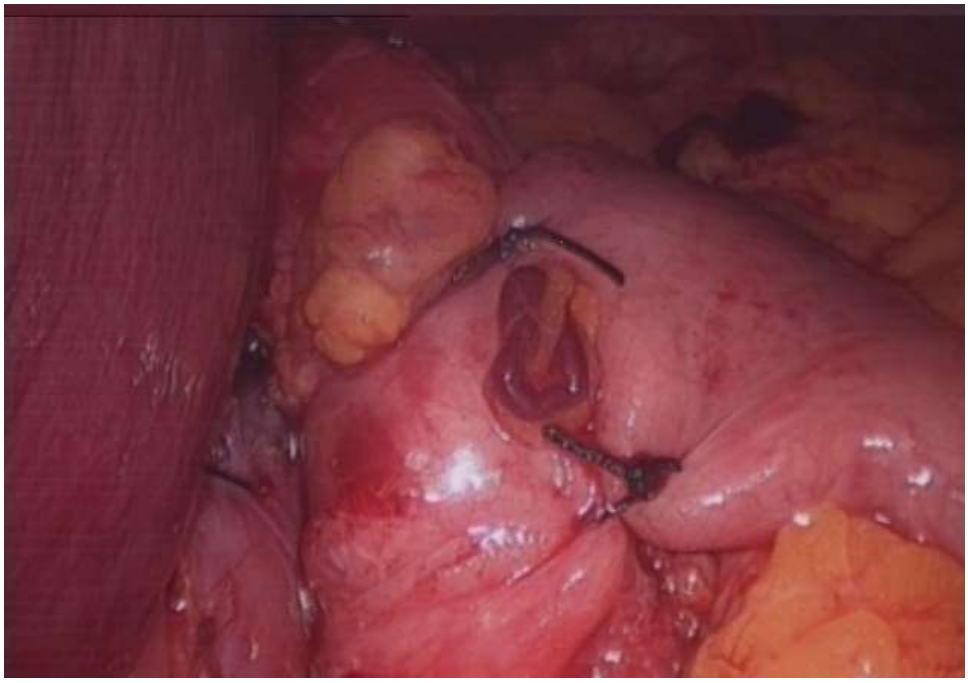
Thank you for the constructive suggestion. Only once omeprazole 40mg i.v. was used for gastric mucosal protection after operation, without any PPI medication during the in-hospital time; and PPI independence was defined as no continual 3 days PPI usage for GERD and LPR symptoms. As suggested by the reviewer, the PPI usage in LNF group and PPI independence was defined in "MATERIALS AND METHODS" part. We add "Only once omeprazole 40mg i.v. was used for gastric mucosal protection after operation." in "Treatment" part (Page 10 line 9-11). and "any patient who treated with any type of PPI medication continually over 3 days to relieve recurrent GERD and LPR symptoms was excluded from the medication independence from LNF or PPI group" in "Assessment" part (Page 11 line 7-9).

8) Grammar and spelling mistakes are encountered throughout the manuscript.

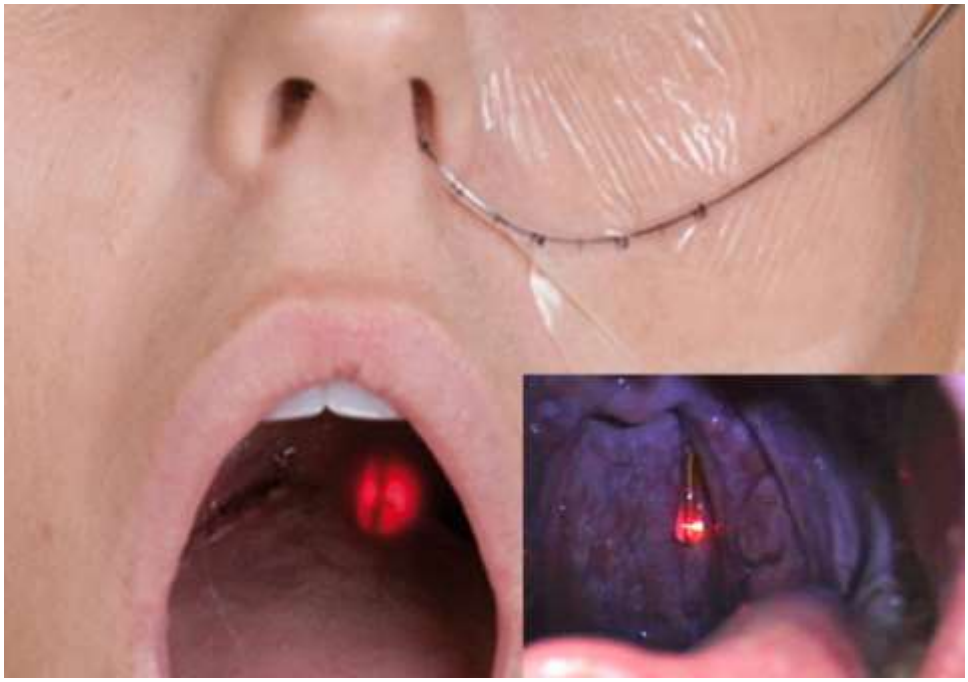
Thank you for the comments in detail, this manuscript has been proofread by language certificate by professional English language editor of Edanz Editing Company, again.

Attachments

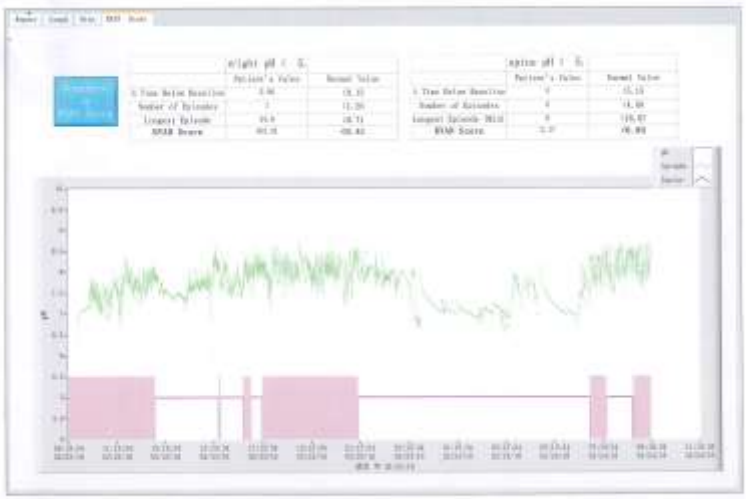
Laparoscopic Nissen fundoplication



Oropharyngeal pH-monitoring



Ryan report



Reference

[1] Lehmann EL, ed. Nonparametrics : Statistical Methods Based on Ranks, 1998.

[2] Ford CN. Evaluation and management of laryngopharyngeal reflux. JAMA, 2005, 294: 1534-1540.

[3] Aviv JE, Liu H, Parides M, et al. Laryngopharyngeal sensory deficits in patients with laryngopharyngeal reflux and dysphagia. Ann Otol Rhinol Laryngol, 2000, 109: 1000-1006.