REVIEW ARTICLE

PHILIPPINE JOURNAL OF OTOLARYNGOLOGY-HEAD AND NECK SURGERY VOL. 37 NO. 1 JANUARY – JUNE 2022

Effect of Proton Pump Inhibitors on Reflux Symptom Index (RSI) and Reflux Finding Score (RFS) in Patients with Laryngopharyngeal Reflux: A Systematic Review and Meta-Analysis

Patricia Ann U. Soriano, MD[1]
Erasmo Gonzalo D.V. Llanes, MD[1,2,3]
Anna Pamela C. Dela Cruz, MD[1]
Kevin Michael L. Mendoza, MD[1]

[1]*Department of Otolaryngology – Head and Neck Surgery*
Philippine General Hospital
University of the Philippines Manila

[2]*Department of Otolaryngology - Head and Neck Surgery*
College of Medicine
University of the Philippines Manila

[3]*Philippine National Ear Institute*
National Institutes of Health
University of the Philippines Manila

ABSTRACT

Objective: The purpose of this study was to determine the efficacy of proton pump inhibitor (PPI) therapy over placebo in treating the symptoms and laryngeal findings among adult patients with laryngopharyngeal reflux (LPR) using the Reflux Symptom Index (RSI) or Reflux Finding Score (RFS).

Methods: Placebo-controlled, randomized clinical trials published after June 2001 to January 2021 which used PPI as the sole intervention and the RSI or RFS as outcome measures were eligible for inclusion. Studies that were published prior to June 2001, those which only made use of questionnaires other than the RSI or RFS, those which used PPI in combination with other treatments, or those with unavailable full-text manuscripts were excluded. These studies were identified from MEDLINE, Scopus, Cochrane Library, Embase, and HERDIN Plus databases which were searched from May 21 to 26, 2020. The primary outcome was the mean difference between baseline/pre-treatment and post-treatment RSI scores for both PPI and placebo groups. The secondary outcome was the mean difference between pre-treatment and post-treatment RFS scores for PPI and placebo groups. Aggregate results of these outcomes were analyzed using forest plots. Heterogeneity was determined through prediction intervals. Risk of bias of individual studies was assessed using the Cochrane Collaboration's Tool in Assessing Risk of Bias.

Results: Nine randomized control trials were included with a total of 737 patients randomized and 595 patients analyzed – 294 from the PPI group and 301 from the placebo group. There were notable variations among the studies in terms of choice of PPI, dosage and frequency. Out of nine studies, four used both RSI and RFS in their analysis. Two studies used RSI alone and three used the RFS in combination with symptom questionnaires other than the RSI. There was a significant decrease in the RSI of the PPI group versus the placebo group with a mean difference of -2.83 (95% CI, -5.13 to -0.53, p = .02). However, there was no significant decrease in the RFS between PPI and placebo groups with a mean difference of -0.84 (95% CI, -2.66 to 0.98, p = .37). For two clinical trials which only reported post-treatment RFS, there was also no significant difference between the two treatment groups with a mean difference of 1.27 (95% CI, -0.22 to 2.76, p = .10).

Conclusion: This meta-analysis found that, although a statistically significant benefit in RSI was noted with PPI therapy, this difference may not translate to a clinically significant change in symptoms; therefore, there is insufficient evidence to recommend for or against the treatment of LPR with PPIs.

Keywords: *laryngopharyngeal reflux; proton pump inhibitors; laryngitis; hoarseness*

Correspondence: Dr. Anna Pamela C. Dela Cruz
Department of Otolaryngology - Head and Neck Surgery
Ward 10, Philippine General Hospital
Taft Avenue, Ermita Manila, 1000
Philippines
Phone: (632) 8554 8467
Email: acdelacruz14@up.edu.ph

The authors declared that this represents original material that is not being considered for publication or has not been published or accepted for publication elsewhere in full or in part, in print or electronic media; that the requirements for authorship have been met by all the authors, and that each author believes that the manuscript represents honest work.

Disclosures: The authors signed a disclosure that there are no financial or other (including personal) relationships, intellectual passion, political or religious beliefs, and institutional affiliations that might lead to a conflict of interest. This study was not funded or given financial support by any grant or agency.

Creative Commons (CC BY-NC-ND 4.0)
Attribution - NonCommercial - NoDerivatives 4.0 International

Philipp J Otolaryngol Head Neck Surg 2022; 37 (2): 6-14

© Philippine Society of Otolaryngology – Head and Neck Surgery, Inc.

Laryngopharyngeal reflux is a commonly diagnosed condition in the out-patient department.[1] Several problems, however, are encountered with respect to diagnosing and treating LPR. The true prevalence of gastroesophageal reflux among those with laryngeal symptoms varies widely which leads to a question of whether LPR has indeed been accurately diagnosed.[2] To date, there is no ideal test to clinch the diagnosis of LPR.[3] Twenty-four hour pH monitoring, although considered the gold standard for LPR, has not been shown to be a reliable test for diagnosing LPR given its low sensitivity.[1-3] The two validated instruments – the Reflux Symptom Index (RSI) and Reflux Finding Score (RFS) – were developed from patients diagnosed with LPR confirmed by 24-hour dual probe pH monitoring.[4-5] The RSI and RFS have been used to diagnose LPR among those who obtain scores greater than 13 and 7, respectively although Belafsky noted that RSI scores greater than or equal to 10 and RFS greater than or equal to 5 are clinically significant.[6] The empiric use of PPIs has been suggested as initial management for patients presenting with laryngeal symptoms in the background of gastroesopheageal reflux (GER).[1] Although there has been no consensus regarding the optimum diagnosis and a precise treatment protocol, the current recommendation states that twice-daily dosing of PPIs may be given for three to six months but with no consensus on the specific PPI or dosage.[1,3,7]

There have been nine previous systematic reviews and meta-analyses to date studying the effect of PPIs on LPR management.[2,8-15] All of them measured successful LPR treatment as a 50% reduction in symptoms. These studies showed conflicting results since six of them concluded that the use of proton pump inhibitors did not yield differences that were statistically significant when compared with placebo with an overall risk ratio ranging from 1.18-1.21 but with lower and upper values of the confidence interval less than and greater than 1, respectively.[2,8-12] The three remaining systematic reviews and meta-analyses concluded that the use of PPIs could improve reflux symptoms although with varying clinical significance.[13-15] The limitations of all these previous studies were in the various sources of heterogeneity – the inclusion criteria, the specific PPI chosen, the dosing, the duration, and the outcomes measured. The individual trials included in the systematic reviews and meta-analyses chose different laryngeal symptoms to assess with some studies creating their own questionnaires which may have contributed to differences in results among studies. Among the meta-analyses listed, 2 of them conducted subgroup analyses using changes in RSI and RFS as the measured outcome; however, on close inspection, both included studies which did not actually make use of the RSI or RFS, thus making their conclusions questionable.[13,14] With the introduction of the validated RSI and the endoscopy-based RFS, a more standardized manner of assessing relief may be used which may also lead to more comparable studies and definitive conclusions.

The main objective of this study was to determine the effectiveness of PPI over placebo in the treatment of LPR among adult patients, defined as a statistically significant change in the RSI and RFS from pre-treatment to post-treatment after a treatment period of at least one month. A minor objective was to validate conclusions from previous systematic reviews or meta-analyses by specifically selecting RSI and RFS as the outcome measures. Because these two instruments were used to assess efficacy of treatment outcome, one source of variation was eliminated. Findings of this meta-analysis aimed to assess if the conclusions drawn would support majority of the previous reviews which did not find that PPIs were effective in treating LPR versus placebo.

METHODS

This systematic review and meta-analysis was conducted from May 21, 2020, to April 22, 2021 with University of the Philippines Manila Research Ethics Board exemption number 2020-714-EX. This review was registered in the Research Grants Administration Office, University of the Philippines, Manila (RGAO-2020-0048).

Eligibility criteria

Studies eligible for inclusion were randomized, placebo-controlled trials involving adult patients with LPR which used the RSI and/or the RFS as outcome measurements. Studies which used proton pump inhibitors as the sole intervention for at least one month were included. Randomized controlled trials conducted after June 2001 to January 2021 were eligible as the RFS was published first in June 2001. Studies that compared PPIs with other drugs or in combination with other non-pharmacologic interventions were excluded. Studies which used acid reflux measurement or gastroesophageal symptom relief as the sole outcomes were excluded. Studies which only made use of questionnaires other than the standardized RSI or RFS were excluded. Studies that used the RFS or RSI but did not report mean scores, standard deviations, or standard error were excluded. Studies which were published prior to 2001 were excluded. Abstracts, reports, and unpublished manuscripts were eligible for inclusion if personal correspondence with the studies' primary authors yielded full-text manuscripts for analysis; otherwise, these were excluded. Included studies were later divided into PPI and placebo groups for synthesis based on either RSI or RFS measurements.

Information Sources

A search of randomized control studies using PPIs and placebo was done independently by the principal investigator (PAUS) and a

co-investigator (KMLM) using MEDLINE (through PubMed), Scopus, Cochrane Library, Embase (through Ovid@journal), HERDIN Plus, and reference lists of existing systematic reviews and meta-analyses from May 21-26, 2020. A final search of each database was done from December 18-20, 2020 to assess if new articles had been added that might be eligible for the meta-analysis.

Search Strategy

To create a thorough search, different keywords using laryngopharyngeal reflux and proton pump inhibitors such as "laryngopharyngeal reflux," "reflux laryngitis," "laryngitis," "chronic cough," "hoarseness," "proton pump inhibitors," "omeprazole," "rabeprazole," "esomeprazole," "lansoprazole," "pantoprazole," and "dexlansoprazole" were used. A sample search included a line search of the MeSH Term "proton pump inhibitors" and each specific type of PPI under the filter of Title/Abstract using the conjunction "OR." The next line search included the MeSH Term "laryngopharyngeal reflux" and the chosen specific symptom phrases aforementioned under the filter of Title/Abstract using the conjunction "OR." These two lines were joined together in one search combined by the conjunction "AND" to obtain all possible articles with this intersection of search terms.

Other studies were sought by searching for previous or ongoing trials registered in ClinicalTrials.gov to determine if there were any unpublished but relevant studies. The final database search for each was conducted in December 18 to 21, 2020. Only studies after June 2001 until January 2021 were included in the search. Among all the studies included in the final analysis, the two most recently published RCTs were obtained by contacting the author of the published protocol through e-mail. Studies were not strictly limited to those written in English; however, no potential study published in another language fit the eligibility criteria.

Data Extraction

All studies regarding LPR and PPI treatment independently found by two researchers (PAUS and KMLM) were listed with duplicate studies combined using the Mendeley Desktop program version 1.19.4 for Windows (Mendeley Ltd., Elsevier, Amsterdam, Netherlands). Each study's abstract was then analyzed for eligibility. The full text of each seemingly relevant study was read thereafter to assess if it could be included in the meta-analysis. Disagreements between the two researchers regarding the inclusion of a study were settled by consensus. Data from each study such as number of participants, specific drug choice and dosing, treatment duration, outcomes measured, and methodology were extracted independently. Risk of bias was carried out by the primary author using the Cochrane Collaboration's Tool in Assessing Risk of Bias and the Review Manager (RevMan) Version 5.3, 2014 (The Nordic Cochrane Centre, The Cochrane Collaboration, Copenhagen, Denmark).[16]

Statistical Analysis

All articles gathered from databases and other sources were screened through reading their respective abstracts. Those with information that appeared to fit the eligibility criteria were retained, and available full-text articles of each were obtained. Study characteristics such as participant characteristics, specific treatment regimen, treatment duration, control variable, mean differences and SD of RSI, mean differences and SD of RFS were then summarized and tabulated to ensure that eligibility criteria were met prior to inclusion in further analysis.

The primary outcome identified in each study was the mean difference between baseline/pre-treatment and post-treatment RSI scores for both treatment (PPI) and placebo groups. The secondary outcome was the mean difference in RFS from pre- to post-treatment between the 2 treatment groups. Duration of treatment was dependent on what was stated in each study as the timeframe prior to primary outcome measurement. Standard deviations (SD) of the mean differences were also reported. In the event that mean differences and standard deviations were not reported, the authors were contacted via e-mail to obtain the missing data.

Two studies reported standard error of the mean instead of standard deviation. [3,7] Standard deviation was then calculated from the data provided using the following formula:

$SD = SE \sqrt{n}$

SE = standard error of the mean

n = sample size

The authors of two studies were unable to supply the mean differences, so a separate analysis was conducted using the post-treatment mean scores and SD as advised by the consultant statistician.[17,18] One study reported a median score and range instead of mean scores and SD.[18] The median score was reported as the mean score. The SD was derived from the range and was computed as:

$SD = ¼ (max – min)$

We used forest plots to depict the summary of findings of the studies using a 95% confidence interval (CI) using Review Manager (RevMan) Version 5.3, 2014 (The Nordic Cochrane Centre, The Cochrane Collaboration, Copenhagen, Denmark). To assess for heterogeneity, Higgins I^2 coefficient and prediction intervals were identified. Random effects modelling was applied for the analysis based on the assumption that the mean and effect size likely varied across studies. Publication

bias was also assessed through the generation of funnel plots using the Comprehensive Meta-Analysis Program version 3.3.070, 2014 (Biostat, Englewood, NJ, USA).

RESULTS

Literature Search

The literature search process yielded nine articles that were included in the analysis. *(Figure 1)* The initial search among databases and other sources resulted in a total of 525 articles. After adjusting for duplicates, 441 articles remained. A further 421 articles were excluded since these were mostly reviews, observational studies, uncontrolled studies, or studies which did not make use of the RSI or RFS. The full-text of the remaining 20 articles were read for eligibility resulting in nine articles fitting all the criteria for inclusion and exclusion. *Figure 1* also shows more detailed descriptions regarding reasons for excluding the other 11 articles. Although nine articles were included in this meta-analysis, two of the articles were published using the exact same protocol and participant data in two different journals but under different primary authors.[19,20] This duplication of published data ultimately resulted in only eight independent data sets being included in the meta-analysis.

Study Characteristics

Among the nine RCTs included in the analysis, there was a total of 737 patients randomized but only 595 analyzed – 294 of whom were part of the treatment (PPI) group and 301 of whom belonged to the control (placebo) group. Characteristics of each individual study were summarized and tabulated. *(Table 1)* Out of the nine included studies, four studies made use of both the RSI and RFS and were included in both analyses – primary and secondary outcome. Two studies used only the RSI; however, as previously stated, these two studies used a single data set. The remaining three studies analyzed the RFS but used symptom questionnaires other than the RSI in their methodology and analysis. It is notable that there were variations among studies in terms of choice of PPI, dosage, and frequency although majority of the studies (6 out of 8 data sets) used a twice daily regimen. Treatment duration for primary outcome measurement also varied among studies ranging from 6 weeks to 16 weeks. Most of the studies reported low drop-out rates except for the study of Wilson[21]/O'Hara[22] which reported 22.8% (79 out of 346) lost to follow up at the time of primary outcome measurement.

Primary Outcome

The primary outcome measured was the mean difference of the RSI per treatment group at baseline and post-treatment which included six studies (five data sets). There was a significant decrease in RSI score in the PPI group versus the placebo group with a mean difference of -2.83 (95% CI, -5.13 to -0.53, p = .02) as noted in the forest plot. *(Figure 2)* Most studies' mean differences lay on the side favoring PPIs except for the data set with the largest sample size which showed contrary findings regarding the efficacy of PPIs in reducing RSI scores. Random effects modeling was employed; however, it can be noted that fixed effects modeling resulted in a similar overall effect with a mean difference of 2.94 (95% CI, 2.02 to 3.8, p<.00001). The resulting I^2 coefficient showed that 81% of the variance would have remained if sampling error could be removed.[24] This high I^2 value was contributed solely by the outlier study of Wilson[21] and O'Hara[22] since removal of these articles brought the I^2 coefficient down to zero. To show the heterogeneity more accurately between studies, a prediction interval was computed showing that the true effect size for 95% of all populations would lie somewhere from -10.973 to 5.313.[24] This wide range shows that there is significant heterogeneity among studies.

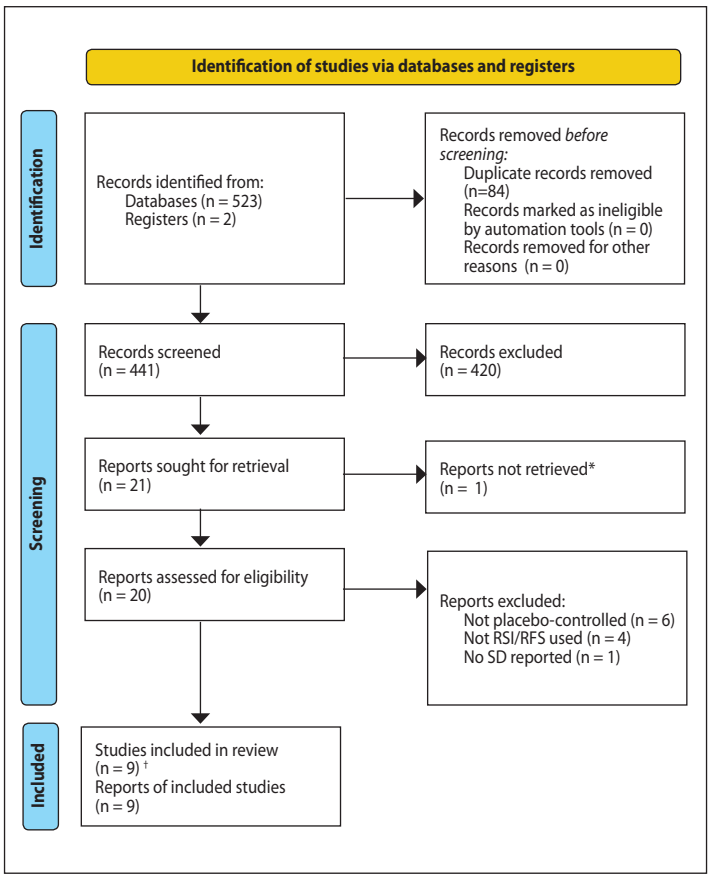

Figure 1. PRISMA Preferred Reporting Items for Systematic Reviews and Meta-Analyses for literature search process
*Full-text article could not be retrieved online or by contacting the primary author through e-mail
† 9 full-text articles included in meta-analysis with 2 studies sharing the same data
RSI Reflux Symptom Index; RFS Reflux Finding Score; SD standard deviation

Secondary Outcome

The secondary outcome measured was the mean difference of the RFS pre- and post-treatment between the PPI and placebo groups. The objective changes in laryngopharyngeal reflux showed the same tendency for benefit of PPI treatment as in symptomatic relief; however, this was not significant with a mean difference of -0.84 (95% CI, -2.66 to 0.98, p = .37) which can be seen in the forest plot. *(Figure 3)* There was once again significant heterogeneity seen. It was estimated that 85% of the observed variance would remain if variance due to sampling error could be removed (I^2 = 85%). On analysis of the five studies included, the two studies reporting a significant mean difference in RFS favoring PPI contributed greatly to this as removal of these studies decreased the I^2 coefficient to 24%. The prediction interval was computed which showed that the true effect size for 95% of all populations would be found within the range of means from -7.5927 to 5.9127, once again showing significant heterogeneity evidenced by the wide range of dispersion.

As stated previously, two studies included in the meta-analysis did not provide mean differences between pre-treatment and post-treatment RFS. A separate analysis was conducted using the post-treatment RFS mean of the PPI and placebo groups. *(Figure 4)* In contrast, there was a tendency for the placebo to cause a reduction in RFS; however, this was again not significant with a mean difference of 1.27 (95% CI, -0.22 to 2.76, p = .10).

Risk of Bias

The risk of bias of individual studies was assessed by the primary author. Anzić's study was identified as the study with the highest risk of bias with issues in selection bias, incomplete outcome data, and selective reporting.[23] The authors of the study stated a limitation of their study being the lack of stratification of baseline characteristics of study participants. It was revealed that, although random allocation was done, more patients with higher baseline RSI scores were assigned to the PPI group. Another bias of this particular study was the apparent incomplete outcome reporting. The study stated in their methodology that other objective measurements were taken, such as pH probe measurements and microbiopsy of inferior turbinates, none of which were reported in the results. The study also chose to only report the mean scores of each group at baseline and at the end of treatment, as well as a corresponding p-value with no mention of mean differences pre- and post-treatment. However, this was remedied by e-mailing the author who provided the missing data. Still, analysis was repeated after removing this study with the highest risk of bias which resulted in an overall RSI mean difference of -2.56 (-5.74, 0.62), resulting in a result that was no longer statistically significant. It was noted that significant

Table 1. Study Characteristics

Source	Location	Patients Randomized	Positive pH Probe Test Required*	Patients Analyzed (PPI/ placebo)	PPI dosage and frequency	RSI/ RFS	Treatment duration‡	Drop-outs/ Lost to follow up
Wo, 2006[20]	University of Louisville, USA	39	Yes	20/19	Pantoprazole 40mg, once in the morning	RFS†	12 weeks	-
Pawar, 2007[23]	Medical College of Wisconsin	53	No¶	21/26	Rabeprazole 20mg, twice daily	RSI, RFS	90 days	6
Reichel, 2008[7]	Ludwig Maximilians University, Munich	62	No	30/28	Esomeprazole 20 mg, twice daily	RSI, RFS	3 months	4
Lam, 2010[3]	University of Hong Kong, Queen Mary Hospital	86	No¶	42/40	Rabeprazole 20mg, twice daily	RSI, RFS	12 weeks	4
Faruqi, 2011[19]	Hull Cough Clinic, Cottingham, UK	51	No	24/25	Esomeprazole 20mg, twice daily	RFS†	8 weeks	2
Shaheen, 2011[24]	Chapel Hill, North Carolina	40	No¶	22/18	Esomeprazole 40mg, twice daily	RFS†	12 weeks	0
Anzić, 2018[25]	University of Zagreb, Croatia	60	No¶	33/27	Omeprazole 20mg, once daily	RSI, RFS	8 weeks	0
Wilson, 2021[21]/ O'Hara, 2021[22]	Eight UK NHS sites	346	No	102/118§	Lansoprazole 30mg, twice daily	RSI	16 weeks	79‖

* pH probe value <4 required for inclusion
† Symptom questionnaires other than RSI used
‡ Primary outcome measurement as stated in each study
¶ pH probe testing done and participants divided into positive or negative
§Compliant Intention to treat analysis (Pragmatic intention to treat sample size was 127/140)
‖Missing data at time of primary outcome measurement: 63 (37 for PPI and 26 for placebo) did not attend at 16 weeks; 16 (8 for each intervention group) attended but no RSI recorded

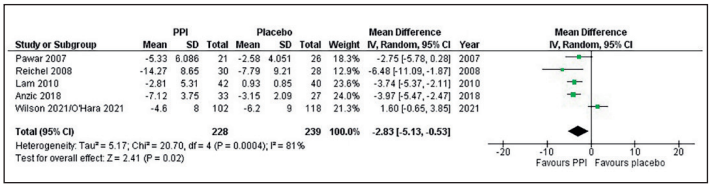

Figure 2. Forest plot showing the mean difference in Reflux Symptom Index between PPI and placebo

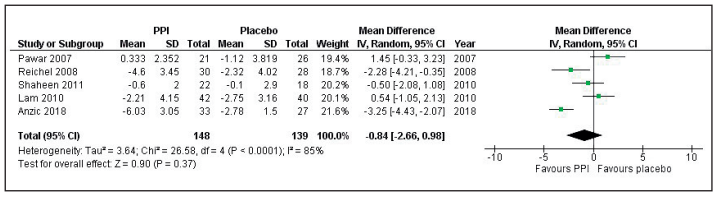

Figure 3. Forest plot showing the mean difference in Reflux Finding Score between PPI and placebo

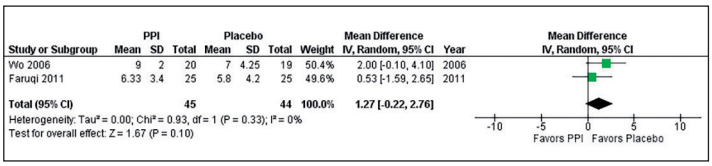

Figure 4. Forest plot showing the post-treatment Reflux Finding Score mean between PPI and placebo

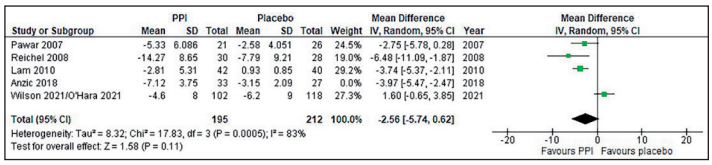

Figure 5. Forest plot showing the mean difference in RSI with Anzić's study removed

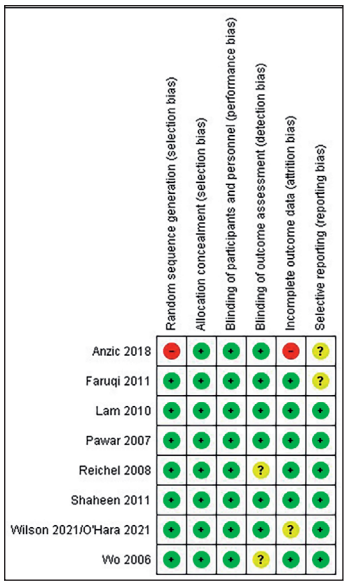

Figure 6. Risk of bias summary: judgements about each risk of bias item for each included study.

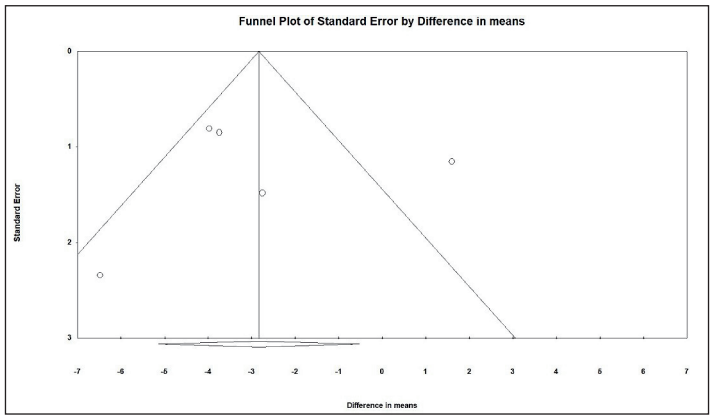

Figure 7. Funnel plot of studies assessing the effect on RSI

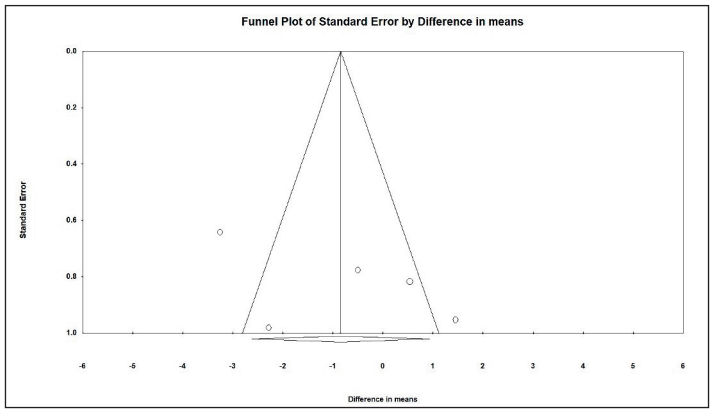

Figure 8. Funnel plot of studies assessing the effect on RFS

heterogeneity remained despite removal of this study. *(Figure 5)*

Two studies with unclear but potentially significant bias came from the most recent RCT due to the high dropout rate where only 267 out of 346 participants had data from the initial end-of-treatment assessment.[19,20] Additionally, only 220 participants were included in the main statistical tests of analysis. However, the same statistical tests were done for the pragmatic-intention-to-treat sample, and their data was reported in the appendices which showed similar findings. All other studies included in the meta-analysis were generally deemed to have low risk of bias. *(Figure 6)*

In order to assess publication bias, funnel plots were created and analyzed. Visual inspection showed asymmetry of findings for both sets of studies which assessed the RSI and the RFS. Both funnel plots showed outlier studies which did not lie within the funnel *(Figures 7 and 8)*. For studies which assessed the effect of PPIs on RSI, there is a gap at the bottom of the funnel which shows a paucity of data from smaller studies with non-significant effects. This points to the possible existence of publication bias.

On the other hand, the funnel plot for the studies assessing PPI effect on RFS seems to show some symmetry; however, it can be concluded that more precise studies are lacking as shown by the gap in the upper half of the triangle. Two studies are also shown to lie outside of the funnel plot due to the extreme value of the mean difference found in these studies.[21,23] A funnel plot for the second analysis involving the remaining studies was not generated due to there being only two studies in the analysis.

Ideally, more tests to quantify and assess funnel plot asymmetry would have been done. However, these tests are not used in the event of fewer than ten studies since conclusions from these results cannot be drawn or relied upon.[25]

DISCUSSION

The use of twice-daily dosing of PPI in the empiric treatment for LPR has been advocated for years; however, this recommendation has been drawn mostly from non-randomized and uncontrolled studies.[26] There are no current clinical practice guidelines (CPG) from societies of otolaryngology that specifically or extensively discuss the disease entity LPR. The only available CPG regarding dysphonia/hoarseness by the American Academy of Otolaryngology-Head and Neck Surgery (AAO-HNS) recommends against the use of PPIs for isolated dysphonia without evidence of laryngeal changes or symptoms of gastroesophageal reflux disease (GERD).[27] A separate position statement from the AAO-HNS simply states that empirical medical treatment may be diagnostic of LPR.[28] There was no mention regarding the application of Belafsky's validated instruments (RSI or RFS) in the CPG or position statement,

and the position statement did not specify medication as the explicit prescription of PPIs.

Several prospective studies that do report the effectiveness of PPIs on significantly reducing RSI and/or RFS are either not placebo-controlled studies or do not have control groups for comparison.[29-33] Extensive search for placebo-controlled RCTs only yielded eight independent data sets (nine articles) that met the eligibility criteria for our meta-analysis, most of which had relatively small sample sizes. This clearly shows the need for the conduction of RCTs with larger samples to create a more precise analysis of the true effect of PPIs in reducing symptoms and laryngeal findings of LPR. It is also notable that even with the publication of validated instruments for LPR (i.e. RSI and RFS), several placebo-controlled RCTs published after 2002 continued to make use of other symptom questionnaires (such as cough or GERD questionnaires) and did not apply the RFS.[34-37] Three of the RCTs included in this meta-analysis applied the RFS but not the RSI despite being published after the creation of the RSI.[17,18,22] This is contrary to the range of observational studies that did make use of both the RSI and RFS.[29-33] Instead of creating a larger pool of studies which could have resulted in a potentially more homogenous set of data and a more precise and powerful meta-analysis, those studies decreased the number of RCTs that could have been included.

The primary outcome showed that PPI therapy is effective in symptomatic relief of LPR as shown by the overall mean effect among studies. It must be stated that, although the overall mean difference showed a benefit for PPI therapy, the point estimate and corresponding confidence interval may not necessarily translate to a clinically significant change. Severe reflux disease was defined as an RSI > 13 and RFS > 7.[4,5] If the mean difference of -2.83 were applied, the change would result in an RSI > 10, which would still be deemed clinically significant.[6] This lack of efficacy was further highlighted when the study with the highest risk of bias was removed resulting in an overall mean difference in RSI that was not statistically significant. The same inference may be said for the RFS. If an overall mean difference of -0.84 were applied, this would not result in a change that would be clinically significant (RFS > 5).[6]

The greatest source of heterogeneity was contributed by the most recently conduct RCT which was also the largest trial to date and the only published study that did not find PPIs to be effective in the treatment of LPR.[19-20] This may be due to the smaller studies having poorer methodological quality resulting in larger or inflated effect sizes.[25] It is also possible that studies with similar sample sizes which did not yield significant effects were not published and, thus, did not contribute to the pool of known data regarding PPI efficacy. This is evidenced by the funnel plot generated. *(Figure 7)* This possible list of unpublished studies, however, was not found during the literature search of the reviewers.

Contrary to the popular usage of Higgins I^2 as the marker for between-study heterogeneity, Borenstein *et al.* clarify that it is a proportion between the variance of the true effect versus the variance in the observed effect, and that it is not automatically used as a surrogate for heterogeneity.[24] Hence, the use of prediction intervals has been suggested to illustrate the range of variance of true effects instead of the I^2 value as the former would lead to a more accurate depiction of dispersion of effects between studies.[24,38] Due to the presence of the outlier articles,[19-20] heterogeneity was found to be significant as evidenced by the wide range of the prediction interval which showed that the true effect size could actually lie on the side of no benefit (upper limit value of 5.313). The outlier articles prove the need for future RCTs with larger sample sizes and more rigorous methodology to determine whether the consistent positive findings of the smaller studies might, in fact, be due to inherent bias rather than the true effectiveness of PPIs.

The secondary outcome of the effectiveness of PPI therapy on laryngeal findings (RFS) showed the same tendency for benefit albeit very small. However, the confidence interval of the overall mean difference for the two analyses of RFS showed that the benefit was not significant. This translates to a difference which is also not clinically significant. Heterogeneity was also noted to be significant as shown by the wide range of the prediction interval. This heterogeneity may be due to the differences in the protocols used per study. Although four out five studies in the first RFS analysis made use of both RSI and RFS, only two studies[7,23] applied cut-off values for both the RSI and RFS (RSI >13 and RFS > 7) as part of their inclusion criteria while one study only applied the cutoff value of RFS > 7 but did not take into account the RSI score in the inclusion criteria.[3] Although Pawar et al. used both instruments in their analysis, the inclusion criteria was not stringent as no cut-off values were used in the recruitment process.[21] The study of Shaheen *et al.* made use of entirely different questionnaires and did not factor in the RFS in its inclusion criteria.[22] The exclusion criteria among studies also varied widely with one study[21] excluding all patients with signs of acute or chronic sinus disease while two studies specifically included patients with signs of postnasal drip or chronic sinus disease.[22,23] All of these variations in inclusion and exclusion criteria may have contributed to the heterogeneity.

It is interesting to note that this heterogeneity is absent in the analysis of RSI for the four studies which analyzed both RSI and RFS.[3,7,21,23] If prediction intervals were to be computed for these four common studies alone, the range for the RSI is narrow with a lower limit of -6.0872 and an upper limit of -1.6528 suggesting a narrow dispersion or variation between studies. In contrast, the prediction interval for the RFS of the four same studies is quite wide with a lower limit of -11.7547

and an upper limit of 9.9147 suggesting a very wide between-study variation. In this case, it is most probable that the low number of studies (n=4) actually causes the falsely narrow and falsely wide prediction interval values for the RSI and RFS, respectively.[38]

In assessing the possibility of publication bias for studies assessing RFS, the funnel plot generated suggested that there was little publication bias since most studies were found on the bottom of the funnel and on both sides of the mean difference. *(Figure 8)* This was in contrast to the funnel plot generated for the RSI which suggested possible publication bias. It is possible that the apparent absence of publication bias for RFS is because the RFS as an outcome measure is usually a secondary outcome while any symptom questionnaire was reported as the primary outcome. Given this, a small study which reports significant findings with respect to symptomatic relief would still have to report findings of the RFS even in the event of non-significant findings.[3] However, it was noted that among the seven studies that made use of the RFS, four studies actually did not find that PPI therapy resulted in a significant effect in their primary outcome measurement whether it was the RSI or another questionnaire.[17,18,21,22]

It was noted that none of the RCTs included applied the cut-off values for the RSI or RFS suggested by Belafsky *et al*.[4,5] in assessing response or resolution of LPR; rather, all studies computed for the overall mean score per treatment group with or without the mean difference from baseline to post-treatment. After contacting the authors of the studies, only data from Anzić's study were obtained regarding the proportion of patients with final RFS and RSI scores less than the cut-off. Surprisingly, only two participants (6.06%) in the PPI group had a post-treatment RSI less than 13, and eight (24.24%) had an RFS less than 7. It was noted that zero (0%) and four participants (14.81%) in the placebo group had an RSI less than 13 and RFS less than 7, respectively. Since only one out of nine studies has data on proportions of patients' scores, no further subgroup analysis can be carried out. It may be suggested then for future researchers to apply these cut-off values as published by Belafsky *et al*.[4,5] A further recommendation would be for future researchers to stratify patients into those with severe disease with RSI greater than 13 and RFS greater than 7 from those with clinically significant disease with RSI greater than 10 and RFS greater than 5.[4-6]

It is clear that larger, stringent, placebo-controlled, RCTs studying the effectiveness of PPI therapy in LPR are still lacking. More so, it is recommended that future researchers make use of the RSI and RFS which are validated instruments specifically for LPR instead of other questionnaires. Although this meta-analysis found that there was a statistically significant difference in post-treatment RSI scores favoring PPIs over placebo, this does not necessarily translate to a clinically significant effect among patients. Given the presence of great heterogeneity among studies, possible risk of bias in at least one RCT, and a lack of reporting using cut-off values suggested by Belafsky for the RSI and RFS, there is currently insufficient evidence to recommend for or against the use of PPIs in the treatment of LPR.[6]

The conduction of larger RCTs with rigorous methodology should address these current issues in order to create a more scientifically sound recommendation regarding the treatment of LPR. These future trials can possibly address the question of PPI efficacy by stratifying patients based on cut-off values for severe disease (RSI>13 and RFS>7) and for clinically significant disease (RSI>10 and RFS>5) in their inclusion criteria and outcome measurements instead of purely reporting mean scores and mean differences. Other areas of variability that should be addressed by future trials are the exact type of PPI, dosage, frequency, and minimum duration of therapy.

With regard to side effects, only one RCT reported a single serious adverse effect possibly related to the treatment.[19,20] All others reported no serious adverse events related to the intake of PPIs. Still, prescribing PPIs is not without risk, not to mention the potential financial strain of prolonged treatment. These factors must also be taken into consideration when weighing the benefit of empiric treatment given the lack of robust evidence to support it.

As stated, the AAO-HNS recommends against prescribing PPIs for isolated dysphonia without documenting laryngeal findings suggestive of reflux disease.[27] In this age of new out-patient consultation procedures with the advent of virtual consultations or telemedicine/telehealth, new layers to the decision-making process for practitioners are added. One must not only decide whether to prescribe PPIs at all if a diagnosis of LPR is being considered but, if so, when – during a virtual consult without the aid of a physical examination or after? We expect that recommendations that will guide these decisions will continue to evolve as more studies are developed and published.

Registration and Protocol

A protocol was created and exempted from the authors' institution ethical board of review (UP Manila Research Ethics Board (UPMREB 2020-714-EX). This review was registered in the Research Grants Administration Office, University of the Philippines, Manila on January 15, 2020 (RGAO-2020-0048).

ACKNOWLEDGEMENTS

We thank Dr. Louella Patricia Carpio for answering clarifications on statistical computations and approaches for this meta-analysis and Professor Davor Plavec and Mr. James O'Hara for responding to queries and providing any additional information requested regarding their respective studies.

REFERENCES

1. Kavitt RT, Vaezi MF. Diseases of the esophagus. In: Flint PW, Haughey BH, Lund V, Niparko JK, Robbins KT, Thomas JR, et al., editors. Cummings otolaryngology head and neck surgery. 6th ed. Philadelphia: Saunders Elsevier; 2014. p. 1009-1010.
2. Karkos PD, Wilson JA. Empiric treatment of laryngopharyngeal reflux with proton pump inhibitors: a systematic review. *Laryngoscope*. 2006 Jan;116(1):144-148. DOI:10.1097/01.mlg.0000191463.67692.36 PubMed PMID: 16481828.
3. Lam PKY, Manwa LN, Cheung TK, Wong BYH, Tan VPY, Fong DYT, et al. Rabeprazole is effective in treating laryngopharyngeal reflux in a randomized placebo-controlled trial. *Clin Gastroenterol Hepatol*. 2010 Sep;8(9):770-776. DOI:10.1016/j.cgh.2010.03.009 PubMed PMID: 20303417.
4. Belafsky PC, Postma GN, Koufman JA. Validity and reliability of the reflux symptom index (RSI). *J Voice*. 2002 Jun;16(2):274-277. DOI:10.1016/s0892-1997(02)00097-8. PubMed PMID: 12150380.
5. Belafsky PC, Postma GN, Koufman JA. The validity and reliability of the reflux finding score (RFS). *Laryngoscope*. 2001 Aug;111(8):1313-1317. DOI:10.1097/00005537-200108000-00001 PubMed PMID: 11568561.
6. Belafsky PC, Postma GN, Koufman JA. Laryngopharyngeal reflux symptoms improve before changes in physical findings. *Laryngoscope*. 2001 Jun;111(6):979-981. DOI:10.1097/00005537-200106000-00009 PubMed PMID: 11404607.
7. Reichel O, Dressel H, Wiederänders K, Issing WJ. Double-blind, placebo-controlled trial with esomeprazole for symptoms and signs associated with laryngopharyngeal reflux. *Otolaryngol Head Neck Surg*. 2008 Sep;139(3):414-420. DOI:10.1016/j.otohns.2008.06.003 PubMed PMID: 18722223.
8. Liu C, Wang H, Liu K. Meta-analysis of the efficacy of proton pump inhibitors for the symptoms of laryngopharyngeal reflux. *Braz J Med Biol Res*. 2016 Jul;49(7):e5149. DOI:10.1590/1414-431X20165149 PubMed PMID: 27383119 PubMed Central PMCID: PMC4942224.
9. Megwalu UC. A systematic review of proton pump inhibitor therapy for laryngopharyngeal reflux. *Ear Nose Throat J*. 2013 Aug;92(8):364-371. DOI:10.1177/014556131309200811 PubMed PMID: 23975490.
10. Spantideas N, Drosou E, Bougea A, AlAbdulwahed R. Proton Pump Inhibitors for the Treatment of Laryngopharyngeal Reflux. A Systematic Review. *J Voice*. 2020 Nov;34(6):918-929. DOI:10.1016/j.jvoice.2019.05.005 PubMed PMID: 31160182.
11. Sen P, Georgalas C, Bhattacharyya AK. A systematic review of the role of proton pump inhibitors for symptoms of laryngopharyngeal reflux. *Clin Otolaryngol*. 2006 Feb;31(1):20-24. DOI:10.1111/j.1749-4486.2006.01134.x. PubMed PMID: 16441796.
12. Gatta L, Vaira D, Sorrenti G, Zucchini S, Sama C, Vakil N. Meta-analysis: the efficacy of proton pump inhibitors for laryngeal symptoms attributed to gastro-oesophageal reflux disease. *Aliment Pharmacol Ther.* 2007 Feb;25(4):385-392. DOI:10.1111/j.1365-2036.2006.03213.x. PubMed PMID: 17269993.
13. Guo H, Ma H, Wang J. Proton Pump Inhibitor Therapy for the Treatment of Laryngopharyngeal Reflux: A Meta-Analysis of Randomized Controlled Trials. *J Clin Gastroenterol*. 2016 Apr;50(4):295-300. DOI:10.1097/MCG.0000000000000324. PubMed PMID: 25906028.
14. Wei C. A meta-analysis for the role of proton pump inhibitor therapy in patients with laryngopharyngeal reflux. *Eur Arch Otorhinolaryngol*. 2016 Nov;273(11):3795-3801. DOI:10.1007/s00405-016-4142-y PubMed PMID: 27312992.
15. Lechien JR, Saussez S, Schindler A, Karkos PD, Hamdan AL, Harmegnies B, et al. Clinical outcomes of laryngopharyngeal reflux treatment: A systematic review and meta-analysis. *Laryngoscope*. 2019 May;129(5):1174-1187. DOI:10.1002/lary.27591 PubMed PMID: 30597577.
16. Higgins JP, Altman DG, Gøtzsche PC, Jüni P, Moher D, Oxman AD, et al. The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ*. 2011;343:d5928. DOI:10.1136/bmj.d5928 PubMed PMID: 22008217 PubMed Central PMCID: PMC3196245.
17. Faruqi S, Molyneux ID, Fathi H, Wright C, Thompson R, Morice AH. Chronic cough and esomeprazole: a double-blind placebo-controlled parallel study. *Respirology*. 2011 Oct;16(7):1150-1156. DOI:10.1111/j.1440-1843.2011.02014.x PubMed PMID: 21707852.
18. Wo JM, Koopman J, Harrell SP, Parker K, Winstead W, Lentsch E. Double-blind, placebo-controlled trial with single-dose pantoprazole for laryngopharyngeal reflux. *Am J Gastroenterol*. 2006 Sep;101(9):1972-2169. DOI:10.1111/j.1572-0241.2006.00693.x PubMed PMID: 16968502.
19. Wilson JA, Stocken DD, Watson GC, Fouweather T, McGlashan J, Mackenzie K, et al. Lansoprazole for persistent throat symptoms in secondary care: the TOPPITS RCT. *Health Technol Assess*. 2021 Jan;25(3):1-118. DOI:10.3310/hta25030 PubMed PMID: 33492208 PubMed Central PMCID: PMC7869007.
20. O'Hara J, Stocken DD, Watson GC, Fouweather T, McGlashan J, Mackenzie K, *et al.* Use of proton pump inhibitors to treat persistent throat symptoms: multicentre, double blind, randomised, placebo controlled trial. *BMJ*. 2021;372:m4903. DOI: 10.1136/bmj.m4903 PubMed PMID: 33414239 PubMed Central PMCID: PMC7789994.
21. Pawar S, Lim HJ, Gill M, Smith TL, Merati A, Toohill RJ, et al. Treatment of postnasal drip with proton pump inhibitors: a prospective, randomized, placebo-controlled study. *Am J Rhinol*. 2007 Nov-Dec;21(6):695-701. DOI:10.2500/ajr.2007.21.3098 PubMed PMID: 18201449.
22. Shaheen NJ, Crockett SD, Bright SD, Madanick RD, Buckmire R, Couch M, et al. Randomised clinical trial: high-dose acid suppression for chronic cough - a double-blind, placebo-controlled study. *Aliment Pharmacol Ther*. 2011 Jan;33(2):225-234. DOI:10.1111/j.1365-2036.2010.04511.x. PubMed PMID: 21083673.
23. Anzić SA, Turkalj M, Župan A, Labor M, Plavec D, Baudoin T. Eight weeks of omeprazole 20 mg significantly reduces both laryngopharyngeal reflux and comorbid chronic rhinosinusitis signs and symptoms: Randomised, double-blind, placebo-controlled trial. *Clin Otolaryngol*. 2018 Apr;43(2):496-501. DOI:10.1111/coa.13005 PubMed PMID: 29024410.
24. Borenstein M, Higgins JP, Hedges LV, Rothstein HR. Basics of meta-analysis: I^2 is not an absolute measure of heterogeneity. *Res Synth Methods*. 2017 Mar;8(1):5-18. DOI:10.1002/jrsm.1230. PubMed PMID: 28058794.
25. Page MJ, Higgins JPT, Sterne JAC. Chapter 13: Assessing risk of bias due to missing results in a synthesis. In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al., editors. Cochrane handbook for systematic reviews of interventions version 6.2 [Internet] [updated 2021 Feb; cited 2021 Apr 2]. Available from www.training.cochrane.org/handbook/current/chapter-13.
26. Altman KW, Prufer N, Vaezi MF. A review of clinical practice guidelines for reflux disease: toward creating a clinical protocol for the otolaryngologist. *Laryngoscope*. 2011 Apr;121(4):717-723. DOI:10.1002/lary.21429. PubMed PMID: 21298646.
27. Stachler RJ, Francis DO, Schwartz SR, Damask CC, Digoy GP, Krouse HJ, et al. Clinical practice guideline: Hoarseness (dysphonia) (update). *Otolaryngol Head Neck Surg.* 2018 Mar;158(1_suppl):S1-S42. DOI:10.1177/0194599817751030. PubMed PMID: 29494321. Erratum in: *Otolaryngol Head Neck Surg*. 2018 Aug;159(2):403. DOI: 10.1177/0194599818766900. PubMed PMID: 29494321.
28. American Academy of Otolaryngology-Head and Neck Surgery. Position statement: Laryngopharyngeal reflux [Internet]. Virginia, USA: AAO-HNS/F; 2006 [updated 2020 Aug 11; cited 2021 Apr 2]. Available from: https://www.entnet.org/content/laryngopharyngeal-reflux.
29. Lee JS, Lee YC, Kim SW, Kwon KH, Eun YG. Changes in the quality of life of patients with laryngopharyngeal reflux after treatment. *J Voice*. 2014 Jul;28(4):487-491. DOI:10.1016/j.jvoice.2013.12.015. PubMed PMID: 24598356.
30. Yoon YH, Park KW, Lee SH, Park HS, Chang JW, Koo BS. Efficacy of three proton-pump inhibitor therapeutic strategies on laryngopharyngeal reflux disease; a prospective randomized double-blind study. *Clin Otolaryngol*. 2019 Jul;44(4):612-618. DOI:10.1111/coa.13345. PubMed PMID: 31002475.
31. Doshi K, Varghese A, Badyal DK. Evaluation of omeprazole in the treatment of laryngopharyngeal reflux disease: A single center, prospective and randomized study. *Int J Otorhinolaryngol Clin.* 2015 Jan-Apr;7(2):45-50. DOI:10.5005/jp-journals-10003-1186 Corpus ID: 36862984.
32. Bhargava A, Faiz SM, Srivastava MR, Shakeel M, Singh NJ. Role of Proton Pump Inhibitors in Laryngopharyngeal Reflux: Clinical Evaluation in a North Indian Population. *Indian J Otolaryngol Head Neck Surg*. 2019 Sep;71(3):371-377. DOI:10.1007/s12070-018-1493-2. PubMed PMID: 31559206 PubMed Central PMCID: PMC6737127.
33. Chugh R. The role of empirical treatment with proton pump inhibitor as a diagnostic tool in laryngopharyngeal reflux. *Indian J Otolaryngol Head Neck Surg*. 2019 Oct:1-8. DOI:10.1007/s12070-019-01750-7.
34. Eherer AJ, Habermann W, Hammer HF, Kiesler K, Friedrich G, Krejs GJ. Effect of pantoprazole on the course of reflux-associated laryngitis: a placebo-controlled double-blind crossover study. *Scand J Gastroenterol.* 2003 May;38:462–467. DOI:10.1080/00365520310001860 PubMed PMID: 12795454.
35. Steward DL, Wilson KM, Kelly DH, Patil MS, Schwartzbauer HR, Long JD, et al. Proton pump inhibitor therapy for chronic laryngo-pharyngitis: a randomized placebo-control trial. *Otolaryngol Head Neck Surg.* 2004 Oct;131:342–350. DOI:10.1016/j.otohns.2004.03.037 PubMed PMID: 15467597.
36. Vaezi MF, Richter JE, Stasney CR, Spiegel JR, Iannuzzi RA, Crawley JA, et al. Treatment of chronic posterior laryngitis with esomeprazole. *Laryngoscope.* 2006 Feb;116:254–260. DOI:10.1097/01.mlg.0000192173.00498.ba PubMed PMID: 16467715.
37. Ezzat WF, Fawaz SA, Fathey H, El Demerdash A. Virtue of adding prokinetics to proton pump inhibitors in the treatment of laryngopharyngeal reflux disease: prospective study. *J Otolaryngol Head Neck Surg.* 2011 Aug;40: 350–356. PubMed PMID: 21777555.
38. Page MJ, Higgins JPT, Sterne JAC. Chapter 10: Analysing data and undertaking meta-analyses. In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al., editors. Cochrane handbook for systematic reviews of interventions version 6.2 [Internet] [updated 2021 Feb; cited 2021 Apr 2]. Available from www.training.cochrane.org/handbook/current/chapter-10.