

RESEARCH ARTICLE

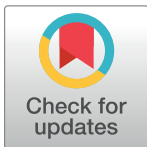
# Evaluation of the Quality of Reporting of Observational Studies in Otorhinolaryngology - Based on the STROBE Statement

Martine Hendriksma<sup>1</sup>✉, Michiel H. M. A. Joosten<sup>1</sup>✉, Jeroen P. M. Peters<sup>1,2</sup>, Wilko Grolman<sup>1,2</sup>, Inge Stegeman<sup>1,2\*</sup>

**1** Department of Otorhinolaryngology and Head & Neck Surgery, University Medical Center Utrecht, Utrecht, the Netherlands, **2** Brain Center Rudolf Magnus, University Medical Center Utrecht, Utrecht, the Netherlands

✉ These authors contributed equally to this work.

\* [i.stegeman@umcutrecht.nl](mailto:i.stegeman@umcutrecht.nl)



## Abstract

### Background

Observational studies are the most frequently published studies in literature. When randomized controlled trials cannot be conducted because of ethical or practical considerations, an observational study design is the first choice. The STROBE Statement (*STrengthening the Reporting of OBservational studies in Epidemiology*) was developed to provide guidance on how to adequately report observational studies.

### Objectives

The objectives were 1) to evaluate the quality of reporting of observational studies of otorhinolaryngologic literature using the STROBE Statement checklist, 2) to compare the quality of reporting of observational studies in the top 5 Ear, Nose, Throat (ENT) journals versus the top 5 general medical journals and 3) to formulate recommendations to improve adequate reporting of observational research in otorhinolaryngologic literature.

### Methods

The top 5 general medical journals and top 5 otorhinolaryngologic journals were selected based on their ISI Web of Knowledge impact factors. On August 3<sup>rd</sup>, 2015, we performed a PubMed search using different filters to retrieve observational articles from these journals. Studies were selected from 2010 to 2014 for the general medical journals and from 2015 for the ENT journals. We assessed all STROBE items to examine how many items were reported adequately for each journal type.

### Results

The articles in the top 5 general medical journals (n = 11) reported a mean of 69.2% (95% confidence interval (CI): 65.8%–72.7%; median 70.6%), whereas the top 5 ENT journals (n = 29) reported a mean of 51.4% (95% CI: 47.7%–55.0%; median 50.0%). The two journal types reported STROBE items significantly different ( $p < .001$ ).

## OPEN ACCESS

**Citation:** Hendriksma M, Joosten MHMA, Peters JPM, Grolman W, Stegeman I (2017) Evaluation of the Quality of Reporting of Observational Studies in Otorhinolaryngology - Based on the STROBE Statement. PLoS ONE 12(1): e0169316. doi:10.1371/journal.pone.0169316

**Editor:** Pei-Yi Chu, Fu Jen Catholic University, TAIWAN

**Received:** April 15, 2016

**Accepted:** November 30, 2016

**Published:** January 6, 2017

**Copyright:** © 2017 Hendriksma et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

**Data Availability Statement:** All relevant data are within the paper and its Supporting Information files.

**Funding:** The authors received no specific funding for this work.

**Competing Interests:** The authors have declared that no competing interests exist.

## Conclusion

Quality of reporting of observational studies in otorhinolaryngologic articles can considerably enhance. The quality of reporting was better in general medical journals compared to ENT journals. To improve the quality of reporting of observational studies, we recommend authors and editors to endorse and actively implement the STROBE Statement.

## Introduction

Most published studies in literature are observational studies. Observational studies are recognized as a useful insight into identifying best practices and addressing new hypotheses. Sometimes observational studies initiate the need for further investigation and lead to the emergence of randomized controlled trials (RCTs) [1]. Observational studies also offer the opportunity to establish high external validity in a practical setting, which is difficult to achieve in RCTs [2]. While RCTs are being advocated as the gold standard for the evaluation of treatment oriented interventions, well designed observational studies have been shown to provide similar results [3–5]. Because of their design, observational studies are prone to bias, confounding, cause and chance [6]. However, when RCTs cannot be conducted because of ethical or practical considerations, an observational study design is the first choice [7,8]. Clear presentation of its methods, execution and analyses is crucial for its valid implementation in clinical care. Poor reporting of observational studies can further reduce their usefulness.

Therefore, in 2004 [9] the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement was introduced to improve the quality of reporting of observational studies. The aim of the STROBE statement is to increase transparency in reporting [9]. A checklist made of 22 items was constructed in order to assess the quality of reporting of observational studies. The STROBE Statement is currently being endorsed by a growing number of biomedical journals [10].

To our knowledge, this quality of reporting in otorhinolaryngology has never been assessed before. In otorhinolaryngology, observational studies are an often used study design. Therefore, our objective was to assess the quality of reporting of observational research in otorhinolaryngologic literature using the STROBE Checklist. As a second objective, we aimed to compare the quality of reporting between the top 5 otorhinolaryngologic journals and the top 5 general medical journals, as a benchmark for presumably high quality of reporting. Finally, based on our findings, recommendations were forged in order to improve the reporting of observational studies in otorhinolaryngology.

## Methods

### Journals

We selected the top 5 medical journals (*New England Journal of Medicine* (NEJM), *The Lancet*, *Journal of the American Medical Association* (JAMA), *British Medical Journal* (BMJ) and *PLOS Medicine* (PLOS Med)) and the top 5 Ear Nose Throat (ENT) journals (*Rhinology*, *Hearing Research* (Hear Res), *Ear & Hearing* (Ear Hear), *Head & Neck* and *Journal of the Association for Research in Otolaryngology* (JARO)) based on their ISI Web of Knowledge impact factors.

**Table 1** presents the ISI Web of Knowledge impact factors of 2012 for general medical journals and of 2014 for ENT journals ([www.webofknowledge.com](http://www.webofknowledge.com), date of access August 3<sup>rd</sup>, 2015).

**Table 1. Top 5 general medical journals and ENT journals impact factors and STROBE endorsement.**

Rank	Title	Impact factor*	Endorse STROBE?
<b>General medical journals</b>			
1	New England Journal of Medicine (NEJM)	51.658	No
2	The Lancet	39.060	Yes
3	Journal of the American Medical Association (JAMA)	29.978	No
5	The British Medical Journal (BMJ)	17.215	Yes
7	PLOS Medicine (PLOS Med)	15.976	Yes
<b>ENT Journals</b>			
1	Rhinology	3.761	No
2	Hearing Research (Hearing Res)	2.968	No
3	Ear & Hearing (Ear Hear)	2.842	No
4	Head & Neck (Head Neck)	2.641	No
5	Journal of the Association for Research in Otolaryngology (JARO)	2.598	No

\* Source: ISI Web of Knowledge. For general medical journals impact factors of 2012. For ENT journals impact factors of 2014. Both assessed on August 3<sup>rd</sup>, 2015.

ENT = Ear Nose Throat.

doi:10.1371/journal.pone.0169316.t001

In addition, we evaluated the ‘Instruction to authors’ section on the websites of the included journals to check if they endorsed the STROBE statement (see also [Table 1](#)).

### Search

We performed a PubMed search on August 3<sup>rd</sup> 2015 using several search syntaxes. First, to retrieve only observational studies, a syntax developed by the Scottish Intercollegiate Guidelines Network was used [11]. Second, an adapted version of the Cochrane ENT search syntax was used to retrieve otorhinolaryngologic articles [12]. Third, a filter was used to retrieve articles published in the top 5 general medical journals ([Table 1](#)) and articles published in the top 5 ENT journals ([Table 1](#)) separately. Fourth, a date restriction filter was applied. To retrieve sufficient otorhinolaryngologic articles published in general medical journals, we searched from January 1<sup>st</sup>, 2010 until December 31<sup>st</sup>, 2014. We did not search for articles earlier than 2010, since the STROBE Statement was first published in 2007. Hereby, we allowed for sufficient time for implementation of the STROBE Statement by study authors. Naturally, it is easier to retrieve otorhinolaryngologic articles in ENT journals, so for articles published in ENT journals we applied a date restriction from January 1<sup>st</sup>, 2015 until August 3<sup>rd</sup>, 2015. Finally, a combination of search syntaxes was made to retrieve observational studies from general medical journals and ENT journals respectively (see [S1 File](#)).

### Study selection

Two authors (MH and MHMAJ) independently assessed titles, abstracts and full texts of the retrieved articles to check if the study was indeed an observational study and if it was conducted in the otorhinolaryngologic field. To be considered an observational study, studies must have had a cross-sectional, cohort or case control design. To be considered as a study in the otorhinolaryngologic field, studies must have assessed patient populations generally treated by ENT physicians or procedures generally performed by ENT physicians, including procedures performed by head and neck surgeons. Discrepancies between the two reviewers were discussed until consensus was reached.

## Strobe statement adherence

To score the quality of reporting, the most recent version of the STROBE statement was used [9]. The included articles were read and scored independently by two authors (MH and MHMAJ). We evaluated the items of the STROBE checklist that were adequately reported. The total number of items on the STROBE checklist is 34 (subitems included). Some items (6a, 6b, 12d, 14c, 15) are specific for some study designs only (e.g. cohort or case control). Consequently, if an item was not applicable for the study design, it was scored as 'not applicable'. For a more detailed description of the requirements to score 'adequately reported', see [S2 File](#).

Because of an imbalance in the number of articles per journal category, we scored one article from the general medical journals for every three articles from the ENT journals, so a possible learning effect in use of the STROBE checklist was distributed evenly across both journal categories. Differences in opinion were discussed until consensus was reached.

## Data analysis

We divided the number of adequately reported items by the total number of applicable items, which resulted in a proportion of adequately reported items. A higher proportion reflects that the item was reported more adequately. Mean, medians and 95% confidence intervals (CI) were calculated per item.

The 2-tailed Mann Whitney U test for two independent samples was used to test if there was a statistically significant difference between the STROBE scores for articles published in the top 5 general medical journals and in the top 5 ENT journals.

Furthermore, the interobserver agreement (Cohen's kappa) was calculated to determine if there were large differences in the scoring of items between the two authors.

Statistical tests were performed using SPSS v21 statistics package. A  $p$ -value of  $< .05$  was considered statistically significant.

## Results

### Search

The search process is shown in [Fig 1](#). The combined search syntaxes yielded 42 articles from general medical journals and 44 from ENT journals.

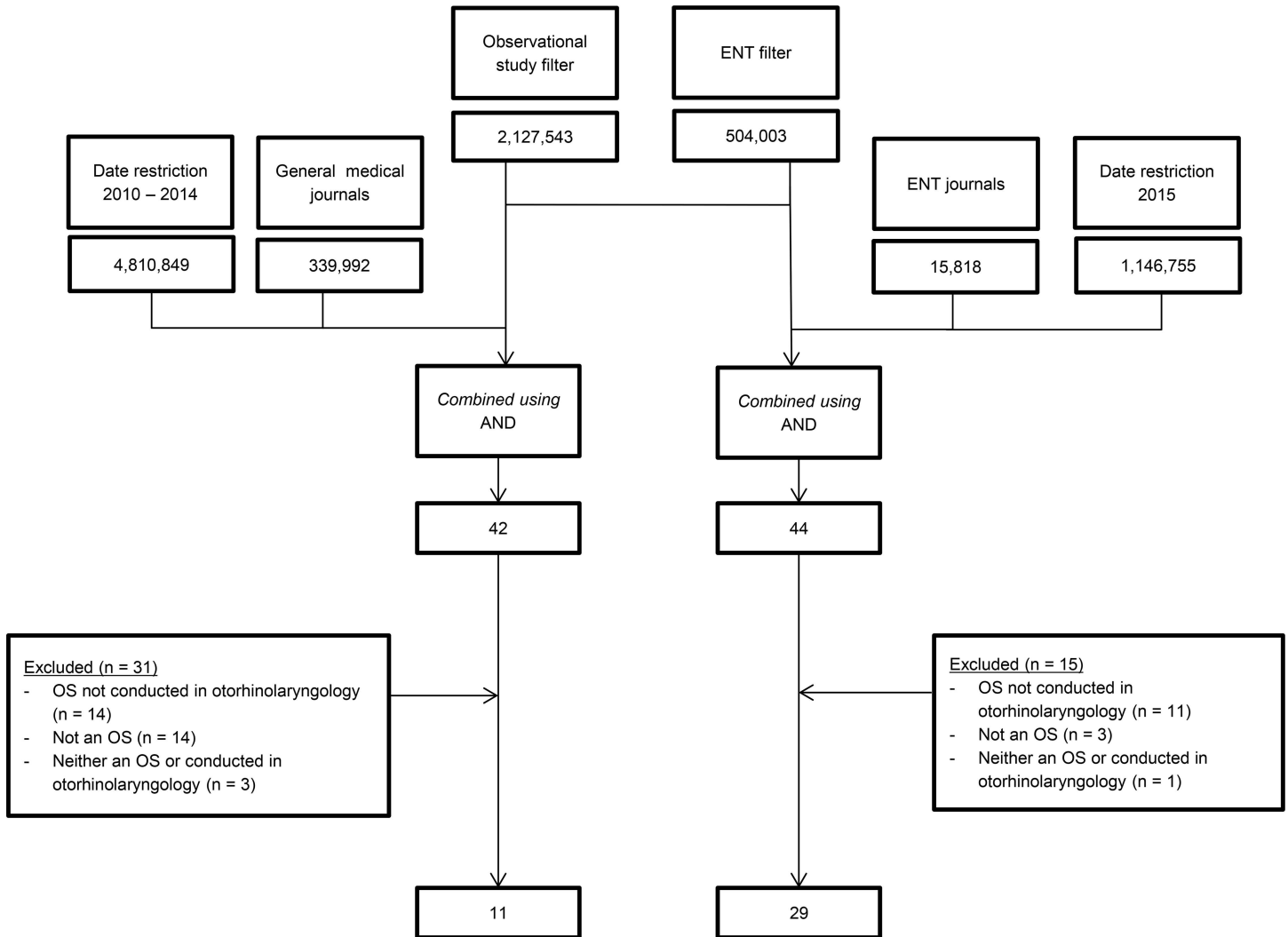
### Study selection

Of the 42 articles in the general medical journals, 14 articles were not considered otorhinolaryngologic articles, 14 articles were considered not to be an observational study and 3 articles were neither considered an otorhinolaryngologic nor an observational study. Finally, we included 11 observational studies on otorhinolaryngologic topics from the general medical journals.

Of the 44 articles in the ENT journals, 11 articles were not considered to be an observational study, 3 articles were not considered otorhinolaryngologic articles, and 1 article was considered neither an otorhinolaryngologic nor an observational study. Finally, we included 29 observational studies on otorhinolaryngologic topics from ENT journals.

### STROBE Statement adherence

The 11 articles published in general medical journals (NEJM = 1, Lancet = 1, JAMA = 4, BMJ = 2, PLOS Med = 3) reported a mean of 69.2% (95% CI: 65.8%–72.7%; median 70.6%) of STROBE items adequately. The 29 articles published in ENT journals (Rhinology = 1, Hear Res = 0, Ear & Hearing = 8, Head & Neck = 20 and JARO = 0) reported a mean of 51.4% (95%



**Fig 1. Flowchart of search.** Date of search: August 3<sup>rd</sup>, 2015. For complete search syntax, see [S1 File](#). ENT = Ear, Nose, Throat, OS: observational study, N = number.

doi:10.1371/journal.pone.0169316.g001

CI: 47.7%–55.0%; median 50.0%) of STROBE items adequately. The exact percentage of adequately reported STROBE items for studies published in general medical journals and ENT journals are presented in [Table 2](#). A graphic illustration is provided in [Fig 2](#).

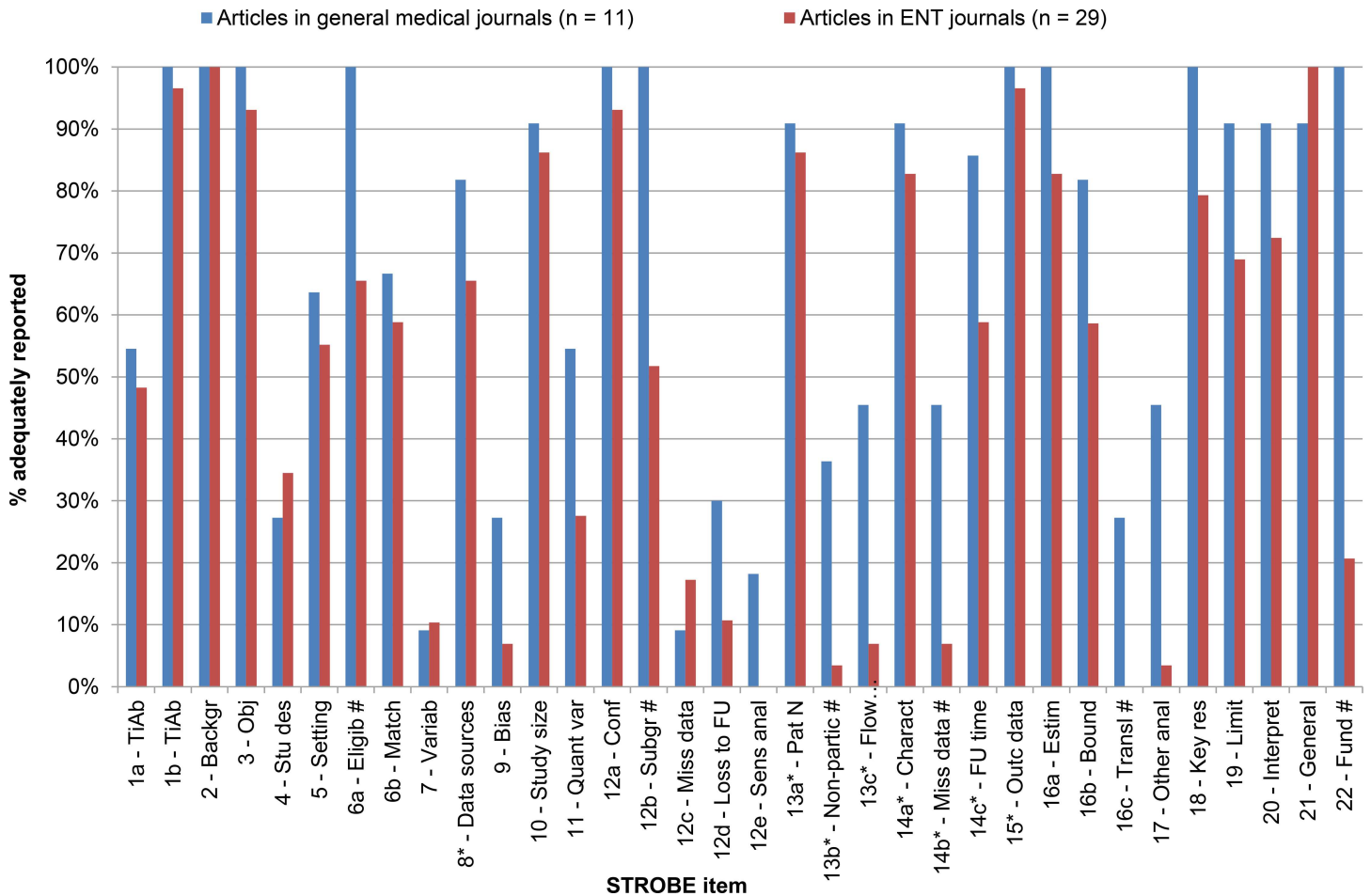
We compared the difference in reporting individual items between observational studies published in general medical journals and ENT journals. The ENT journals reported several items inadequately more frequently than general medical journals. The difference in reporting between journal categories was statistically significant for 8 out of 34 items (highlighted with a # in [Fig 2](#)). First, less than 25% of the articles published in ENT journals reported funding and the role of the funders, whereas this was reported in 100% of the articles published in general medical journals (item 22). Second, several other items were scored significantly better in general medical journals than in ENT journals. Item 6a (eligibility criteria and the sources of methods of selection of participation) and item 8 (data sources and measurement for each variable of interest) were scored better by general medical journals, as well as item 12b (methods used to examine subgroups and interaction). Likewise for items 13a, 13b and 13c (reporting of

**Table 2. data table of Fig 2.** Percentage of adequately reported STROBE items per journal category.

Item	1a - TiAb	1b - TiAb	2 - Backgr	3 - Obj	4 - Studes	5 - Setting	6a - Eligib	6b - Match	7 - Variab	8 - Data sourc	9 - Bias	10 - Study size	11 - Quant var	12a - Conf	12b - Subgr	12c - Miss data	12d - Loss to FU
Articles in general medical journals (n = 11)	55%	100%	100%	100%	27%	64%	100%	67%	9%	82%	27%	91%	55%	100%	100%	9%	30%
Articles in ENT journals (n = 29)	48%	97%	100%	93%	34%	55%	66%	59%	10%	66%	7%	86%	28%	93%	52%	17%	11%
p-value Chi2 (2-tailed)	1,000	1,000	N/A	1,000	1,000	0,730	<b>0,038</b>	1,000	1,000	0,451	0,117	1,000	0,147	1,000	<b>0,004</b>	1,000	0,310
Item	12e - Sens anal	13a - PatN	13b - Non-partic	13c - Flow Diagram	14a - Charact	14b - Miss data	14c - FU time	15 - Outc data	16a - Estim	16b - Bound	16c - Transl	17 - Other anal	18 - Key res	19 - Limit	20 - Interpret	21 - General	22 - Fund
Articles in general medical journals (n = 11)	18%	91%	36%	45%	91%	45%	86%	100%	100%	82%	27%	45%	100%	91%	91%	91%	100%
Articles in ENT journals (n = 29)	0%	86%	3%	7%	83%	7%	59%	97%	83%	59%	0%	3%	79%	69%	72%	100%	21%
p-value Chi2 (2-tailed)	0,071	1,000	<b>0,015</b>	<b>0,011</b>	1,000	<b>0,011</b>	0,352	1,000	0,298	0,270	<b>0,017</b>	<b>0,004</b>	0,162	0,233	0,399	0,275	<b>0,000</b>

**Legend:** Yellow cell and bold font = significant difference for quality of reporting between journal types (Chi2, p < 0.05). For an overview of all STROBE items see [S2 File](#). ENT = Ear Nose Throat.

doi:10.1371/journal.pone.0169316.t002



**Fig 2. Adequately reported STROBE items.** For an overview of all STROBE items see [S2 File](#). Items highlighted with a # are statistically significant between journal categories. ENT = Ear, Nose, Throat.

doi:10.1371/journal.pone.0169316.g002

numbers of individuals at each stage of the study, reasons for non-participation at each stage and the use of a flow diagram). Also items 14a, 14b and 14c (descriptive data of the study participants) and at last item 15 (the outcome data) were all scored significantly better in general medical journals.

Besides these differences, both types of journals scored insufficient on several other items. For example, both journal categories scored poorly in reporting the variables (item 7); also both categories of journals reported less than 20% on items 9, 12c, 12d and 12e. Moreover, none of the articles published in ENT journals reported sensitivity analyses (item 12e) or translated estimates of relative risk into absolute risk (item 16c).

The overall interobserver agreement kappa is 0.64. A kappa score between 0.61 and 0.80 means there is a good agreement [13]. [S3 File](#) shows the interobserver agreement per STROBE item for both journal categories, as well as both journal categories together. Items scored with the most discrepancy between the two authors, were the defining of outcomes (item 7), the handling of quantitative variables (item 11), the reporting of individuals at each stage of the study (item 13a), the study characteristics of participants (item 14a), the reporting of unadjusted estimates (item 16c), the overall interpretation of results (item 20) and the generalizability (item 21).

## Discussion

To the best of our knowledge, this study is the first to assess the quality of reporting of observational studies in otorhinolaryngologic literature. Furthermore, we made a comparison between the top 5 general medical journals and the top 5 ENT journals, by using the STROBE checklist. This study shows a substantial difference in the quality of reporting between ENT journals and general medical journals.

### Interpretation of results

First, notable differences were reported regarding the eligibility criteria (item 6a). Eligibility criteria help readers understand the applicability and generalizability of the reported results. In observational studies, these characteristics are often not adequately reported [14].

Second, sub-analyses are only reported in about 50% of the top 5 ENT journals. Being one of the items of the STROBE checklist, there is debate about this item (item 12b) [15,16]. The reporting of sub-analyses allows one to examine whether effects or associations differed across groups [17].

Third, missing data is common in observational research (item 14b). It is essential to describe for which variable of interest data was missing, because (in)voluntary exclusion of patients can lead to a distorted image in the results.

Fourth, other analyses were only reported in 3% of the articles published in ENT journals (item 17). However, these analyses are important, because they may address specific sub-groups, display potential interaction between risk factors, calculate attributable risks or use alternative definitions of study variables in sensitivity analyses [17].

Fifth, almost 80% of the ENT journals do not report funding (item 22), although there are strong associations between the source of funding and conclusions [18,19]. It is therefore important to be transparent about financial support, as specific types of sponsorships may be associated with positive study results [20].

Sixth, the use of a flow diagram is an easy way to display extensive information in a compact design (item 13c); 7% of the articles published in ENT journals made use of a flow diagram compared to 45% of the articles published in general medical journals.

Seventh, reasons for non-participation (item 13b) was reported in 3% of the ENT journals. This item lets readers judge whether the study population was representative of the target population [17].

Last, translating estimates of relative risk into absolute risk (item 16c) was not reported in any of the ENT articles. However, in some cases this item may not be relevant.

As for the top 5 general medical journals, 11 items scored particular lower than expected, i.e. less than 50%; items 4, 7, 9, 12c, 12d, 12e, 13b, 13c, 14b, 16c and 17. Of these, 6 items (item 4, 7, 9, 12c, 12d and 12e) scored similar poor compared to the top 5 ENT journals (no statistical difference).

### Comparison with literature

In other research fields, like ophthalmology and hand surgery, items like bias (item 9), the explanation of missing data (item 12c), use of a flow-diagram (item 13c) and indicating missing data for each variable of interest (item 14b) scored similarly insufficient to the articles published in ENT journals [21,22]. Items which scored better in these research fields were presenting key elements of study design early in the paper (item 4), defining all outcomes, exposures, predictors, potential confounders and effect modifiers (item 7), describing loss to follow-up/matching of cases and controls/sampling (item 12d) and displaying characteristics of study participants (item 14c). Only item 10, the explanation of how study size was arrived



at, was scored better in the articles published in ENT journals in our sample. In another study, Delaney et al. evaluated platelet transfusion studies according to the STROBE criteria [23]. They also scored missing data as reported inadequately, in accordance with our results. Unfortunately, an average compliance score was not calculated [23]. Parsons et al. [24] evaluated general orthopedic journals and found an overall average compliance to the STROBE checklist of 58% (95% CI: 56%–60%), close to our findings. There were 9 items that scored similarly poor (items 9, 12c, 12d, 12e, 13b, 13c, 14b and 16c). It was unclear whether the journals in which the included articles were published endorsed the STROBE Statement. Since the articles included by Parsons et al. were selected from the years 2005 to 2010, active endorsement was not entirely possible, as the STROBE Statement was first published in 2007.

Similar results were seen in the research field of plastic surgery, where the 9 items mentioned above along with item 17 scored similarly poor in reference to our results [25]. This study also suggested that reporting could possibly be improved by making the STROBE checklist mandatory at submission [25].

One other study showed that the quality of reporting of confounding improved in some aspects, but the authors concluded that the overall quality remains suboptimal [26]. Quantitative bias analysis (item 9) was scored very rarely, which we also observed.

Two reporting guidelines (Consolidated Standards of Reporting Trials (CONSORT) and Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA)) have been published before the STROBE Statement. Several studies found that the endorsement of the CONSORT and PRISMA Statements led to a better quality of reporting [27–34]. Two other articles evaluating the CONSORT Statement concluded that active endorsing would further increase the quality of reporting [35,36]. Unfortunately, no studies have been conducted evaluating the STROBE Statement for multiple disciplines simultaneously or investigating the implementation of the STROBE Statement. For better applicability of the STROBE Statement, some papers have adjusted the STROBE statement to their specific research field [37,38].

## Methodological considerations

Strengths of our study include the independent assessment of all articles by two authors. Besides, to accumulate the learning effect in reference to the imbalance in the number of articles per journal type, we scored one article from general medical journals for every three articles from ENT journals. All details of our search were clear and transparent and our search can therefore easily be reproduced. Furthermore, some of the authors were involved in the assessment of the PRISMA and CONSORT Statement checklists [39,40]. This scientific context and experience could have improved the overall quality of our study. Additionally an inter-observer agreement (Cohen's kappa) was calculated to determine which items were often disputed (these items deserve extra attention when assessing, and possibly even rephrasing by the STROBE workgroup, because they may be hard to interpret).

However, our study also has limitations. First, scoring of items remains a subjective task, with differences between observers. Several items were prone to discussion between the scoring authors. However, our analysis showed that we reached good interobserver agreement (Cohen's kappa 0.64). Second, only eleven studies were found for the top 5 general medical journals. This small sample size makes it more difficult to draw appropriate conclusions for this journal category. Third, the inclusion periods (top 5 general medical journals: January 1<sup>st</sup>, 2010 –December 31<sup>st</sup>, 2014; top 5 ENT journals January 1<sup>st</sup>, 2015 –August 3<sup>rd</sup>, 2015) of both journals types were not identical. As expected, less otorhinolaryngologic articles were published in general medical journals than in ENT journals. Even with a broader time frame in the search syntax for articles from general medical journals, this still resulted in less articles

( $n = 11$ ) than the number of articles included from ENT journals ( $n = 29$ ). However, retrieving earlier studies would not lead to a more accurate comparison, since the STROBE Statement was first published in 2007. Consequently, earlier studies would not have had time to implement the STROBE Statement. On the other hand, we did not broaden our date restriction for ENT journals, since 29 articles form a good sample to base conclusions on. Including otorhinolaryngologic articles from ENT journals from 2010–2014 would result in many more articles to score, leading to incomparable samples. Moreover, including more articles published in ENT journals would probably not have changed conclusions. Finally, by choosing the most recent articles published in ENT journals, we would expect maximum improved quality of reporting, because more time was available to implement the STROBE Statement. It would be interesting to evaluate the quality of reporting over a period of time. However, eleven articles from five years do not provide enough statistical power for such a comparison.

## Recommendations

Clear and transparent reporting in research papers facilitates clinicians and researchers that they can evaluate the validity of articles. The STROBE Statement was developed to aid researchers to report their observational studies adequately. Therefore, we think it is highly plausible that using the STROBE Statement will improve the quality of reporting. We recommend authors of otorhinolaryngologic articles to use the STROBE Statement and recommend editors of otorhinolaryngologic journals to actively endorse this statement in their guidelines for authors.

## Conclusion

Current quality of reporting of observational studies in otorhinolaryngologic journals is sub-optimal. According to the STROBE Statement checklist, the reporting of otorhinolaryngologic observational studies is significantly better in articles published in general medical journals (2010–2014) than in articles published in ENT journals (2015). We recommend authors of otorhinolaryngologic articles to use the STROBE Checklist to help them to improve the quality of reporting. Furthermore, we suggest editors of ENT journals to actively endorse the STROBE Statement in their submission process.

## Supporting Information

**S1 File. Search Syntaxes.** Date of search: August 3<sup>rd</sup>, 2015. We used a study syntax developed by the Scottish Intercollegiate Guidelines Network [11]. An adapted version of the Cochrane ENT search syntax was used to retrieve otorhinolaryngologic [12] articles. Finally, a date restriction was applied from January 1<sup>st</sup>, 2010 until December 31<sup>st</sup>, 2014 for the top 5 general medical journals and from January 1<sup>st</sup>, 2015 until August 3<sup>rd</sup>, 2015 for the top 5 ENT journals. (DOCX)

**S2 File. Explanation of reported items of the STROBE checklist.** The original STROBE Statement can be assessed by visiting the STROBE website [10] or see Von Elm et al. [9]. For an explanatory and elaborated view see Vandenbroucke et al. [17]. Items were either scored as ‘adequately reported’, ‘inadequately reported’ or ‘not applicable’. Five items of the STROBE Statement are specific for study design (6a, 6b, 12d, 14c, 15). If any of these items were not applicable for the study design, it was scored as ‘not applicable’. ‘Not applicable’ items were not added to the amount of items to score. (DOCX)

**S3 File. Kappa per STROBE item.** The interobserver agreement for each STROBE item was calculated for both journal categories, as well as both journal categories together. For an explanation of scale division for the interobserver agreement, see Altman [13]. (DOCX)

## Author Contributions

**Conceptualization:** IS JPMP.

**Data curation:** MHMAJ MH JPMP.

**Formal analysis:** JPMP.

**Investigation:** MHMAJ MH.

**Methodology:** IS JPMP.

**Project administration:** MHMAJ MH JPMP IS.

**Resources:** MHMAJ MH JPMP IS.

**Supervision:** JPMP IS WG.

**Validation:** IS JPMP.

**Visualization:** MHMAJ MH JPMP.

**Writing – original draft:** MHMAJ MH.

**Writing – review & editing:** MHMAJ MH JPMP IS WG.

## References

1. Glasziou P, Vandenbroucke JP, Chalmers I. Assessing the quality of research. *BMJ*. 2004; 328: 39–41. doi: [10.1136/bmj.328.7430.39](https://doi.org/10.1136/bmj.328.7430.39) PMID: [14703546](https://pubmed.ncbi.nlm.nih.gov/14703546/)
2. Black N. Why we need observational studies to evaluate the effectiveness of health care. *BMJ*. 1996; 312: 1215–1218. PMID: [8634569](https://pubmed.ncbi.nlm.nih.gov/8634569/)
3. Song JW, Chung KC. Observational studies: cohort and case-control studies. *Plast Reconstr Surg*. 2010; 126: 2234–2242. doi: [10.1097/PRS.0b013e3181f44abc](https://doi.org/10.1097/PRS.0b013e3181f44abc) PMID: [20697313](https://pubmed.ncbi.nlm.nih.gov/20697313/)
4. Benson K, Hartz AJ. A comparison of observational studies and randomized, controlled trials. *N Engl J Med*. 2000; 342: 1878–1886. doi: [10.1056/NEJM200006223422506](https://doi.org/10.1056/NEJM200006223422506) PMID: [10861324](https://pubmed.ncbi.nlm.nih.gov/10861324/)
5. Concato J, Shah N, Horwitz RI. Randomized, controlled trials, observational studies, and the hierarchy of research designs. *N Engl J Med*. 2000; 342: 1887–1892. doi: [10.1056/NEJM200006223422507](https://doi.org/10.1056/NEJM200006223422507) PMID: [10861325](https://pubmed.ncbi.nlm.nih.gov/10861325/)
6. Jepsen P, Johnsen SP, Gillman MW, Sorensen HT. Interpretation of observational studies. *Heart*. 2004; 90: 956–960. doi: [10.1136/hrt.2003.017269](https://doi.org/10.1136/hrt.2003.017269) PMID: [15253985](https://pubmed.ncbi.nlm.nih.gov/15253985/)
7. The periodic health examination. Canadian Task Force on the Periodic Health Examination. *Can Med Assoc J*. 1979; 121: 1193–1254. PMID: [115569](https://pubmed.ncbi.nlm.nih.gov/115569/)
8. Sackett DL. Rules of evidence and clinical recommendations on the use of antithrombotic agents. *Chest*. 1989; 95: 2S–4S. PMID: [2914516](https://pubmed.ncbi.nlm.nih.gov/2914516/)
9. von Elm E, Altman DG, Egger M, Pocock SJ, Gotsche PC, Vandenbroucke JP, et al. The Strengthening of Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *PLoS Med*. 2007; 4: e296. [pii]. doi: [10.1371/journal.pmed.0040296](https://doi.org/10.1371/journal.pmed.0040296) PMID: [17941714](https://pubmed.ncbi.nlm.nih.gov/17941714/)
10. STROBE statement Strengthening the reporting of observational studies in epidemiology. Jan 2016. Available: <http://www.strobe-statement.org>.
11. Scottish Intercollegiate Guidelines Network. Search filter for observational studies (adapted). Accessed August 3rd, 2015. Available: <http://www.sign.ac.uk/methodology/filters.html#obs>.

12. The editorial team, Cochrane Ear Nose and Throat Disorders Group. About the Cochrane Collaboration (Cochrane Reviews Group (CRGs)). 2012, issue 7, art. no.: ENT. CENTRAL search strategy.
13. Altman DG. Practical Statistics for medical research. Chapman and Hall. 1991.
14. Tooth L, Ware R, Bain C, Purdie DM, Dobson A. Quality of reporting of observational longitudinal research. *Am J Epidemiol*. 2005; 161: 280–288. [pii]. doi: [10.1093/aje/kwi042](https://doi.org/10.1093/aje/kwi042) PMID: [15671260](https://pubmed.ncbi.nlm.nih.gov/15671260/)
15. Pocock SJ, Collier TJ, Dandreo KJ, de Stavola BL, Goldman MB, Kalish LA, et al. Issues in the reporting of epidemiological studies: a survey of recent practice. *BMJ*. 2004; 329: 883. [pii]. doi: [10.1136/bmj.38250.571088.55](https://doi.org/10.1136/bmj.38250.571088.55) PMID: [15469946](https://pubmed.ncbi.nlm.nih.gov/15469946/)
16. Gotzsche PC. Believability of relative risks and odds ratios in abstracts: cross sectional study. *BMJ*. 2006; 333: 231–234. [pii]. doi: [10.1136/bmj.38895.410451.79](https://doi.org/10.1136/bmj.38895.410451.79) PMID: [16854948](https://pubmed.ncbi.nlm.nih.gov/16854948/)
17. Vandenberghe JP, von Elm E, Altman DG, Gotzsche PC, Mulrow CD, Pocock SJ, et al. Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): explanation and elaboration. *PLoS Med*. 2007; 4: e297. [pii]. doi: [10.1371/journal.pmed.0040297](https://doi.org/10.1371/journal.pmed.0040297) PMID: [17941715](https://pubmed.ncbi.nlm.nih.gov/17941715/)
18. Bekelman JE, Li Y, Gross CP. Scope and impact of financial conflicts of interest in biomedical research: a systematic review. *JAMA*. 2003; 289: 454–465. [pii]. PMID: [12533125](https://pubmed.ncbi.nlm.nih.gov/12533125/)
19. Lexchin J, Bero LA, Djulbegovic B, Clark O. Pharmaceutical industry sponsorship and research outcome and quality: systematic review. *BMJ*. 2003; 326: 1167–1170. doi: [10.1136/bmj.326.7400.1167](https://doi.org/10.1136/bmj.326.7400.1167) PMID: [12775614](https://pubmed.ncbi.nlm.nih.gov/12775614/)
20. Sun GH, Houlton JJ, MacEachern MP, Bradford CR, Hayward RA. Influence of study sponsorship on head and neck cancer randomized trial results. *Head Neck*. 2013; 35: 1515–1520. doi: [10.1002/hed.23151](https://doi.org/10.1002/hed.23151) PMID: [22987508](https://pubmed.ncbi.nlm.nih.gov/22987508/)
21. Fung AE, Palanki R, Bakri SJ, Depperschmidt E, Gibson A. Applying the CONSORT and STROBE statements to evaluate the reporting quality of neovascular age-related macular degeneration studies. *Ophthalmology*. 2009; 116: 286–296. doi: [10.1016/j.ophtha.2008.09.014](https://doi.org/10.1016/j.ophtha.2008.09.014) PMID: [19091408](https://pubmed.ncbi.nlm.nih.gov/19091408/)
22. Sorensen AA, Wojahn RD, Manske MC, Calfee RP. Using the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) Statement to assess reporting of observational trials in hand surgery. *J Hand Surg Am*. 2013; 38: 1584–9.e2. doi: [10.1016/j.jhsa.2013.05.008](https://doi.org/10.1016/j.jhsa.2013.05.008) PMID: [23845586](https://pubmed.ncbi.nlm.nih.gov/23845586/)
23. Delaney M, Meyer E, Cserti-Gazdewich C, Haspel RL, Lin Y, Morris A, et al. A systematic assessment of the quality of reporting for platelet transfusion studies. *Transfusion*. 2010; 50: 2135–2144. doi: [10.1111/j.1537-2995.2010.02691.x](https://doi.org/10.1111/j.1537-2995.2010.02691.x) PMID: [20497518](https://pubmed.ncbi.nlm.nih.gov/20497518/)
24. Parsons NR, Hiskens R, Price CL, Achten J, Costa ML. A systematic survey of the quality of research reporting in general orthopaedic journals. *J Bone Joint Surg Br*. 2011; 93: 1154–1159 doi: [10.1302/0301-620X.93B9.27193](https://doi.org/10.1302/0301-620X.93B9.27193) PMID: [21911523](https://pubmed.ncbi.nlm.nih.gov/21911523/)
25. Agha RA, Lee SY, Jeong KJ, Fowler AJ, Orgill DP. Reporting Quality of Observational Studies in Plastic Surgery Needs Improvement: A Systematic Review. *Ann Plast Surg*. 2016; 76: 585–589 doi: [10.1097/SAP.0000000000000419](https://doi.org/10.1097/SAP.0000000000000419) PMID: [25643190](https://pubmed.ncbi.nlm.nih.gov/25643190/)
26. Pouwels KB, Widyakusuma NN, Groenwold RH, Hak E. Quality of reporting of confounding remained suboptimal after the STROBE guideline. *J Clin Epidemiol*. 2016; 69: 217–224. doi: [10.1016/j.jclinepi.2015.08.009](https://doi.org/10.1016/j.jclinepi.2015.08.009) PMID: [26327488](https://pubmed.ncbi.nlm.nih.gov/26327488/)
27. Plint AC, Moher D, Morrison A, Schulz K, Altman DG, Hill C, et al. Does the CONSORT checklist improve the quality of reports of randomised controlled trials? A systematic review. *Med J Aust*. 2006; 185: 263–267. [pii]. PMID: [16948622](https://pubmed.ncbi.nlm.nih.gov/16948622/)
28. Moher D, Jones A, Lepage L, CONSORT Group (Consolidated Standards for Reporting of Trials). Use of the CONSORT statement and quality of reports of randomized trials: a comparative before-and-after evaluation. *JAMA*. 2001; 285: 1992–1995. [pii]. PMID: [11308436](https://pubmed.ncbi.nlm.nih.gov/11308436/)
29. Turner L, Shamseer L, Altman DG, Weeks L, Peters J, Kober T, et al. Consolidated standards of reporting trials (CONSORT) and the completeness of reporting of randomised controlled trials (RCTs) published in medical journals. *Cochrane Database Syst Rev*. 2012; 11: MR000030. doi: [10.1002/14651858.MR000030.pub2](https://doi.org/10.1002/14651858.MR000030.pub2) PMID: [23152285](https://pubmed.ncbi.nlm.nih.gov/23152285/)
30. Tao KM, Li XQ, Zhou QH, Moher D, Ling CQ, Yu WF. From QUOROM to PRISMA: a survey of high-impact medical journals' instructions to authors and a review of systematic reviews in anesthesia literature. *PLoS One*. 2011; 6: e27611. doi: [10.1371/journal.pone.0027611](https://doi.org/10.1371/journal.pone.0027611) PMID: [22110690](https://pubmed.ncbi.nlm.nih.gov/22110690/)
31. Fleming PS, Seehra J, Polychronopoulou A, Fedorowicz Z, Pandis N. A PRISMA assessment of the reporting quality of systematic reviews in orthodontics. *Angle Orthod*. 2013; 83: 158–163. doi: [10.2319/032612-251.1](https://doi.org/10.2319/032612-251.1) PMID: [22720835](https://pubmed.ncbi.nlm.nih.gov/22720835/)
32. Panic N, Leoncini E, de Belvis G, Ricciardi W, Boccia S. Evaluation of the endorsement of the preferred reporting items for systematic reviews and meta-analysis (PRISMA) statement on the quality of published systematic review and meta-analyses. *PLoS One*. 2013; 8: e83138. doi: [10.1371/journal.pone.0083138](https://doi.org/10.1371/journal.pone.0083138) PMID: [24386151](https://pubmed.ncbi.nlm.nih.gov/24386151/)

33. Tunis AS, McInnes MD, Hanna R, Esmail K. Association of study quality with completeness of reporting: have completeness of reporting and quality of systematic reviews and meta-analyses in major radiology journals changed since publication of the PRISMA statement? *Radiology*. 2013; 269: 413–426. doi: [10.1148/radiol.13130273](https://doi.org/10.1148/radiol.13130273) PMID: [23824992](https://pubmed.ncbi.nlm.nih.gov/23824992/)
34. Stevens A, Shamseer L, Weinstein E, Yazdi F, Turner L, Thielman J, et al. Relation of completeness of reporting of health research to journals' endorsement of reporting guidelines: systematic review. *BMJ*. 2014; 348: g3804. doi: [10.1136/bmj.g3804](https://doi.org/10.1136/bmj.g3804) PMID: [24965222](https://pubmed.ncbi.nlm.nih.gov/24965222/)
35. Pandis N, Shamseer L, Kokich VG, Fleming PS, Moher D. Active implementation strategy of CONSORT adherence by a dental specialty journal improved randomized clinical trial reporting. *J Clin Epidemiol*. 2014; 67: 1044–1048. doi: [10.1016/j.jclinepi.2014.04.001](https://doi.org/10.1016/j.jclinepi.2014.04.001) PMID: [24837296](https://pubmed.ncbi.nlm.nih.gov/24837296/)
36. Hopewell S, Ravaud P, Baron G, Boutron I. Effect of editors' implementation of CONSORT guidelines on the reporting of abstracts in high impact medical journals: interrupted time series analysis. *BMJ*. 2012; 344: e4178. doi: [10.1136/bmj.e4178](https://doi.org/10.1136/bmj.e4178) PMID: [22730543](https://pubmed.ncbi.nlm.nih.gov/22730543/)
37. White RG, Hakim AJ, Salganik MJ, Spiller MW, Johnston LG, Kerr L, et al. Strengthening the Reporting of Observational Studies in Epidemiology for respondent-driven sampling studies: "STROBE-RDS" statement. *J Clin Epidemiol*. 2015; 68: 1463–1471. doi: [10.1016/j.jclinepi.2015.04.002](https://doi.org/10.1016/j.jclinepi.2015.04.002) PMID: [26112433](https://pubmed.ncbi.nlm.nih.gov/26112433/)
38. Nicholls SG, Quach P, von Elm E, Guttman A, Moher D, Petersen I, et al. The REporting of Studies Conducted Using Observational Routinely-Collected Health Data (RECORD) Statement: Methods for Arriving at Consensus and Developing Reporting Guidelines. *PLoS One*. 2015; 10: e0125620. doi: [10.1371/journal.pone.0125620](https://doi.org/10.1371/journal.pone.0125620) PMID: [25965407](https://pubmed.ncbi.nlm.nih.gov/25965407/)
39. Peters JP, Hooft L, Grolman W, Stegeman I. Assessment of the quality of reporting of randomised controlled trials in otorhinolaryngologic literature—adherence to the CONSORT statement. *PLoS One*. 2015; 10: e0122328. doi: [10.1371/journal.pone.0122328](https://doi.org/10.1371/journal.pone.0122328) PMID: [25793517](https://pubmed.ncbi.nlm.nih.gov/25793517/)
40. Peters JP, Hooft L, Grolman W, Stegeman I. Reporting Quality of Systematic Reviews and Meta-Analyses of Otorhinolaryngologic Articles Based on the PRISMA Statement. *PLoS One*. 2015; 10: e0136540. doi: [10.1371/journal.pone.0136540](https://doi.org/10.1371/journal.pone.0136540) PMID: [26317406](https://pubmed.ncbi.nlm.nih.gov/26317406/)