This is a post-print of: Netten, A.P., Dekker, F.W., Rieffe, C., Soede, W., Briaire, J.J., & Frijns, J.H.M. (2017). Missing Data in the Field of Otorhinolaryngology and Head & Neck Surgery: Need for Improvement. *Ear and Hearing, 38*, 1-6, which was published at: http://dx.doi.org/ 10.1097/AUD.0000000000000346.

Missing Data in the Field of Otorhinolaryngology and Head & Neck Surgery: Need for Improvement.

Anouk P. Netten,[1] Friedo W. Dekker,[2] Carolien Rieffe,[3,4] Wim Soede,[1] Jeroen J. Briaire,[1] and Johan H.M. Frijns[1,5]

[1]Department of Otorhinolaryngology and Head & Neck Surgery, Leiden University Medical Center, The Netherlands
[2]Department of Epidemiology, Leiden University Medical Center, The Netherlands
[3]Department of Developmental Psychology, Leiden University, The Netherlands
[4]Dutch Foundation for the Deaf and Hard of Hearing Child, Amsterdam, The Netherlands
[5]Leiden Institute for Brain and Cognition, The Netherlands

Corresponding author: A.P. Netten, MD., Department of Otorhinolaryngology and Head & Neck Surgery, Leiden University Medical Center, PO Box 9600, 2300 RC Leiden, The Netherlands, tell: +31 715262440, Fax: +31 715248201, e-mail: a.p.netten@lumc.nl

Abbreviations: MCAR – Missing Completely At Random, MAR – Missing At Random, MNAR – Missing Not At Random, MI – Multiple Imputations, DHH – deaf or hard of hearing

Keywords: Missing data, multiple imputations, review, otorhinolaryngology, head & neck surgery

Source of Funding: This research was financially supported by Stichting het Heinsius-Houbolt Fonds.

Conflict of Interest: None declared.

ABSTRACT

Objective Clinical studies are often facing missing data. Data can be missing for various reasons, e.g., patients moved, certain measurements are only administered in high-risk groups, patients are unable to attend clinic because of their health status. There are various ways to handle these missing data (e.g., complete cases analyses, mean substitution). Each of these techniques potentially influences both the analyses and the results of a study. The first aim of this structured review was to analyze how often researchers in the field of otorhinolaryngology / head & neck surgery report missing data. The second aim was to systematically describe how researchers handle missing data in their analyses. The third aim was to provide a solution on how to deal with missing data by means of the multiple imputation technique. With this review we aim to contribute to a higher quality of reporting in otorhinolaryngology research.

Design Clinical studies among the 398 most recently published research articles in three major journals in the field of otorhinolaryngology / head & neck surgery were analyzed based on how researchers reported and handled missing data.

Results Of the 316 clinical studies, 85 studies reported some form of missing data. Of those 85, only a small number (12 studies, 3.8%) actively handled the missingness in their data. The majority of researchers exclude incomplete cases, which results in biased outcomes and a drop in statistical power.

Conclusions Within otorhinolaryngology research, missing data are largely ignored and underreported, and consequently, handled inadequately. This has major impact on the results and conclusions drawn from this research. Based on the outcomes of this review, we provide solutions on how to deal with missing data. To illustrate, we clarify the use of multiple imputation techniques, which recently became widely available in standard statistical programs.

INTRODUCTION

"When dealing with real data, the practicing statistician should explicitly consider the process that causes missing data far more often than he does."

Rubin (p.589, 26)(Rubin 1976)

Missing data are almost inevitable when conducting research using patient information (Rubin 1976; Schafer et al. 2002; Wood et al. 2004; Van Buuren 2012). For numerous reasons, databases are incomplete and researchers have to decide how to deal with this issue. Most often in medical research, this problem is overlooked and missing data are underreported (Wood et al. 2004; Sterne et al. 2009). However, it is important for researchers to realize that standard analyzing techniques assume complete cases and consequently remove incomplete cases from the analyses. Ignoring missing data through complete case analyses introduces bias and a drop in statistical power as it insufficiently uses the available data (Schafer and Graham 2002). The first aim of this structured review was to evaluate the (under)reporting of missing data in the otorhinolaryngology research field. The second aim was to analyze how researchers deal with missing data and highlight the consequences this potentially has. The third aim was to provide solutions on how to deal with missing data using modern techniques that are widely available nowadays.

The quality of medical research reports is of increasing interest to assure valid outcomes and generalizability. A growing number of journals requests authors to complete checklists such as the Consolidated Standards of Reporting Trials (CONSORT) for randomized controlled trials and the Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) for observational studies (Moher et al. 2001; Vandenbroucke et al. 2007). These checklists provide a guideline for the concise report of medical research. Among other things, checklists like STROBE emphasize the importance of reporting missing data in all variables of interest and strongly recommend to give reasons for missing data where possible.

Types of missing data

What to do when confronted with missing data largely depends on under what assumption the data are incomplete. In other words, what are the characteristics of the missing data and do we know the reason why a value is missing? Epidemiologists assume three types of missing data: i.e., Missing Completely At Random (MCAR), Missing At Random (MAR), and Missing Not At Random (MNAR) (Van Buuren 2012).

Missing Completely At Random (MCAR)

The reason for missingness is completely independent of the (missing) true value, and from any other variables that are or are not included in the dataset. An example of MCAR is a questionnaire that was lost in the mail, or a broken freezer that contained frozen patient specimens. In the case of MCAR, the observed values are a random selection of the sample and thus, are representative for that population.

Missing At Random (MAR)

In the MAR condition, the reason for missingness is related to other factors that are measured within the dataset. This term can be confusing as it suggests that there is no relation between the missing values and other factors, albeit there is. For instance, in a dataset, spoken language scores are more often missing from Deaf and Hard of Hearing (DHH) children that prefer to use sign-supported language as their mode of communication. Likely, the missing scores for children that prefer to use sign language are lower than for children who prefer spoken language. In the MAR assumption, factors that are related to the missing values (e.g. communication mode) can help to reconstruct the actual level of spoken language scores.

Missing Not At Random (MNAR)

A problem arises when the reason for missing data is related to the true value, or to other unknown factors. Yet, these variables are all unknown. This is the case in data that is MNAR; data it is missing only because of its value. To illustrate, MNAR might happen when asking cancer participants about their quality of life during their out-clinic appointment. The answers might be missing because the patient was too sick to attend to clinic. Another example is patients suffering from depression that are too depressed to complete a questionnaire about their mental wellbeing. Here, the true value of the outcome measure is the reason why the specific value is missing. The difference with both MCAR and MAR is that in the MNAR condition we do not know the reason, nor can we speculate what the true value would have been, because essential information is not available.

Hypothesizing the reason for missingness and under what assumption data are missing is helpful in the process of deciding how to handle this issue. Although it is tempting to assume that data fall under either one of these three assumptions, often the pattern of missing data is a combination of more than one of the assumptions. The missing data of some patients are MCAR, others are MAR, and others are even MNAR. Reporting missing data is essential to assure valid and replicable results. Unfortunately, this is still quite unpopular in medical research. To illustrate this statement, this structured review identified how researchers in the field of otorhinolaryngology reported and handled missing data. Additionally, we explain the multiple imputation technique to adequately handle missing data.

METHODS

A literature review of the most recent articles published in three major Otorhinolaryngology / Head & Neck surgery journals was performed to identify how researchers reported and handled missing data. All articles published between September 1^{st} 2014 and August 31^{st} 2015 in the journals *Ear and Hearing* (159 articles), *Rhinology* (76 articles), and *Head & Neck* (679 articles) were identified. Because the third journal published over 600 articles during that period, we decided to analyze a sub selection and included all articles published between the 1^{st} of May and the 31^{st} of August 2015 (163 articles). A total of 398 articles were identified. Articles were excluded if they did not describe clinical research as is the case in reviews, letters and case-reports. A total of 316 articles describing clinical research were selected for further analysis. For details on exclusion, see figure 1.

All included articles were systematically checked on terms like ‘missing’, ‘unknown’, ‘remove’, ‘exclude’, ‘complete’, ‘absent’, ‘lost’, and ‘imputation’ by the first author. The methods and results section of each article were analyzed based on two questions: i.) did the authors report missing data and if so, ii.) how did they handle the missingness in their analysis? Figures and tables were checked if numbers added up, and whether or not they reported characteristics to be ‘unknown’ or ‘missing’. Statistical analyses were checked as to whether the degrees of freedom were consistent, if imputations were mentioned or applied, and if other likelihood-based methods were used that are able to handle missing data without excluding incomplete cases, such as linear mixed models (Twisk et al. 2013). A second researcher additionally checked 30 randomly selected articles out of the 316 articles and confirmed the findings of the first one.

RESULTS

Of the 316 eligible articles, roughly one-fourth (85 articles) reported some kind of missing data, either in the text, or it was indirectly derived from tables, figures and/or analyses. In 73 of those 85 articles, complete case analyses or pairwise deletions were used. The remaining 12 articles (9 in *Ear and Hearing*, 2 in *Head & Neck*, and 1 in *Rhinology*) actively took action upon their missing data. In eight of these 12 articles, the mean substitution method was used. In two articles complete and incomplete cases were compared on several variables to illustrate that data were MCAR. In one case, a linear mixed model was used and in the remaining case, multiple imputations were performed to handle missing data, see Table 1 and Figure 2 for an overview.

Fifty of the clinical studies in this review had a relatively small sample size (i.e., less than 25 participants). None of these small studies reported missing data. Most of these studies were experiments in the area of cochlear implantation with few participants. Because of the small sample size, these type of studies usually do not encounter missing data related issues and often only perform descriptive statistics. Therefore, we decided to perform a sensitivity analyses and excluded the 50 small studies. Excluding these studies only raised the percentage of studies that reported some kind of missing data (n=85) to nearly one-third of the total sample.

DISCUSSION

This structured review examined how often researchers in the field of Otorhinolaryngology / Head & Neck surgery report missing data in their research. If missing data were reported, the second aim was to analyze how researchers solve missing data-related issues. The outcomes of this review underline the importance of this study. Despite the introduction of checklists (such as the STROBE) to increase the quality of reporting, the majority of researchers do not report missing data, nor step up to act adequately when confronted with missing data. This might be due to the fact that the use of such checklists is not mandatory in many journals, and their use is therefore relatively unknown. We therefore assume that this underreporting of missing data is most likely the result of unfamiliarity with the consequences of missing data assumptions rather than an unwillingness to deal with this issue (Newgard et al. 2015). To increase awareness, we will attempt to explain how several commonly used methods to handle missing data can influence results. Second, we will provide a solution on how to adequately handle missing data using modern, well-established techniques.

Complete case analyses

As can be seen in Figure 2, the majority of researchers who reported missing data did not handle this issue. Not deciding how to handle missing data results in *complete case analyses* (also called *listwise deletion*), i.e. the incomplete cases are removed from the analyses. In programs like *SPSS* (IBM 2013), this is automatically done. When performing a t-test for example, the program removes incomplete cases when conducting the test and reports the amount of cases with incomplete data. It is important to note that this method is only accurate when the cases with complete data are a random selection of the population. In other words, the incomplete cases may not differ systematically from the complete cases. Complete case analyses can thus only be used if missing data are MCAR. Strikingly, the MCAR assumption is very difficult to prove. The researcher has to be sure that there is no common reason why this specific selection of data is missing. Yet, in practice, data are most frequently MAR. Hence, the complete cases analyses technique will rarely produce the most accurate outcomes. To add, removing incomplete cases from the analyses will always result in loss of power and accuracy.

Comparison of complete and incomplete cases

In this review, four research groups attempted to prove the MCAR statement by comparing complete and incomplete cases on several characteristics that could potentially influence the missing variable in order to prove no differences between the two groups (Aarhus et al. 2015; Bulut et al. 2015; Huang et al. 2015; Stam et al. 2015). Yet, it is often impossible to test all possible related variables. As a result, assuming MCAR and removing incomplete cases from the analyses produces biased results and broadens the confidence intervals as a result of lower statistical power if data are MAR or MNAR. Unfortunately, complete case analyses are often used without hypothesizing the reason for missingness. The same goes for *pairwise deletion*. In this technique the complete cases are identified and analyzed separately. This method was identified once in this review (Kumar et al. 2015). Pairwise deletion additionally blurs the outcomes as the number of participants differs per analysis. To illustrate, if correlations are measured but the number of participants per analysis differs, this may yield biased estimates.

<u>Mean substitution</u>

The disadvantages of complete case analyses suggest it might be more convenient to reconstruct the missing data instead of throwing incomplete cases out. Standard techniques can then be used on the reconstructed dataset which solves the power issue. In this review, eight researchers chose to use the *mean substitution* technique, which calculates the mean of the complete cases and imputes (‘fills in’) this mean in all missing fields of that variable (Mackersie et al. 2015). This tool was most often used when data in questionnaires was missing (Aarhus et al. 2015; Barry et al. 2015; Bulut et al. 2015; Hesser et al. 2015; Hornsby et al. 2015; Huang et al. 2015; Kumar et al. 2015). Manuals of validated questionnaires often state that a scale may be measured if *n* % of the items to calculate that scale is missing. For example, if a scale consists of five questions but only four are answered, the mean of these four questions is imputed in the fifth question because the questionnaire assumes a high correlation between the five items within a certain scale (i.e., the internal consistency of the scale). In one other article, zip code-specific socio-economic variables of participants with missing zip codes were replaced by the state average (Schaefer et al. 2015).

However, this method has some disadvantages. Suppose there is a correlation between the outcome and the substituted value. As a result of mean substitution, the strength of this relation alters. To add, it also artificially narrows the confidence interval of the imputed variable because a higher percentage of data lies closer to the mean.

Missing data in longitudinal research

Last observation carried forward (LOCF, also known as baseline observation carried forward) is a method that can be used in longitudinal data. This method was not used in any of the articles in this review but is worthwhile to discuss as longitudinal data is increasingly collected, also in Otorhinolaryngology / Head & Neck surgery research. This method copies the last known observation in a row of observations and imputes it in the missing fields of that case. An advantage of this method is that it is case specific because it acknowledges the fact that every case is different and unique. However, the development over time is seriously biased by this method and special analyzing techniques should follow after LOCF. Especially if one is interested in development over time or a treatment effect, these results are biased by LOCF. An additional problem arises when the baseline measure is missing as these cases will still be excluded in complete cases analyses. In addition, cases with missing data in (one of the) confounders will be excluded when such confounders are added to the analyses.

<u>Likelihood-based approaches</u>

De Kegel et al. use linear mixed models in their longitudinal study to account for missing values (De Kegel et al. 2015). *Likelihood-based methods* such as linear mixed models create a model based on the observed data of both complete and incomplete cases. It calculates the maximum likelihood estimate; the value of a parameter that is most likely to have resulted in the observed data. Both the likelihood estimate of the complete and incomplete cases are calculated and jointly maximized. This method does not impute values and is therefore relatively easy to use. It is a reliable method when confronted with missing data in studies with a longitudinal design. However, likelihood-based approaches are limited to linear models. Another potential pitfall when using this approach is that all the factors that are entered into the model besides the dependent variable should not have missing data. Otherwise these cases will still be excluded from the analyses.

A state of the art solution: Multiple imputation

All the above described methods to handle missing data have their limitations. We will therefore now highlight the abilities of multiple imputations (MI), a well-established technique that has none of the limitations described above. MI is increasingly used since popular statistical programs started to include its possibility in their interface. This technique was used in only one article in this review (Sereda et al. 2015).

Imputation means nothing more than “filling in the data”. Multiple imputations indicate that the imputations were done more than once. To illustrate the mechanism behind MI, we will return to the previously mentioned fictive dataset containing language scores of DHH children in which language scores of some children were missing. In this database, we observed that children who preferred to use sign-supported language often had lower spoken language scores than children that preferred to use spoken language to communicate. If we now decide to use the preferred mode of communication of the child to predict their language scores, this would produce a more accurate result than when imputing the mean language score of the whole sample. In the same line of thinking, we also know from the complete data that children attending mainstream schools show higher language scores than those attending special education. We can therefore decide to include the type of school that the child attended into the prediction model. Additionally, the age of the child is also positively related to its language abilities, and so on. One will notice that the more variables we will put into this so-called prediction model, the more accurate the prediction of the possible language score will turn out. The MI method uses the complete data to compute a prediction model of the variable that has missing data. It then uses characteristics of the missing cases to predict the missing values in the data.

Obviously, the imputation model only calculates an estimation of the unknown value. The true value lies within a certain range that was estimated by the calculated prediction model. We therefore want to insert a certain amount of uncertainty (or variance) for this value. To achieve this, instead of doing this imputation only once, we have the model predict a language score *n* times. This results in one large database containing *n* datasets in which the complete cases remain the same, but the missing values differ within the range that was estimated by the prediction model. All these complete datasets can then be analyzed simultaneously using standard techniques (e.g., t-tests, ANOVA's) which generates *n* outcomes. These outcomes are automatically pooled into one outcome with one *p*-value; the final result of the analysis. Pooling these *n* datasets will give a mean of the *n* imputed values together with its standard error; the uncertainty of our estimation. MI is a robust method that produces valid and unbiased outcomes (Van Buuren 2012; de Goeij et al. 2013). However, its use requires some training and should always be guided by an experienced user of the MI method, especially since there is still debate about what to do when data are MNAR. Sterne and colleagues provided clear guidelines on how to report the use of MI in scientific writing to improve reproducibility and increase transparency (Sterne et al. 2009).

Without any doubt, it would be best to prevent the appearance of missing data. Although almost inevitable, this can partly be achieved by thoroughly overthinking all steps of data-collection during the design of a new study. We would therefore strongly advise researchers to contact an epidemiologist or statistician prior to the start of a new study. Studies entirely devoted to the prevention of missing data provide useful tips such as the use of user-friendly case-report forms, the conduction of a pilot-study, and teaching of research assistants prior to the start of the study (Wisniewski et al. 2006; Scharfstein et al. 2012; Kang

2013). Even if data collection has already finished, contacting an epidemiologist or statistician can be very helpful to discuss the appearance of missing data and possible methods to handle missing data related issues, in order to assure valid outcomes.

CONCLUSION

With this article we want to draw attention to the importance of reporting missing data, and urge researchers to hypothesize about why data are missing. Defining why data is missing is essential in the process of selecting the most reliable technique to solve the missing data issue and prevent researchers from drawing invalid conclusion. We strongly suggest researchers to use available guidelines for reporting research (e.g., STROBE and CONSORT). To add, we highly recommend editorial boards of scientific journals to introduce the use of such checklists to increase their familiarity and ensure high reporting standards. To improve the quality of reporting, we would also like to encourage reviewers to pay attention to missing data and its possible consequences when reviewing articles for publication. As can be seen from this review, in the Otorhinolaryngology / Head & Neck surgery research field most often missing data are not reported and they are rarely handled properly. With this review, we hope to motivate researchers to think about missing data and to use methods such as multiple imputation to maximize the use of their data in order to draw more valid conclusions in future research.

ACKNOWLEDGEMENTS

The authors would like to thank Mrs. Ewa Banat for reviewing a selection of articles. This research was financially supported by Stichting het Heinsius-Houbolt Fonds.

A.P.N. and F.W.D. defined the outlines of this review and wrote the main paper. A.P.N. reviewed all articles and performed the analysis. All authors discussed the results and implications and commented on the manuscript in all stages.

REFERENCES

Aarhus, L., Tambs, K., Kvestad, E., et al. (2015). Childhood Otitis Media: A Cohort Study With 30-Year Follow-Up of Hearing (The HUNT Study). *Ear Hear, 36*, 302-308.

Barry, J. G., Tomlin, D., Moore, D. R., et al. (2015). Use of Questionnaire-Based Measures in the Assessment of Listening Difficulties in School-Aged Children. *Ear Hear*.

Bulut, O. C., Wallner, F., Plinkert, P. K., et al. (2015). Quality of life after septorhinoplasty measured with the Functional Rhinoplasty Outcome Inventory 17 (FROI-17). *Rhinology, 53*, 54-58.

de Goeij, M. C., van Diepen, M., Jager, K. J., et al. (2013). Multiple imputation: dealing with missing data. *Nephrol Dial Transplant, 28*, 2415-2420.

De Kegel, A., Maes, L., Van Waelvelde, H., et al. (2015). Examining the impact of cochlear implantation on the early gross motor development of children with a hearing loss. *Ear Hear, 36*, e113-121.

Hesser, H., Bankestad, E., Andersson, G. (2015). Acceptance of Tinnitus As an Independent Correlate of Tinnitus Severity. *Ear Hear, 36*, e176-182.

Hornsby, B. W., Kipp, A. M. (2015). Subjective Ratings of Fatigue and Vigor in Adults with Hearing Loss Are Driven by Perceived Hearing Difficulties Not Degree of Hearing Loss. *Ear Hear*.

Huang, T. L., Chien, C. Y., Tsai, W. L., et al. (2015). Long-term late toxicities and quality of life for survivors of nasopharyngeal carcinoma treated with intensity-modulated radiotherapy versus non-intensity-modulated radiotherapy. *Head Neck*.

IBM SPSS Statistics for Windows Version 23.0. Armonk, NY: IBM Corp.; 2013.

Kang, H. (2013). The prevention and handling of the missing data. *Korean Journal of Anesthesiology, 64*, 402-406.

Kumar, R., Warner-Czyz, A., Silver, C. H., et al. (2015). American parent perspectives on quality of life in pediatric cochlear implant recipients. *Ear Hear, 36*, 269-278.

Mackersie, C. L., MacPhee, I. X., Heldt, E. W. (2015). Effects of hearing loss on heart rate variability and skin conductance measured during sentence recognition in noise. *Ear Hear, 36*, 145-154.

Moher, D., Schulz, K. F., Altman, D. G. (2001). The CONSORT statement: revised recommendations for improving the quality of reports of parallel-group randomized trials. *J Am Podiatr Med Assoc, 91*, 437-442.

Newgard, C. D., Lewis, R. J. (2015). Missing data: How to best account for what is not known. *JAMA, 314*, 940-941.

Rubin, D. B. (1976). Inference and Missing Data. *Biometrika, 63*, 581-590.

Schaefer, E. W., Wilson, M. Z., Goldenberg, D., et al. (2015). Effect of marriage on outcomes for elderly patients with head and neck cancer. *Head Neck, 37*, 735-742.

Schafer, J. L., Graham, J. W. (2002). Missing data: our view of the state of the art. *Psychol Methods, 7*, 147-177.

Scharfstein, D. O., Hogan, J., Herman, A. (2012). On the prevention and analysis of missing data in randomized clinical trials: the state of the art. *J Bone Joint Surg Am, 94 Suppl 1*, 80-84.

Sereda, M., Hoare, D. J., Nicholson, R., et al. (2015). Consensus on Hearing Aid Candidature and Fitting for Mild Hearing Loss, With and Without Tinnitus: Delphi Review. *Ear Hear, 36*, 417-429.

Stam, M., Smits, C., Twisk, J. W., et al. (2015). Deterioration of Speech Recognition Ability Over a Period of 5 Years in Adults Ages 18 to 70 Years: Results of the Dutch Online Speech-in-Noise Test. *Ear Hear, 36*, e129-137.

Sterne, J. A., White, I. R., Carlin, J. B., et al. (2009). Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *Bmj, 338*, b2393.

Twisk, J., de Boer, M., de Vente, W., et al. (2013). Multiple imputation of missing values was not necessary before performing a longitudinal mixed-model analysis. *J Clin Epidemiol, 66*, 1022-1028.

Van Buuren, S. (2012). *Flexible Imputation of Missing Data*. Boca Raton: CRC Press.

Vandenbroucke, J. P., von Elm, E., Altman, D. G., et al. (2007). Strengthening the Reporting of Observational Studies in Epidemiology (STROBE): Explanation and Elaboration. *PLoS Med, 4*, e297.

Wisniewski, S. R., Leon, A. C., Otto, M. W., et al. (2006). Prevention of missing data in clinical research studies. *Biol Psychiatry, 59*, 997-1000.

Wood, A. M., White, I. R., Thompson, S. G. (2004). Are missing outcome data adequately handled? A review of published randomized controlled trials in major medical journals. *Clinical Trials, 1*, 368-376.

Figure 1 Flow chart of **structured** review

Figure 2 Proportion of papers that reported missing data

Table 1. Characteristics of selected studies that actively handled missing data

Author	**Type of study**	**Imputation method**	**Detail**	**Journal**
(Aarhus et al. 2015)	Longitudinal cohort	Mean substitution	Comparison of responders vs. non responders on many characteristics, report loss to follow-up and discuss the probability of selection bias	*Ear and Hearing*
(Barry et al. 2015)	Cross-sectional case-control	Mean substitution	Within different questionnaires, missing data were replaced by mean data	*Ear and Hearing*
(Bulut et al. 2015)	Cross-sectional cohort	Mean substitution	Comparison of responders vs. non responders on two characteristics, mean substitution in one questionnaire	*Rhinology*
(De Kegel et al. 2015)	Longitudinal case-control	Likelihood-based approach	Do not report missing data, no. of participants increases with follow-up time	*Ear and Hearing*
(Hesser et al. 2015)	Cross-sectional cohort	Mean substitution	Within different questionnaires, missing data were replaced by mean data if < 20% of items per scale was missing, followed by complete case analyses	*Ear and Hearing*
(Hornsby and Kipp 2015)	Cross-sectional cohort	Mean substitution	Missing data were replaced by mean data in one questionnaire, followed by complete case analyses	*Ear and Hearing*
(Huang et al. 2015)	Cross-sectional cohort	Mean substitution	Comparison of responders vs. non responders on several characteristics to account for selection bias, in one questionnaire, missing data were replaced by mean data if < 50% of items per scale was missing	*Head & Neck*
(Kumar et al. 2015)	Cross-sectional cohort	Mean substitution	Within one questionnaires, missing data were replaced by mean data, followed by pairwise deletions	*Ear and Hearing*
(Mackersie et al. 2015)	Cross-sectional case-control	Mean substitution	In ECG: artifacts were removed and missing intervals were interpolated from the adjacent interbeat interval values (<1%)	*Ear and Hearing*
(Schaefer et al. 2015)	Cross-sectional cohort	Mean substitution	For missing zip codes, the state average was imputed. Bootstrapping was used to obtain confidence intervals of the built model	*Head & Neck*
(Sereda et al. 2015)	Longitudinal cohort	Multiple Imputation	No information	*Ear and Hearing*
(Stam et al. 2015)	Longitudinal case-control	None	Comparison of responders vs. non responders, report selection bias because of loss to follow-up	*Ear and Hearing*