Voice Rehabilitation by Voice Prostheses After Total Laryngectomy: A Systematic Review and Network Meta-Analysis for 11,918 Patients

19 authors, including:

Gehad Mohamed Tawfik
Ain Shams University
60 PUBLICATIONS **467** CITATIONS

SEE PROFILE

Omar Elsayed
Houston Methodist Hospital
25 PUBLICATIONS **568** CITATIONS

SEE PROFILE

Ahmad Helmy Zayan
Menoufia University
25 PUBLICATIONS **143** CITATIONS

SEE PROFILE

Sherief Ghozy
Mayo Clinic - Rochester
229 PUBLICATIONS **2,521** CITATIONS

SEE PROFILE

Some of the authors of this publication are also working on these related projects:

Safety and efficacy of Nitazoxanide on diarrhea. A systematic review and meta-analysis of randomized controlled trials View project

Dengue neurological complications: a prospective cohort study View project

Review Article

Voice Rehabilitation by Voice Prostheses After Total Laryngectomy: A Systematic Review and Network Meta-Analysis for 11,918 Patients

Gehad Mohamed Tawfik,[a,b] (iD) Omar Mohamed Makram,[b,c] (iD) Ahmad Helmy Zayan,[b,d]
Sherief Ghozy,[b,e] (iD) Peter Samuel Eid,[a,b] (iD) Mona Hanafy Mahmoud,[a,b] Abdelaziz Abdelaal,[b,f]
Seif Mahmoud Abdelghany,[b,g] Ahmed M. Sayed,[b,h] To Kim Sang,[b,i] Mahmoud Kassem,[b,j]
Quoc Le Minh Ho,[b,k] Heba Hussien Eltanany,[b,l] Amira Farghaly Ali,[b,m] Osama Gamal Sweiny,[b,n]
Khaled Essam Elsherbiny,[b,o] Amr G. Shafik,[p] Kenji Hirayama,[q] (iD) and Nguyen Tien Huy[r] (iD)

Purpose: Our aim was to assess the different voice prostheses (VPs) to identify the most efficient, safest, patient-tailored, longest lifetime, and inexpensive VPs and assess the different factors affecting their quality.
Method: In September 2017, 15 databases were searched to include all randomized controlled trials. A new search was done in May 2019 to include all other study design articles, which include all the new-era VPs subtypes. Network meta-analysis (NMA) was applied to all 27 outcomes, besides NMA overall and partial order setting was done by using Hasse scatter plots. p values were used in NMA, where the best VPs are approaching one and the least approaches zero. Meta-analysis was done for the rest of the outcomes.
Results: Two hundred one articles were eligible for inclusion in our study (N = 11,918). Provox-2 was significantly the most efficient and safest device concerning the most patient preference (odds ratio [OR] = 33.88 [0.65, 1762.24]; p = .92), the least dislodgement (risk ratio [RR] = 0.27 [0.13, 0.57]; p = .79), the least airflow resistance (RR = 0.42 [0.08, 2.11]; p =

.84), the least granulation formation (RR = 0.73 [0.02, 26.32]; p = .60), and the least VPs' inaccurate size (RR = 0.77 (0.23, 2.61]; p = .66). Heat and moisture exchanger addition showed a significant increase in maximum phonation time and breathing experience, with p values (1 and .59), respectively. While heat and moisture exchanger addition showed a significant decline in stoma cleaning frequency, coughing frequency, forced expectoration, sputum production, sleeping problems, and loosening of adhesive, with p values (.99, .72, .69, .96, 1, and 0.96), respectively, Groningen low resistance and Nijdam were considered the worst devices with both overall mean p value of .44.
Conclusions: Provox-2 is considered the best choice as being the most preferable for patients, with the least airflow resistance, dislodgment, granulation formation, and prosthesis inaccurate size. Groningen low resistance and Nijdam were considered the worst devices according to our analysis.
Supplemental Material: https://doi.org/10.23641/asha.14802903

[a]Faculty of Medicine, Ain Shams University, Cairo, Egypt
[b]Online Research Club, Nagasaki, Japan
[c]Faculty of Medicine, October 6 University, Giza, Egypt
[d]Department of Otolaryngology, Menoufia University, Egypt
[e]Neurosurgery Department, El Sheikh Zayed Specialized Hospital, Giza, Egypt
[f]Faculty of Medicine, Tanta University, Gharbia, Egypt
[g]Faculty of Medicine, Fayoum University, Egypt
[h]Faculty of Pharmacy, Al-Azhar University, Cairo, Egypt
[i]Ho Chi Minh City Oncology Hospital, Vietnam
[j]The Ohio State University Wexner Medical Center, Columbus
[k]University Medical Center of Ho Chi Minh City, Vietnam
[l]Faculty of Medicine, Al-Azhar University, Cairo, Egypt
[m]Faculty of Medicine, Minia University, Egypt
[n]Faculty of Medicine, South Valley University, Qena, Egypt
[o]Faculty of Medicine, Port-Said University, Egypt
[p]Department of Otorhinolaryngology, Faculty of Medicine, Ain Shams University, Cairo, Egypt

[q]Department of Immunogenetics, Institute of Tropical Medicine (NEKKEN), Leading Graduate School Program and Graduate School of Biomedical Sciences, Nagasaki University, Sakamoto, Japan
[r]School of Tropical Medicine and Global Health, Nagasaki University, Japan

Correspondence to Gehad Mohamed Tawfik: dr.gehadmohamed@hotmail.com or Nguyen Tien Huy: tienhuy@nagasaki-u.ac.jp

Gehad Mohamed Tawfik and Omar Mohamed Makram contributed equally to this work.

Editor-in-Chief: Bharath Chandrasekaran
Editor: Jack J. Jiang

Laryngeal cancer is the most common malignancy of the head and neck, ranking the 19th globally with an estimated 13,150 new cases in the United States in 2018, as reported by the American Cancer Society (Global Burden of Disease Cancer Collaboration, 2015; Hernandez et al., 2010; Jemal et al., 2007). Prognosis of laryngectomy surgery remained relatively favorable, with a 5-year survival rate of 60.9%, raising it as a treatment of choice for advanced stages of laryngeal carcinoma (Stankovic et al., 2017). Nevertheless, laryngectomy is usually accompanied by voice loss with a subsequent negative impact on patients' quality of life (Brown et al., 2003; Krishnan & Maclean, 2013). Consequently, an early rehabilitative measure represent a need to overcome these adverse postlaryngectomy outcomes. Patients undergoing laryngectomy featured psychological problems because of the inability to communicate with others (Lorenz, 2017). Formation of a tracheoesophageal fistula with a voice prosthesis (VPs) insertion is now considered as the gold standard approach and a routine procedure to restore voice after total laryngectomy (Brown et al., 2003; Koch, 2002; Mahieu et al., 1986). Multiple factors determine the choice of VPs, including availability, cost affordability, patient preference, and device lifetime (K. L. Hancock et al., 2013). In addition, the device lifetime remains the most important factor, as the shorter the device lifetime, the more frequent replacement is, hence much more expenses (Ameye et al., 2005; K. L. Hancock et al., 2013). Moreover, complications can result in prosthesis replacement. These include silicone degradation, candida, incorrect length or even accidental extrusion, increased airflow resistance during speech, and less frequently excessive granulation tissue formation (Brown et al., 2003; Koch, 2002; Slavícek et al., 2000). Each type of VPs has its own benefits and limitations, which have been assessed in several studies (K. L. Hancock et al., 2013; Harms et al., 2011; Van Den Hoogen et al., 1996). However, there is a lack of a systematic review and network meta-analysis (NMA) comparing all aspects of these prostheses and ordering them accordingly. In our study, we aimed to analyze the different aspects concerning VPs; rank them according to each outcome; and determine the most efficient, safest, patient-tailored, longest lifetime, least complications, and inexpensive VP for speech rehabilitation. In addition, we aimed to determine the various factors affecting these VPs, for example, heat and moisture exchanger (HME), antifungal drugs for colonization, cost, and postoperative follow-up.

Method

Study Protocol Registration

This systematic review and NMA was conducted following the widely accepted Preferred Reporting Items for Systematic Review and Meta-Analysis Statement 2009 (Moher et al., 2009). The detailed steps of methods are shown in the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) checklist (see Supplemental Table S1) and other summarized articles (Tawfik et al., 2019). We registered our protocol at PROSPERO with Registration Number CRD42017080110.

Search Strategy

On September 14, 2017, we searched for relevant articles throughout 15 databases: PubMed, Google Scholar, Scopus, Web of Science, EMBASE, VHL, WHO GHL, Cochrane, Clinical trials.gov, mRCT, Science Direct, WHO, CINAHL, POPLINE, and SIGLE. We performed a manual search of possibly missed articles. There were no restrictions regarding age, sex, race, country, publication date, or language. We used a standard search strategy, which was later modified according to each unique database. Detailed search strategy for each database is placed in Supplemental Table S2. Then, a new search was done on May 11, 2019, in same databases to include all study design articles for inclusion of all the new-era VP subtypes, which were not included in randomized controlled trials (RCTs) trials, due to lack of RCTs in the new era of this topic. Thus, we included other observational studies with the existing RCTs to make the most out of the existing literature and do not miss any important device used in the market, which might help with physicians' decisions upon choosing the right device for the different patients.

Study Eligibility Criteria

We included all articles that assess VPs or factors affecting their outcomes. Our exclusion criteria were (a) single case articles; (b) studies with unavailable full text; (c) unreliable, incomplete, or unextractable, duplicated data; and (d) animal, or in vitro studies only. Three independent reviewers assessed the initially retrieved titles and abstracts for eligibility and further for full-text screening. When necessary, discrepancies were resolved by the opinion of a senior reviewer (fourth reviewer). The selection procedure of the study is summarized in the PRISMA flow diagram shown in Figure 1.

Data Extraction

An extraction sheet was developed by the senior reviewers. Three reviewers independently extracted reliable data from included articles. According to the primary outcome of the study, we compare devices—Provox-1, Provox-2, Provox ActiValve (AV), Provox non-indwelling device (NID), Provox Vega, Sound-Producing Voice Prosthesis (SPVP), Nijdam, Groningen Low Resistance (Rücker, Schwarzer, Krahn, & König), Groningen Ultra Low Resistance (ULR; Rücker et al.), BS low pressure (BS-LP)—together to assess five outcomes by NMA to know better the device in each outcome, which are patients' preferences, devices' lifetime, devices' replacement, maximum phonation time (MPT), leakage, follow-up survival rate, prosthesis
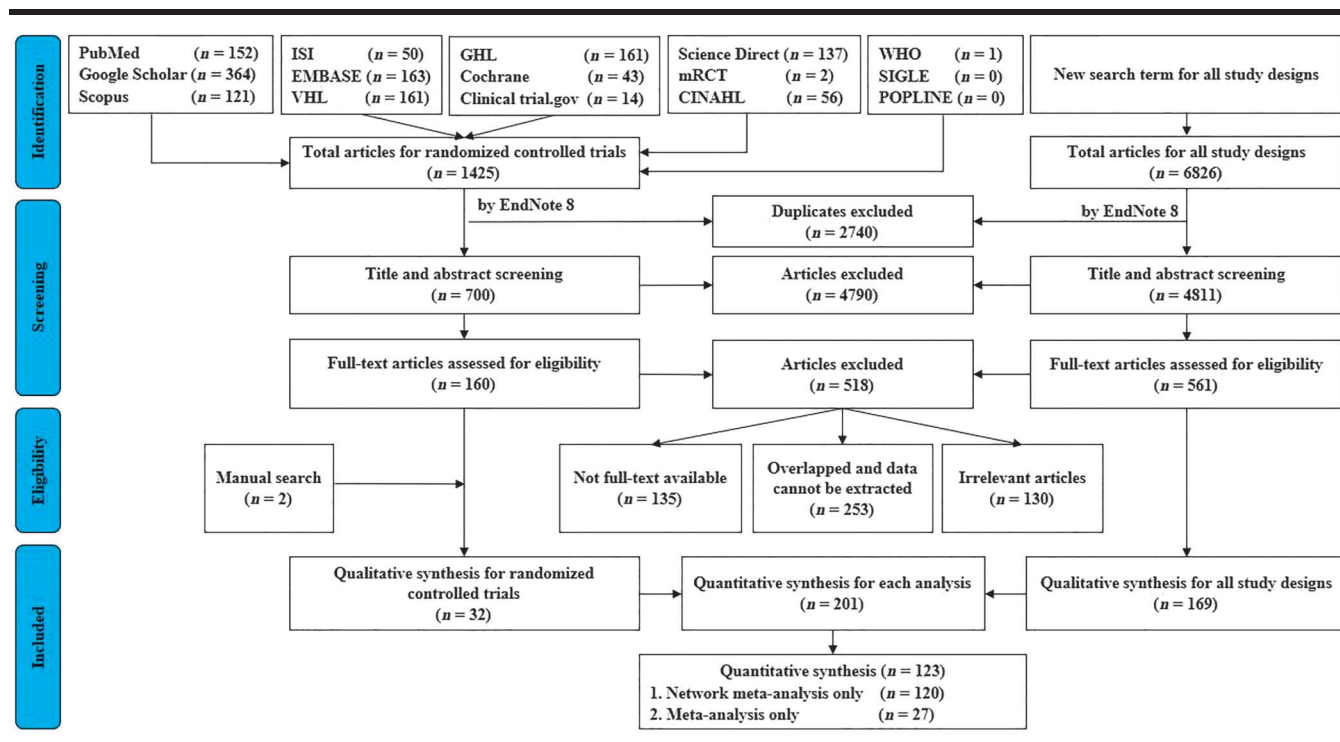
deterioration, prosthesis inaccurate size, granulation, fistula problems, dislodgement, stoma stenosis, speech intelligibility, voice loudness, fundamental frequency, increase phonatory effort, speech rate, and airflow resistance. The detailed definitions of each outcome are shown in Supplemental Table S3.

Quality Assessment

Three independent reviewers have assessed each article while doing the data extraction. We used the Cochrane Collaboration's tool for assessing the risk of bias of the 32 included RCT trials (J. P. Higgins et al., 2011). The assessment was categorized into "low risk," "high risk," or "unclear risk." The remaining 169 non-RCT articles were assessed by the different quality assessment tools provided by the National Institutes of Health for observational cohort, cross sectional studies, and case-series studies National Heart, Lung, and Blood Institute, n.d.). Regarding cross-sectional and cohort studies, each study was given a score out of 14 based on answering each question (Yes = 1, No = 0, NA = 0). A score of 10–14 indicated a good quality article, 5–9 for fair, and 1–4 for poor quality article. For the case-series studies, total evaluation score was 9, a score from 7 to 9 indicated a good quality article, whereas a score from 4 to 6 for fair and 1–3 for poor quality articles.

Statistical Analysis

All equations used in extraction prior to analysis and estimation of standard deviation from other variables is found in Supplemental Table S4, with their references as Hozo et al. (2005), Huy et al. (2013), Van Rijkom et al. (1998), and Wan et al. (2014). Whenever two included studies have the same sample size/data sets of patients, we merged their results and selected mainly the newest with newer data to overcome overinflation and overlap in our analysis results, as some included studies were overlapped data sets from the same patients (K. Hancock et al., 2012; Ward et al., 2011; Zuur et al., 2008, 2009).

NMA

Frequentist NMA was used to compare the investigated VPs' outcomes within a single analytical framework (J. Higgins et al., 2012). NMA was performed by R statistical software Version 3.4.3 with the help of package "netmeta" (Rücker et al., 2015). The NMA was reported in accordance with the modified PRISMA guidelines for NMAs (Hutton et al., 2015). A p value of 0 hypothesizes the worst efficacy where 1 indicates the best. Due to the presence of 27 different outcomes, overall NMA analysis was done to get the best and worst device regarding overall outcomes analyzed together using Hasse scatter plots (Rücker & Schwarzer, 2017). We provided NMA supplementary figures that included all the analyses performed between all the devices in the different outcomes, which consisted of three parts. The parts were (see Supplemental Figure S1) network plots that show direct and indirect comparisons. The size of the nodes represents the total sample size of patients using those VPs. The lines' thickness corresponds to the number of included

studies (see Supplemental Figure S2): League table showed the results of NMA comparing effect size of different types of devices, in which the top-left type was considered the most efficient and safest VPs in that outcome (see Supplemental Figure S3): forest plot, which arranges compared devices according to *p* values, effect sizes of each of them, and its 95% confidence interval.

Partial Order Setting of NMA

Additional NMA partial order setting was done by using overall Hasse scatter plots, partial order settings, and ranked the devices using *p* values to compare patient preference with other VP outcomes (Rücker & Schwarzer, 2017).

Meta-Analysis

Finally, we made pairwise meta-analyses of secondary outcomes—factors affecting VPs, using Comprehensive Meta-Analysis software (Version 3.9; Borenstein et al., 2005). Heterogeneity was assessed with Q statistics and I^2-test, considering it significant with I^2 value > 50% or $p < .10$. Publication bias was assessed with Egger's regression test (Begg & Mazumdar, 1994) and represented graphically by Begg's funnel plot, where Egger's regression $p < .05$ was considered significant, with applying trim and fill method of Duvall and Tweedie (Sue & Richard, 2000). We performed a sensitivity analysis by removing one study to detect any elimination of the significance of the results. We reported uncertainty of our comparisons through reporting point estimate with 95% confidence interval (Tonin et al., 2017). Due to different follow-up endpoints relating to different aspects of VPs' outcomes, we did analysis upon the most repeated point in all studies of this outcome.

Results

Literature Search and Study Characteristics

The 201 included articles reported 11,918 laryngectomized patients (see Figure 1). One hundred twenty articles were included in our NMA; meanwhile, 27 articles were included in our meta-analysis. The mean age ranged between 17 and 90 years old, with males constituting the majority of participants (8546 = 71.7%). Most of the patients received radiotherapy 4340 (65%), with a follow-up range of 0.5–133 months. Detailed characteristics of the included RCTs are shown in Supplemental Table S5.

Risk of Bias Assessment

Risk of bias assessment of 169 non-RCT articles were assessed by National Institutes of Health tool and revealed 32 articles were ranked good, 107 were fair, and 30 were poor (see Supplemental Table S5). Regarding risk of bias for included 32, RCT trials were assessed by ROB Cochrane tool and revealed nine with low risk, 23 with unclear risk, and none with high risk (see Supplemental Figure S1). The overall incomplete outcome data and selective reporting bias remained low risk across studies, while allocation concealment, blinding of participants, and other bias were not adequately reported. In summary, all included RCTs were rated to be of either low or moderate risk (see Supplemental Figure S2).

Primary Outcomes

Devices' Replacements (27 Articles/5,724 Patients)

Through the 5,861 replacements that were done—some patients replaced more than one device (Van Den Hoogen et al., 1996)—2,256 were for Provox-1, 2,048 for Provox-2, 1,065 for Groningen LR, 272 for Blom-Singer low pressure, and 220 for Nijdam. The network of comparisons on rates of device replacements was shown in Supplemental Figures S3A–S3C, where Provox-1 had the lowest replacement rates than others (risk ratio [RR] = 0.69 [0.19, 2.51]; $p = .74$), followed by Nijdam (RR = 0.78 [0.09, 6.73]; $p = .64$) and Groningen ultra low resistance (ULR; RR = 0.81 [0.07, 9.14]; $p = .62$), respectively. The highest rate of replacement was figured in the SPVP (RR = 10.10 [0.71, 144.07]; $p = .05$) compared to BS-LP (see Supplemental Figure S3C).

Devices' Lifetime (33 Articles/4,777 Patients)

About 2,052 patients using Provox-1, 2,048 using Provox-2, 1,065 using Groningen LR, 272 using BS-LP, and 220 using Nijdam were included in the network of comparisons on mean devices' lifetime (see Supplemental Figures S4A–S4C), where Provox AV ranked the most living device (mean difference [MD] 17.25 [0.40, 34.09]; $p = .97$, years), while BS-LP showed the shortest life (MD 4.97 [−11.09, 21.03]; $p = .24$, years) compared to BS Advantage. Additionally, Nijdam was second in long lifetime of device (MD 9.87 [−6.67, 26.42]; $p = .73$, years), followed by Provox-1, Groningen LR, and Groningen ULR (see Supplemental Figure S4C).

Airflow Resistance (Eight Articles/1,850 Patients)

Patients using prosthesis suffered from the airflow resistance mostly while using the Nijdam (RR = 1.31 [0.22, 7.67]; $p = .28$), while the least percent of airflow resistance was clear in Provox 2 (RR = 0.42 [0.08, 2.11]; $p = .84$); those results were compared to Groningen LR (see Supplemental Figures S5A–S5C).

MPT (13 Articles/639 Patients)

Regarding voice control measurement, 229 were using BS-LP, 115 were using Provox-1, 69 were using Provox HME, 14 were using Groningen LR, and six were using Panje. The network of comparisons on mean MPT was shown in Supplemental Figures S6A–S6C, where Provox HME had the highest mean phonation time than others compared to BS Advantage (MD 6.30 [3.34, 9.26]; $p = 1$, seconds), followed by BS adjustable tracheostoma valve (MD 3.00 [−0.35, 6.35]; $p = .87$, seconds) and Provox FreeHands HME (MD −2.90 [−6.07,

0.27]; $p = .59$, seconds), respectively (see Supplemental Figure S6C).

Leakage Rates (40 Articles/1,493 Patients)

Leakage is the most reported complication, with 1,985 patients suffered from leakage around VPs. A network of comparisons on leakage rates was shown in Supplemental Figures S7A–S7C, where Provox Vega had the lowest leakage rate than others compared to BS-LP (RR = 1.87 [0.97, 3.60]; $p = .72$). While Provox-2 (RR = 2.08 [1.11, 3.88]; $p = .53$) and Nijdam (RR = 2.23 [1.27, 3.90]; $p = .47$) followed Provox Vega that recorded the best, respectively (see Supplemental Figure S7C).

Speech Rate (Four Articles/63 Patients)

Through using various devices, we compared speech rate versus BS-LP, where Groningen LR was found worse than BS-LP (MD −1.75 [−24.67, 21.17]; $p = .44$; see Supplemental Figures S8A–S8C).

Patient Device Preference (21 Articles/ 932 Patients)

In our analysis, we compared the devices versus BS-LP, and we figured that most of the patients preferred Provox-2 (OR = 33.88 [0.65, 1762.24]; $p = .92$), followed by Provox-1 (OR = 12.04 [0.27, 538.08]; $p = .78$), and Provox XtraHME (OR = 13.09 [0.18, 974.17]; $p = .78$). Meanwhile, the least preference VPs were for the external humidifier with (OR = 0.04 [0.00, 1.22]; $p = .09$; see Supplemental Figures S9A–S9C).

Stoma Cleaning (Two Articles/137 Patients)

In comparison to Provox-1, we have found that Provox HME needed less stoma cleaning than Provox-1 (MD −1.19 [−2.26, −0.12]; $p = .99$; see Supplemental Figures S10A–S10C).

Breathing Problems (Four Articles/144 Patients)

Our analysis compared those devices versus the external humidifier. Our results showed that Provox XtraHME (RR = 1.70 [0.08, 36.57]; $p = .43$) and Provox HME (RR = 1.70 (0.11, 27.40); p = .43), were associated with more breathing problems (see Supplemental Figures S11A–11C).

Coughing Frequency (Four Articles/246 Patients)

Regarding the coughing frequency, our analysis found the least coughing frequency for Provox XtraHME versus Provox-1 (MD −4.13 [−11.15, 2.88]; $p = .72$), followed by Provox HME (MD −3.63 [−7.53, 0.27]; $p = .70$; see Supplemental Figures S12A–S12C).

Forced Expectorations (Three Articles/193 Patients)

In comparison to Provox-1, Provox HME and Provox XtraHME (MD −13.32 [−35.58, 9.21]; $p = .69$) showed lesser forced expectorations than Provox-1 (MD −13.12 [−50.81, 24.56]; $p = .70$) (see Supplemental Figures S13A–S13C).

Sputum Production (Two Articles/110 Patients)

The production of sputum was investigated along the usage of prosthesis devices versus external humidifier, where Provox HME ranked the lowest device (MD −3.00 [−4.55, −1.45]; $p = .96$), followed by Provox-1 (MD −3.00 [−4.50, 3.90]; $p = 0.32$; see Supplemental Figures S14A–14C).

Sleeping Problems (Two Articles/85 Patients)

Versus the external humidifier, we figured that Provox HME is the lowest device causing sleeping problems (RR = 0.23 [0.09, 0.56]; $p = 1$), followed by Provox-1 (RR = 0.30 [0.12, 0.78]; $p = .50$; see Supplemental Figures S15A–S15C).

Loosening of Adhesive (Three Articles/ 129 Patients)

We figured that Provox HME was better than the external humidifier (MD −3.00 [−4.55, −1.45]; $p = .96$), followed by the Provox-1 (MD −0.30 [−4.50, 3.90]; $p = .32$; see Supplemental Figures S16A–S16C).

Increase Phonatory Effort (Four Articles/ 75 Patients)

Versus BS-LP, we found that Provox Vega has better results (OR = 4.11 [1.29, 13.06]; $p = .99$; see Supplemental Figures S17A–S17C).

Voice Speech Quality (Six Articles/620 Patients)

The quality of the voice speech was found better in Provox Vega versus BS-LP (OR = 16.41 [4.33, 62.22]; $p = 1$), followed by Panje (OR = 1 [0.16, 6.08]; $p = .25$; see Supplemental Figures S18A–S18C).

Fundamental Frequency (Eight Articles/ 148 Patients)

In comparison to Groningen LR, we found that SPVP recorded the highest place (MD 96.33 [17.29, 175.33]; $p = .99$), followed by the Provox-1 (MD 0.08 [−4.21, 4.36]; $p = .26$; see Supplemental Figures S19A–S19C).

Voice Loudness (Seven Articles/247 Patients)

Versus the BS-LP, all other prostheses were worse than BS-LP where Provox NID was the least VP in voice loudness (MD −1.00 [−4.62, 2.62]; $p = .39$; see Supplemental Figures S20A–S20C).

Speech Intelligibility (Seven Articles/692 Patients)

Versus the Groningen LR, we have investigated the speech intelligibility between devices to figure that Nijdam was the first place (MD 3.02 [0.12, 74.99]; $p = .87$), followed

by Groningen LR, and then Provox-1 (MD 0.10 [0.02, 0.55]; *p* = .01) in Supplemental Figures S21A–S21C.

Stoma Stenosis (Eight Articles/437 Patients)

In comparison to BS-LP, we investigated the stoma stenosis and figured out that Provox Vega has more stoma stenosis records (RR = 5.25 [2.06, 13.04]; *p* < .001; see Supplemental Figures S22A–S22C).

Dislodgement (31 Articles/2,977 Patients)

We have found Provox-2 had the least dislodgement results (RR = 0.27 [0.13, 0.57]; *p* = .79), followed by the Provox Vega (RR = 0.28 [0.12, 0.67]; *p* = .71; see Supplemental Figures S23A–S23C).

Fistula Problems (13 Articles/1,767 Patients)

The fistula recorded problems along with the usage of the prosthesis devices; we investigated the reported problems with the usage of various devices. Versus ESKA-Herrmann, Groningen LR was recorded the best VP with the least fistula problems (RR = 0.76 [0.50, 1.18]; *p* = .87), followed by Provox-1 (RR = 0.87 [0.66, 1.15]; *p* = .59), and Nijdam (RR = 0.96 [0.61, 1.51]; *p* = .31; see Supplemental Figures S24A–S24C).

Granulation (23 Articles/3,474 Patients)

Versus BS-LP, Provox-2 was found to have the least granulation formation (RR = 0.73 [0.02, 26.32]; *p* = .60), followed by Provox-1 (RR = 0.95 [0.13, 7.03]; *p* = .53), and Provox Vega (RR = 0.87 [0.01, 63.18]; *p* = .53), respectively, unlike the Nijdam that ranked the last with the highest granulation formation (RR = 1.93 [0.10, 36.18]; *p* = .30; see Supplemental Figures S25A–S25C).

Breathing Experience (Four Articles/116 Patients)

Versus the external humidifier, the Provox HME ranked the first place (*OR* = 2.03 [0.09, 47.12]; *p* = .59) followed by Provox XtraHME (*OR* = 2.03 [0.06, 65.56]; *p* = .58; see Supplemental Figures S26A–S26C).

Prosthesis Inaccurate Size (Four Articles/388 Patients)

Our investigation also extended to include the accuracy of the prosthesis size; versus Provox-1, we figured that Provox-2 had more accurate VP size (RR = 0.77 [0.23, 2.61]; *p* = .66; see Supplemental Figures S27A–S27C).

Prosthesis Deterioration (Three Articles/200 Patients)

In addition, we figured that Provox-2 recorded more VP deterioration than Provox-1 (RR = 2.62 [0.88, 7.81]; *p* = .04; see Supplemental Figures S28A–S28C).

Follow-Up Survival Rate (Three Articles/135 Patients)

Lastly, Provox-1 had better follow-up survival rate than BS-LP (*OR* = 1.99 [0.49, 8.15]; *p* = .83; see Supplemental Figures S29A–S29C).

Overall Hasse Results

Our results showed the devices ranking regarding overall 27 outcomes in terms of *p* values and rounded them to two decimals, where the best nearer to 1 and the worst nearer to zero, as given in Table 1. Provox-2 was the best VP with a mean *p* value of .59 in five outcomes, while Provox AV had the highest mean *p* value in only one outcome. Provox-2 was the most preferable by the patients, the best in airflow resistance, dislodgement, granulation, and prosthesis inaccurate size with *p* values of .89, .84, 1, .61, and .66, respectively (see Table 1). Besides, Provox AV had the best overall mean *p* value of .78 regarding only two outcomes, being the best in device life time only (*p* = 1). Provox-1 was the second best VP following Provox-2, being the best in MPT, device replacements, prosthesis deterioration, and follow-up survival rate with *p* values of .83, .74, .96, and .83, respectively (see Table 1). The lowest scores were settled by the Groningen LR and Nijdam with both mean *p* value of .44 (see Table 1). In the meantime, Groningen LR was the least in case of leakage, speech rate, MPT, and fundamental frequency with *p* values of .29, .44, .09, and .25, respectively (see Table 1). Nijdam was the worst in airflow resistance, patient device preference, fistula problems, and granulation formation with *p* values of .28, .09, .18, and .29, respectively (see Table 1). While Provox NID had the worst overall mean *p* value of .43 in only one outcome, which was voice loudness (*p* = .37; see Table 1).

Partial Order Setting According to Patients' Preferences

To control the confusion of multiple outcomes, we compared features of devices according to patients' preferences in order to determine what are the best choices. Provox-2 was considered the best device for both the airflow resistance and the leakage (see Figures 2A and 2B); nevertheless, the lack of Provox-2 may elect Provox-1 and Groningen LR to be its replacement, as determined by the patients' preference. Regarding phonation time, fundamental frequency, and speech intelligibility, Provox-2 was chosen as the best VP, followed by the Provox-1 according to patients' preference (see Figures 3A and 3B and Figure 4A). In case of dislodgement, the patients preferred the Provox-1, followed by Provox-2 (see Figure 4B). Regarding device replacement, if the Provox-1 was not available, the patients preferred the Provox-2, followed by Groningen (see Figure 5A). However, concerning the device lifetime, the favorable choice to the patients was Provox-1 followed by the Groningen (see Figure 5B). Regarding fistula problems, the patients have favored the Provox-1, followed by Provox-2 and Groningen

Table 1. Overall network meta-analysis Hasse table results for assessing the most efficient, safest, patient-tailored, longest lifetime, least complications voice prosthesis.

Outcomes	Blom-Singer LP	Groningen LR	Groningen ULR	Nijdam	Provox-1	Provox-2	Provox AV	Provox NID	Provox Vega	SPVP
Devices replacements	0.56	0.40	0.62	0.64	0.74	0.44	0.57	0.49	0.48	0.05
Device lifetime	0.10	0.58	0.62	0.72	0.62	0.31	1.00	NA	0.23	NA
Airflow resistance	NA	0.39	0.43	0.28	0.49	0.84	NA	NA	NA	NA
Leakage	0.99	0.29	NA	0.47	0.00	0.53	NA	NA	0.72	NA
Speech rate	0.51	0.44	NA	NA	NA	NA	NA	NA	NA	NA
Maximum phonation time	0.58	0.09	NA	NA	0.83	NA	NA	NA	NA	NA
Patient device preference	NA	0.35	NA	0.09	0.67	0.89	NA	NA	NA	NA
Increase phonatory effort	0.01	NA	NA	NA	NA	NA	NA	NA	0.99	NA
Fundamental frequency	NA	0.25	NA	NA	0.26	NA	NA	NA	NA	0.99
Voice loudness	0.64	NA	NA	NA	0.49	NA	NA	0.37	NA	NA
Speech intelligibility	NA	0.62	NA	0.87	0.01	NA	NA	NA	NA	NA
Stoma stenosis	1.00	NA	NA	NA	NA	NA	NA	NA	0.00	NA
Dislodgement	NA	NA	NA	NA	0.00	1.00	NA	NA	NA	NA
Fistula problems	NA	0.86	NA	0.18	0.46	NA	NA	NA	NA	NA
Granulation	0.51	0.52	NA	0.29	0.54	0.61	NA	NA	0.53	NA
Prosthesis inaccurate size	NA	NA	NA	NA	0.34	0.66	NA	NA	NA	NA
Prosthesis deterioration	NA	NA	NA	NA	0.96	0.04	NA	NA	NA	NA
Follow-up survival rate	0.17	NA	NA	NA	0.83	NA	NA	NA	NA	NA

Note. Groningen LR = Groningen low resistance; Groningen ULR = Groningen ultra low resistance; Provox NID = Provox non-indwelling device; Provox AV = Provox ActiValve; SPVP = Sound-Producing Voice Prosthesis; NA = not applicable.

Figure 2. (A) Partial ranking of devices according to airflow resistance. (B) Partial ranking of devices according to leakage. Groningen LR = Groningen low resistance; Groningen ULR = Groningen ultra low resistance; Provox NID = Provox non-indwelling device; Provox AV = Provox ActiValve; SPVP = Sound-Producing Voice Prosthesis.
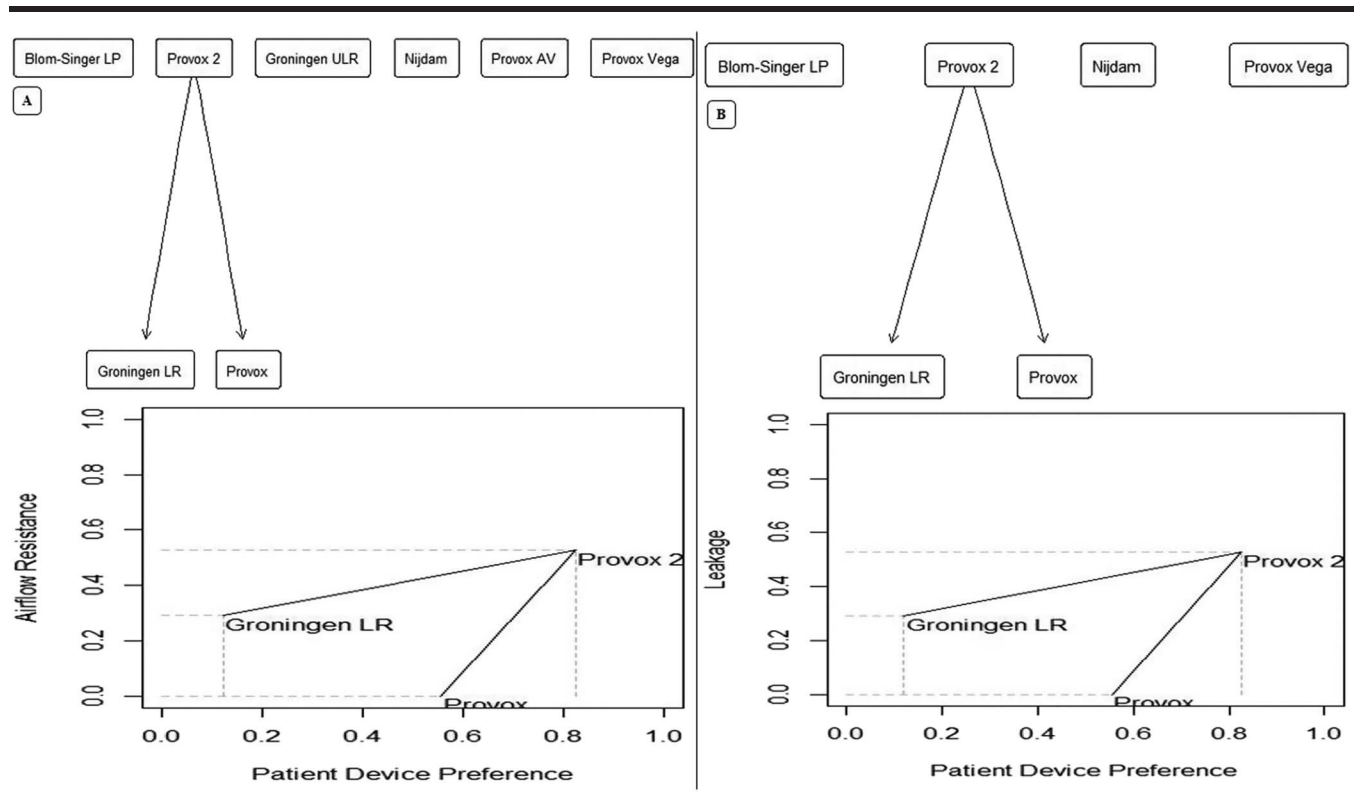
Figure 3. (A) Partial ranking of devices according to maximum phonation time. (B) Partial ranking of devices according to fundamental frequency. Groningen LR = Groningen low resistance; Groningen ULR = Groningen ultra low resistance; Provox NID = Provox non-indwelling device; Provox AV = Provox ActiValve; SPVP = Sound-Producing Voice Prosthesis.
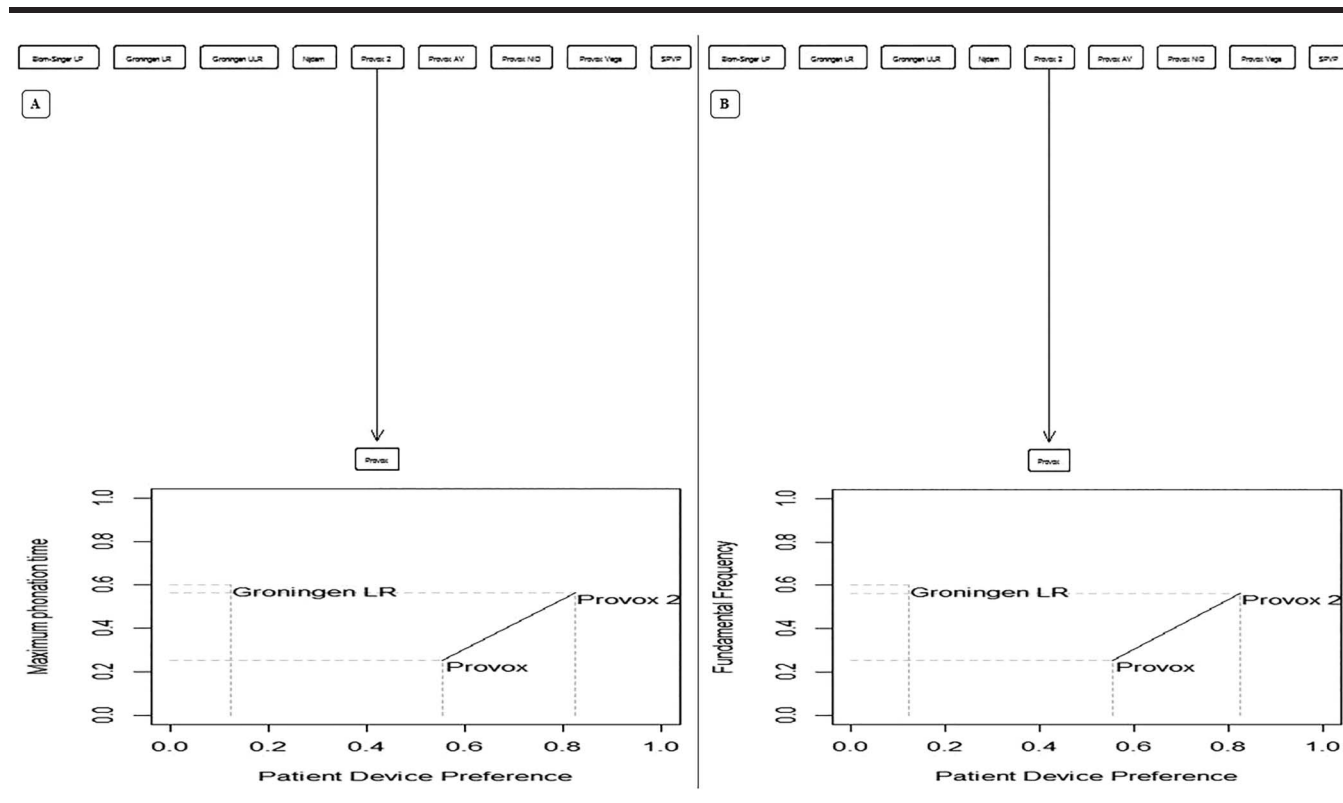

LR (see Figure 6A). For granulation, the patients favored Provox-2 in the first place, followed by the Provox-1 and then Groningen LR (see Figure 6B). For the size accuracy of prosthesis and prosthesis deterioration, Provox-1 was on the top of the list; if not available, patients chose to use Provox-2 and then Groningen LR (see Figures 7A and 7B).

Aspiration Pneumonia (Four Articles/274 Patients)

Regarding aspiration pneumonia, Provox-2 recorded the highest percentage of aspiration results (6.3%). On the other hand, Provox-1 had recorded the least device with aspiration pneumonia (3.4%; see Supplemental Figure S30A).

Fungal Colonization (Six Articles/213 Patients)

The fungus managed to colonize on top of prosthesis of 81% and 65.2% of Provox-2 and Provox-1, respectively. At the same time, half of the BS cases showed also colonization, leaving it as a grave burden for the laryngectomy patients (see Supplemental Figure S30B).

Experience With Speaking (Seven Articles/ 215 Patients)

Most of "Provox Hands-free HME ENB" patients proved they were fluent in speaking (79.2%), while the

"Provox Hands-free HME" were the least fluent patients (20.8%). On the other hand, half of BS-LP patients and Provox HME were fluent in speaking while using these devices (see Supplemental Figure S31A).

Skin Irritation (Nine Articles/417 Patients)

Among adverse effects, our meta-analysis showed that laryngectomized patients using a prosthesis (18.9%) reported skin irritation., where Provox StabiliBase recorded the highest skin irritation (57.1%; see Supplemental Figure S31B).

Qualitative Descriptive Synthesis
Chemoprophylaxis

Some studies have shown that chemoprophylaxis could be used for prevention of fungal infection of VPs. Weissenbruch et al. reported that the use of a buccal bioadhesive slow-release tablet containing an antimycotic agent prevented fungal colonization and deterioration of silicone VPs (Van Weissenbruch et al., 1997). Also, Ameye et al. agreed that these tablets increased significantly the lifetime of devices and more effective than local cleaning (Ameye et al., 2005). Driven et al. had reported that the new moldable external neck brace (ENB 2.0) achieved increasing the lifetime of the

Figure 4. (A) Partial ranking of devices according to speech intelligibility. (B) Partial ranking of devices according to dislodgement. Groningen LR = Groningen low resistance; Groningen ULR = Groningen ultra low resistance; Provox NID = Provox non-indwelling device; Provox AV = Provox ActiValve; SPVP = Sound-Producing Voice Prosthesis.
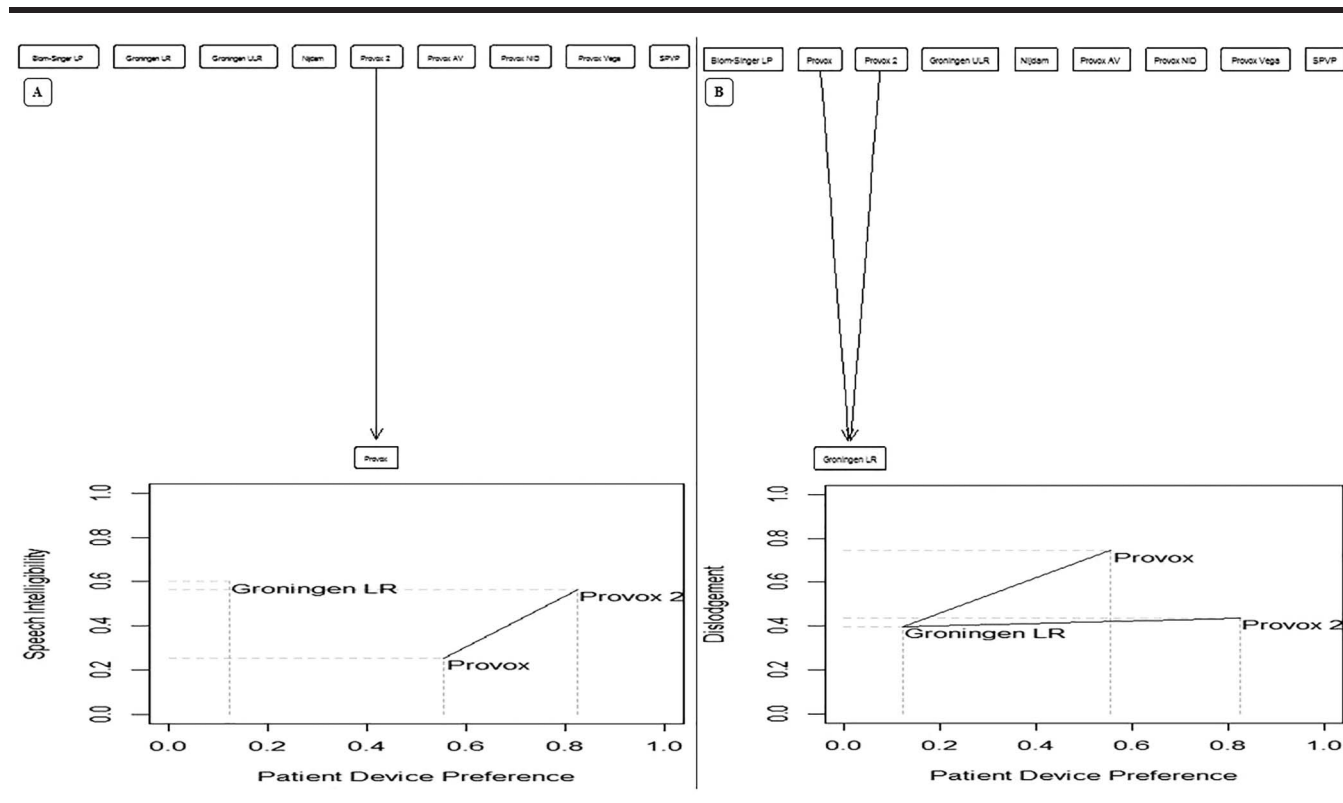
adhesive baseplate of hands-free speech (Dirven et al., 2013, 2012).

Cost

Three articles (Graville et al., 2011; Kesteloot et al., 1994; Kress et al., 2006) hailed the cost of each device with a significant variation; however, they agreed that the highest price is settled by the Provox AV or Provox-2. BS-LP cost the lowest price, according to one study.

Discussion

Our study results revealed that the top-ranked device according to our results was Provox-2, showing the best results regarding most of the VPs' outcomes, patients' device preference, airflow resistance, dislodgement, granulation formation, and VPs' inaccurate size. These results are supported by Harms et al. (2011), Kilic et al. (2014), Mastronikolis et al. (2008), Op de Coul et al. (2000), Terada et al. (2007), and Thylur et al. (2016), whose populations significantly favored Provox-2 results. On the contrary, Delsupehe et al. (1998) results have declared that both Provox-1 and Blom-Singer are equivocal in terms of overall voice quality, lifetime, and patient satisfaction. However, when interpreting these results, we should bear in mind that the Delsupehe study was performed before the Provox-2 invention.

The superiority of the Provox-2 over the other generations is justified by recording the best p value in the airflow resistance, patient device preference, dislodgement, granulation, and prosthesis inaccurate size. Considering that the increased airflow resistance was linked to the formation of biofilms over the device (Ishikawa et al., 2020), to our surprise, the Provox-2 device has shown more biofilm colonization when compared to its silicon control (Leonhard et al., 2018). However, the overall causes of increased airflow resistance were low in Provox-2, including the leakage through the prosthesis and excessive crusting (Laccourreye et al., 1997). Advantageously, while the dislodgement of the VP is thought to be a serious condition as it may cause the fistula tract (Masaany et al., 2009), the Provox-2 is the lowest device to cause fistulas. Besides, the formation of granulation tissue around the tracheoesophageal fistula represented a problem (Hagen, 1990) but Provox-2 was the best device to reduce the granulation tissue formation compared to other devices.

By assessing the device lifetime, Provox AV showed superiority over Provox-2, Provox-1, Groningen, and Nijdam. In contrast to our findings, Hoogen et al. (Van Den Hoogen et al., 1996) reported that Nijdam had higher lifetime over Provox-1 and Groningen upon using the right length of the device. However, with the long-term use of antifungal medications either in the form of a daily application of nystatin buccal bioadhesive tablets (Ameye et al., 2005) or

Figure 5. (A) Partial ranking of devices according to device replacement. (B) Partial ranking of devices according to device lifetime. Groningen LR = Groningen low resistance; Groningen ULR = Groningen ultra low resistance; Provox NID = Provox non-indwelling device; Provox AV = Provox ActiValve; SPVP = Sound-Producing Voice Prosthesis.
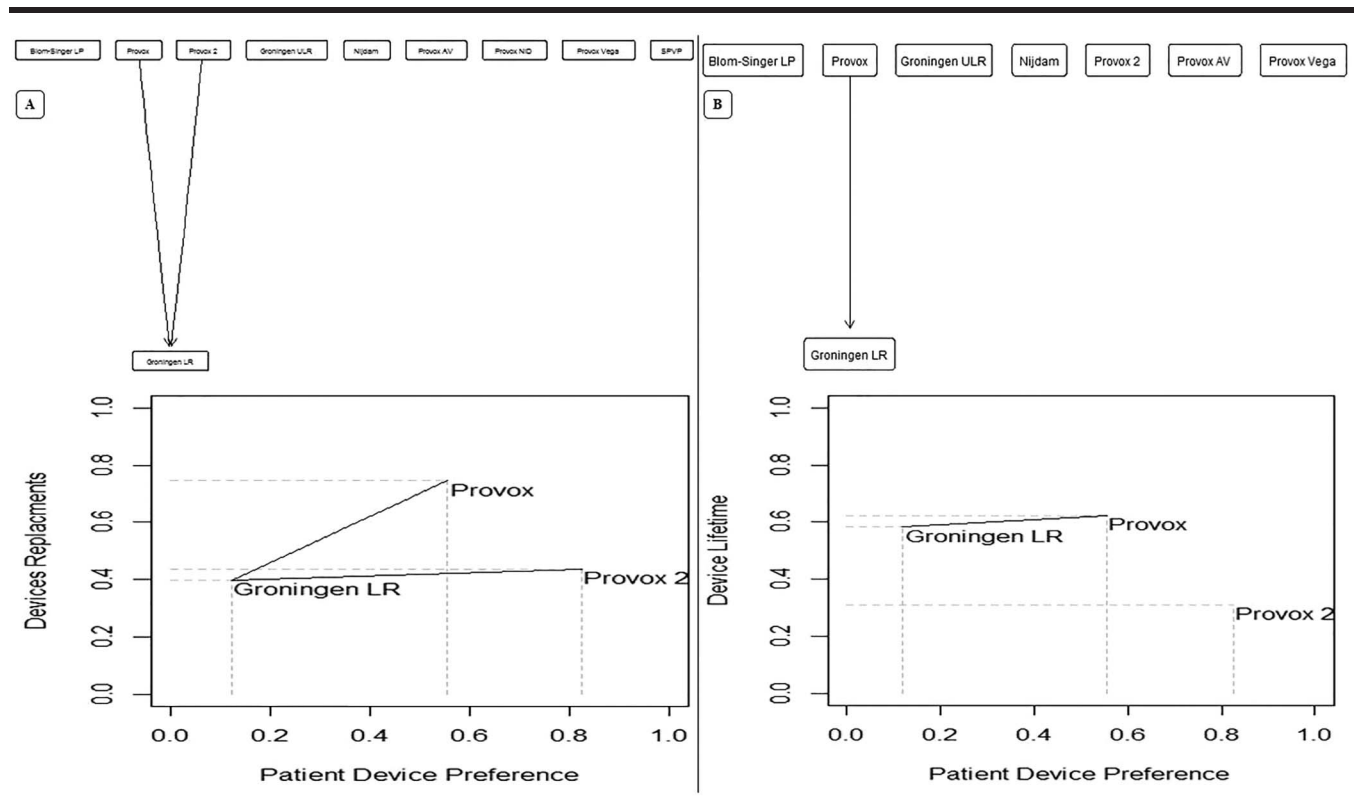
miconazole (Van Weissenbruch et al., 1997), Provox-1 would have a highly extended mean device lifetime compared to controls. Also, we can reply that Hoogen study was implemented in 1996, while the start of Provox and its generations were at the beginning of 1990s (Hilgers & Schouwenburg, 1990), leaving no space to have a well-balanced comparison.

On the other hand, the Groningen LR and Nijdam shared the same position as the worst device according to our analysis. Notably, Groningen LR reported the lowest scores in the leakage, MPT, speech rate, and fundamental frequency. On the other hand, Nijdam showed the highest airflow resistance, granulation, and fistula problems and was considered the least preferable device. Therefore, there was no preference for the Groningen LR nor the Nijdam, which came in concordance with our analysis (Chung et al., 1998; Van Den Hoogen et al., 1996).

Furthermore, Provox-2 had the highest mean MPT followed by Provox-1, Panje, and Groningen, respectively. However, some of the included studies found no significant difference between the assessed devices (Chung et al., 1998; Delsupehe et al., 1998), which may be a result of the low sample size. Moreover, Provox-1 showed the least leakage rates followed by Groningen, Nijdam, and BS, respectively. We can stem that back to the proper intervention to insert the Provox-1 and its design to solve the insertion problem.

Surprisingly, Provox-2 recorded aspiration pneumonia as a complication more than the Provox-1. However, there are numerous risk factors to cause aspiration pneumonia regardless the type of the device like the cough stimulation, the alcohol abuse, prior radiotherapy, and neurological diseases (Conte et al., 2012). Besides, the studies reported aspiration pneumonia represented small samples that may represent a bias. Therefore, a further investigation should find out if there is a true relationship between Provox-2 and the aspiration pneumonia.

Unfortunately, the smooth surfaces of Provox-2 and Provox-1 jeopardized the cleanliness of the prosthesis and left the chance for the bacteria to colonize. In fact, different designs have been used to reduce the biofilm formation over the valve starting from the hinged valve flap, slit valve, tripod ball valve, and the valveless type. Nevertheless, all valves were exposed to the colonization of the bacteria (Leonhard et al., 2010). Therefore, a further modification is needed to reduce the colonization and increase the lifetime of the device.

From the perspective of using supplementary gadgets, HME users had significantly less cough frequency, less sputum production, fewer stoma cleaning attempts, and lower frequency of forced expectoration in addition to less sleeping problems—but this finding is insignificant. However, a trial (Bień et al., 2010) reported a significant reduction in

Figure 6. (A) Partial ranking of devices according to fistula problems. (B) Partial ranking of devices according to granulation. Groningen LR = Groningen low resistance; Groningen ULR = Groningen ultra low resistance; Provox NID = Provox non-indwelling device; Provox AV = Provox ActiValve; SPVP = Sound-Producing Voice Prosthesis.
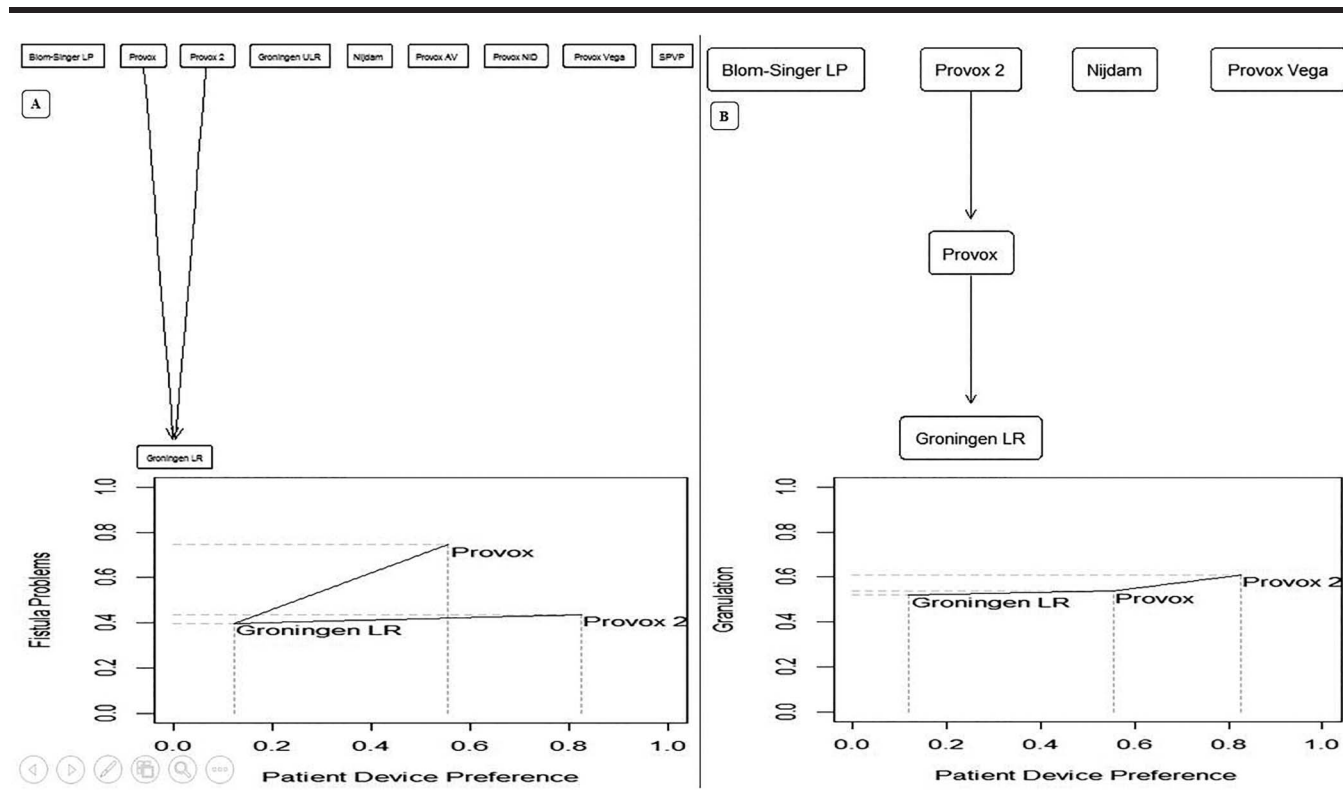
pulmonary problems of cough and forced expectoration and improvement in related aspects among those using HME, noting that compliance is the key to reach these findings. Cost perspective is an essential parameter to compare different VPs (Kress et al., 2014); however, the cost of the Provox-1 and its generations may be considered as a cost-effective (Kress et al., 2014).

Internal validity of our analysis is supported by many factors, by being conducted in strict accordance with the Cochrane handbook of systematic review (J. P. Higgins & Green, 2011) and reported according to the recommendations of the PRISMA statement (Moher et al., 2009). Choosing the NMA approach resulted in applying evidence to a wider range of patients giving more reliable results and usage of *p* values ranking in scatter plots giving a higher applicability according to partial ranking order assumption (Rücker & Schwarzer, 2015).

Limitations

First, infeasibility of some factors as financial status, radiotherapy, dose/duration, and time since radiotherapy to be analyzed in an NMA was the main limitation because of the lack of a common comparator, which made it difficult to assess the impact of such factors on defining various aspects of efficiency of VPs and its effect on the quality of

life. We tried to solve this issue by assessing it using a qualitative analysis after reviewing several related articles in the literature. Besides, choosing the NMA approach to put different kinds of devices in a ranked order resulted in applying evidence to a wider range of individuals giving more reliable results. A huge number of population included from different regions across the world make our study an important stepping stone for further assessment of this problem for future implications.

The second limitation is that the different devices are developed at different years and the newer devices are supposed to be made of better materials to achieve better outcomes, yet this issue was not investigated before and it might be a source of bias. However, we included all types of VPs without any exclusion to any specific one, as for fear of selection bias, and for bias of how to classify exact limit for which are new/old devices. Therefore, we provided all the currently used names from the literature and the manufacturing company. Lastly, the availability of the devices is quite variable based on the country of practice.

Conclusions

Provox-2 was found to be the best device than others regarding patients' preference, airflow resistance, dislodgment, granulation formation, and prosthesis inaccurate size,

Figure 7. (A) Partial ranking of devices according to prosthesis accuracy. (B) Partial ranking of devices according to prosthesis deterioration. Groningen LR = Groningen low resistance; Groningen ULR = Groningen ultra low resistance; Provox NID = Provox non-indwelling device; Provox AV = Provox ActiValve; SPVP = Sound-Producing Voice Prosthesis.
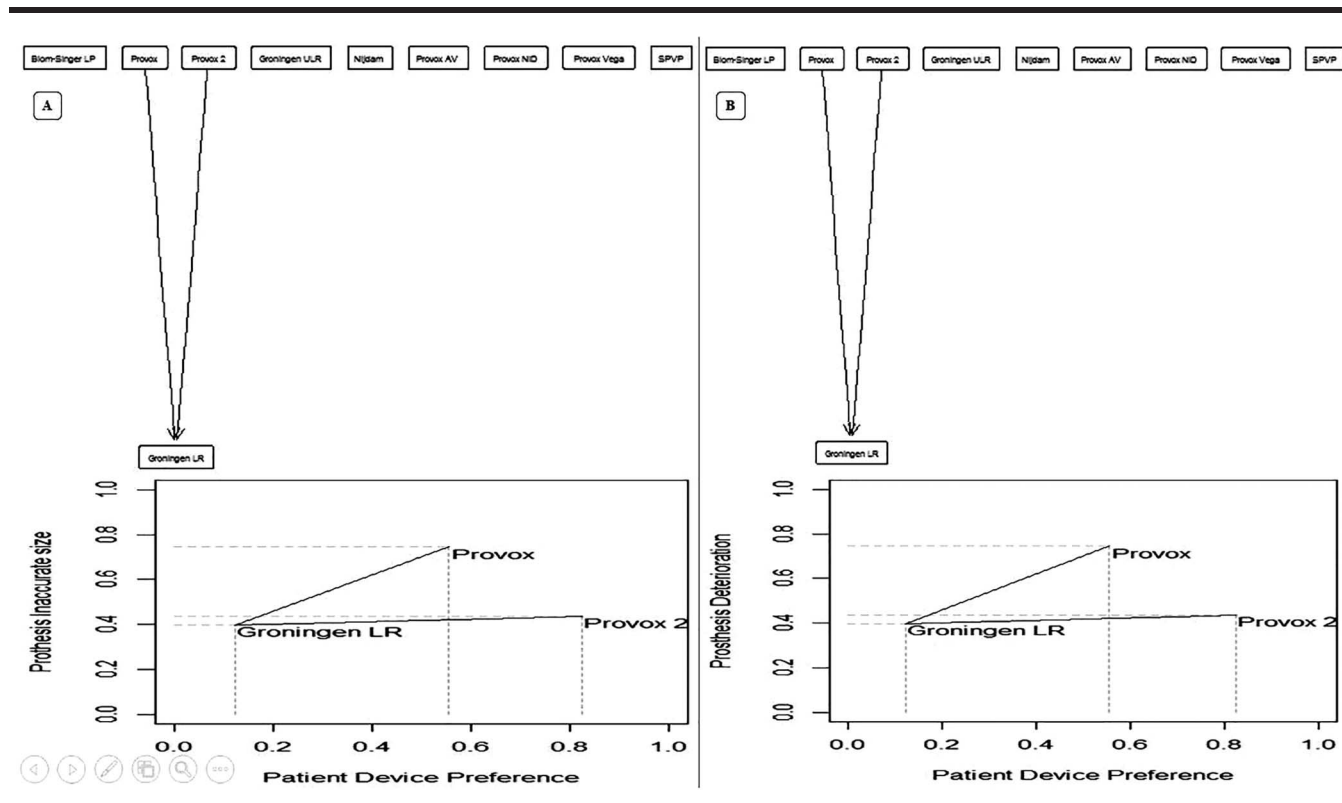
while Provox-1 has shown to be the second choice in terms of MPT, device replacement frequency, voice loudness, devices deterioration, and follow-up survival rates. Patients favored adding HME on top of their devices, as they decrease VP problems such as stoma cleaning, breathing and sleeping problems, coughing frequency, forced expectorations, sputum production, and loosening of adhesives. Besides, HME had better results than other devices regarding MPT and breathing experience. Groningen LR and Nijdam were considered the worst devices according to our analysis.

Author Contributions

G. M. T. and A. G. S. were responsible for the idea under the supervision of N. T. H. Screening and extraction were done by all authors under the supervision of N. T. H. Data analysis and its interpretation were done by G. M. T. and A. H. Z. Tables and figures were done by G. M. T. and O. M. M. All authors contributed to the review article writing and approval of the final version.

Acknowledgments

Thanks to Thuan Minh Tieu, Faculty of Health Sciences, McMaster University, Hamilton, ON, Canada; Võ Thị Trâm Anh, School of Medicine, Viet Nam National University, Ho Chi Minh City, Vietnam; and Nusaiba Adam Yahia, Faculty of Medicine, Omdurman Islamic University, Khartoum, Sudan, for their previous efforts in this study.

References

Ameye, D., Honraet, K., Loose, D., Vermeersch, H., Nelis, H., & Paul Remon, J. (2005). Effect of a buccal bioadhesive nystatin tablet on the lifetime of a Provox silicone tracheoesophageal voice prosthesis. *Acta Oto-Laryngologica, 125*(3), 304–306. https://doi.org/10.1080/00016480410022778

Begg, C. B., & Mazumdar, M. (1994). Operating characteristics of a rank correlation test for publication bias. *Biometrics, 50*(4), 1088–1101. https://doi.org/10.2307/2533446

Bień, S., Okła, S., van As-Brooks, C. J., & Ackerstaff, A. H. (2010). The effect of a heat and moisture exchanger (Provox HME) on pulmonary protection after total laryngectomy: A randomized controlled study. *European Archives of Oto-Rhino-Laryngology, 267*(3), 429–435. https://doi.org/10.1007/s00405-009-1018-4

Borenstein, M., Rothstein, D., & Cohen, J. (2005). *Comprehensive Meta-Analysis: A computer program for research synthesis* [Computer software]. Biostat.

Brown, D. H., Hilgers, F. J., Irish, J. C., & Balm, A. J. (2003). Postlaryngectomy voice rehabilitation: State of the art at the millennium. *World Journal of Surgery, 27*(7), 824–831. https://doi.org/10.1007/s00268-003-7107-4

Chung, R., Patel, P., Ter, M. K., Van, J. L. B., & Mahieu, H. (1998). In vitro and in vivo comparison of the low-resistance

Groningen and the Provox tracheosophageal voice prostheses. *Revue de Laryngologie-Otologie-Rhinologie, 119*(5), 301–306.

Conte, S. C., De Nardi, E., Conte, F., & Nardini, S. (2012). Aspiration of tracheoesophageal prosthesis in a laryngectomized patient. *Multidisciplinary Respiratory Medicine, 7*(1), 25. https://doi.org/10.1186/2049-6958-7-25

Delsupehe, K., Zink, I., Lejaegere, M., & Delaere, P. (1998). Prospective randomized comparative study of tracheoesophageal voice prosthesis: Blom-Singer versus Provox. *The Laryngoscope, 108*(10), 1561–1565. https://doi.org/10.1097/00005537-199810000-00026

Dirven, R., Clark, J. R., Wismans, J. G. F., McGuiness, J., Palme, C. E., Blyth, K., Baxter, C., Stone, D. B., & Marres, H. A. M. (2013). A new custom moldable external neck brace (ENB 2.0) to improve hands-free speech in laryngectomized patients. *The Laryngoscope, 123*(9), 2209–2215. https://doi.org/10.1002/lary.24044

Dirven, R., Kooijman, P. G. C., Wouters, Y., & Marres, H. A. M. (2012). Clinical use of a neck brace to improve hands-free speech in laryngectomized patients. *The Laryngoscope, 122*(6), 1267–1272. https://doi.org/10.1002/lary.23299

Global Burden of Disease Cancer Collaboration. (2015). The global burden of cancer 2013. *JAMA Oncology, 1*(4), 505–527. https://doi.org/10.1001/jamaoncol.2015.0735

Graville, D. J., Palmer, A. D., Andersen, P. E., & Cohen, J. I. (2011). Determining the efficacy and cost-effectiveness of the activalve: Results of a long-term prospective trial. *The Laryngoscope, 121*(4), 769–776. https://doi.org/10.1002/lary.21380

Hagen, R. (1990). Voice rehabilitation following total laryngectomy: Microvascular laryngeal replacement-plasty (laryngoplasty) instead of voice prosthesis. *Laryngo-Rhino-Otologie, 69*(4), 213–216. https://doi.org/10.1055/s-2007-998177

Hancock, K., Ward, E., Lawson, N., & van As-Brooks, C. J. (2012). A prospective, randomized comparative study of patient perceptions and preferences of two types of indwelling voice prostheses. *International Journal of Language & Communication Disorders, 47*(3), 300–309. https://doi.org/10.1111/j.1460-6984.2011.00109.x

Hancock, K. L., Lawson, N. R., & Ward, E. C. (2013). Device life of the Provox Vega voice prosthesis. *European Archives of Oto-Rhino-Laryngology, 270*(4), 1447–1453. https://doi.org/10.1007/s00405-012-2154-9

Harms, K., Post, W. J., van de Laan, K. T., van den Hoogen, F. J., Eerenstein, S. E., & van der Laan, B. F. (2011). A prospective randomized multicenter clinical trial of the Provox2 and Groningen Ultra Low Resistance voice prostheses in the rehabilitation of post-laryngectomy patients: A lifetime and preference study. *Oral Oncology, 47*(9), 895–899. https://doi.org/10.1016/j.oraloncology.2011.04.004

Hernandez, B. Y., Green, M. D., Cassel, K. D., Pobutsky, A. M., Vu, V., & Wilkens, L. R. (2010). Cancer Research Center Hotline: Preview of Hawai'i Cancer Facts and Figures 2010. *Hawaii Medical Journal, 69*(9), 223–224.

Higgins, J., Jackson, D., Barrett, J., Lu, G., Ades, A., & White, I. (2012). Consistency and inconsistency in network meta-analysis: Concepts and models for multi-arm studies. *Research Synthesis Methods, 3*(2), 98–110. https://doi.org/10.1002/jrsm.1044

Higgins, J. P., Altman, D. G., Gøtzsche, P. C., Jüni, P., Moher, D., Oxman, A. D., & Sterne, J. A. J. B. (2011). The Cochrane Collaboration's tool for assessing risk of bias in randomised trials. *BMJ, 343,* d5928. https://doi.org/10.1136/bmj.d5928

Higgins, J. P., & Green, S. (2011). *Cochrane handbook for systematic reviews of interventions (Vol. 4)*. Wiley.

Hilgers, F. J., & Schouwenburg, P. F. (1990). A new low-resistance, self-retaining prosthesis (Provox) for voice rehabilitation after total laryngectomy. *The Laryngoscope, 100*(11), 1202–1207. https://doi.org/10.1288/00005537-199011000-00014

Hozo, S. P., Djulbegovic, B., & Hozo, I. (2005). Estimating the mean and variance from the median, range, and the size of a sample. *BMC Medical Research Methodology, 5*(1), Article 13. https://doi.org/10.1186/1471-2288-5-13

Hutton, B., Salanti, G., Caldwell, D. M., Chaimani, A., Schmid, C. H., Cameron, C., Ioannidis, J. P. A., Straus, S., Thorlund, K., Jansen, J. P., Mulrow, C., Catalá-López, F., Gøtzsche, P. C., Dickersin, K., Boutron, I., Altman, D. G., & Moher, D. (2015). The PRISMA extension statement for reporting of systematic reviews incorporating network meta-analyses of health care interventions: Checklist and explanations. *Annals of Internal Medicine, 162*(11), 777–784. https://doi.org/10.7326/M14-2385

Huy, N. T., Van Giang, T., Thuy, D. H. D., Kikuchi, M., Hien, T. T., Zamora, J., & Hirayama, K. (2013). Factors associated with dengue shock syndrome: A systematic review and meta-analysis. *PLOS Neglected Tropical Diseases, 7*(9), Article e2412. https://doi.org/10.1371/journal.pntd.0002412

Ishikawa, Y., Yanagi, Y., Suzuki, M., & Konomi, U. (2020). A questionnaire to assess phonation and problems in daily living for Provox voice prosthesis users in Japan. *Indian Journal of Otolaryngology and Head & Neck Surgery, 1–10.* https://doi.org/10.1007/s12070-020-01792-2

Jemal, A., Siegel, R., Ward, E., Murray, T., Xu, J., & Thun, M. J. (2007). Cancer statistics, 2007. *CA: A Cancer Journal for Clinicians, 57*(1), 43–66. https://doi.org/10.3322/canjclin.57.1.43

Kesteloot, K., Nolis, I., Huygh, J., Delaere, P., & Feenstra, L. (1994). Costs and effects of tracheoesophageal speech compared with esophageal speech in laryngectomy patients. *Acta Oto-rhino-laryngologica Belgica, 48*(4), 387–394.

Kilic, C., Tuncel, U., & Comert, E. (2014). Provox 2 use for voice restoration after total laryngectomy. *Kulak Burun Bogaz Ihtis Derg, 24*(6), 339–343. https://doi.org/10.5606/kbbihtisas.2014.63497

Koch, W. M. (2002). Total laryngectomy with tracheoesophageal conduit. *Otolaryngologic Clinics of North America, 35*(5), 1081–1096. https://doi.org/10.1016/S0030-6665(02)00040-3

Kress, P., Schäfer, P., & Schwerdtfeger, F. (2006). Clinical use of a voice prosthesis with a flap valve containing silver oxide (Blom-Singer Advantage), biofilm formation, in-situ lifetime and indication. *Laryngo-Rhino-Otologie, 85*(12), 893–896. https://doi.org/10.1055/s-2006-925292

Kress, P., Schäfer, P., Schwerdtfeger, F., & Rösler, S. (2014). Are modern voice prostheses better? A lifetime comparison of 749 voice prostheses. *European Archives of Oto-Rhino-Laryngology, 271*(1), 133–140. https://doi.org/10.1007/s00405-013-2611-0

Krishnan, S., & Maclean, J. (2013). Practice of laryngectomy rehabilitation interventions: A perspective from Australia. *Current Opinion in Otolaryngology & Head and Neck Surgery, 21*(3), 224–229. https://doi.org/10.1097/MOO.0b013e32836118aa

Laccourreye, O., Ménard, M., Crevier-Buchman, L., Couloigner, V., & Brasnu, D. (1997). In situ lifetime, causes for replacement, and complications of the Provox voice prosthesis. *The Laryngoscope, 107*(4), 527–530. https://doi.org/10.1097/00005537-199704000-00018

Leonhard, M., Moser, D., Reumueller, A., Mancusi, G., Bigenzahn, W., & Schneider-Stickler, B. (2010). Comparison of biofilm formation on new Phonax and Provox 2 voice prostheses—A pilot study. *Head & Neck, 32*(7), 886–895. https://doi.org/10.1002/hed.21276

Leonhard, M., Zatorska, B., Tan, Y., Moser, D., & Schneider-Stickler, B. (2018). In vitro biofilm growth on modern voice prostheses. *Head & Neck, 40*(4), 763–769. https://doi.org/10.1002/hed.25053

Lorenz, K. J. (2017). Rehabilitation after total laryngectomy—A tribute to the pioneers of voice restoration in the last two centuries. *Frontiers in Medicine, 4,* 81. https://doi.org/10.3389/fmed.2017.00081

Mahieu, H., Annyas, A., & Nijdam, H. (1986). The Groningen button results. In I. F. Herman (Ed.), *Speech restoration via voice prostheses* (pp. 26–32). Springer. https://doi.org/10.1007/978-3-642-71415-3_7

Masaany, M., Marina, M., Asma, A., & Sani, A. (2009). Prevention of aspiration following recurrent dislodgement of voice prosthesis: A patient's innovation. *The Journal of Laryngology & Otology, 123*(6), 680–682. https://doi.org/10.1017/S0022215108002958

Mastronikolis, N. S., Stathas, T., Papadas, T. A., Spyropoulos, K., Naxakis, S., & Goumas, P. D. (2008). Voice restoration after total laryngectomy using Provox 2 (generation II) prosthesis. *European Review for Medical and Pharmacological Sciences, 12*(5), 327–330.

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & The PRISMA Group. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLOS Medicine, 6*(7), Article e1000097. https://doi.org/10.1371/journal.pmed.1000097

National Heart, Lung, and Blood Institute. (n.d.). Study quality assessment tools. https://www.nhlbi.nih.gov/health-topics/study-quality-assessment-tools

Op de Coul, B. M., Hilgers, F. J., Balm, A. J., Tan, I. B., van den Hoogen, F. J., & van Tinteren, H. (2000). A decade of postlaryngectomy vocal rehabilitation in 318 patients: A single Institution's experience with consistent application of Provox indwelling voice prostheses. *Archives of Otolaryngology—Head & Neck Surgery, 126*(11), 1320–1328. https://doi.org/10.1001/archotol.126.11.1320

Rücker, G., & Schwarzer, G. (2015). Ranking treatments in frequentist network meta-analysis works without resampling methods. *BMC Medical Research Methodology, 15,* Article 58. https://doi.org/10.1186/s12874-015-0060-8

Rücker, G., & Schwarzer, G. (2017). Resolve conflicting rankings of outcomes in network meta-analysis: Partial ordering of treatments. *Research Synthesis Methods, 8*(4), 526–536. https://doi.org/10.1002/jrsm.1270

Rücker, G., Schwarzer, G., Krahn, U., & König, J. (2015). *netmeta: Network meta-analysis using frequentist methods (R package Version 0.8-0)* [Computer software]. https://cran.r-project.org/package=netmeta

Slavícek, A., Betka, J., Novák, V., Kinkal, J., Wasserbauer, R., Vymazalová, Z., & Last'ovicková, B. (2000). Biodegradation of silicone in voice prosthesis. *Sb Lek, 101*(4), 325–339.

Stankovic, I., Milisavljevic, D., & Stankovic, M. (2017). Survival after salvage total laryngectomy: The influence of previous treatment. *Journal of Nuclear Medicine & Radiation Therapy, 8*(322), 2. https://doi.org/10.4172/2155-9619.1000322

Sue, D., & Richard, T. (2000). Trim and fill: A simple funnel-plot–based method of testing and adjusting for publication bias in meta-analysis. *Biometrics, 56*(2), 455–463. https://doi.org/10.1111/j.0006-341X.2000.00455.x

Tawfik, G. M., Dila, K. A. S., Mohamed, M. Y. F., Tam, D. N. H., Kien, N. D., Ahmed, A. M., & Huy, N. T. (2019). A step by step guide for conducting a systematic review and meta-analysis with simulation data. *Tropical Medicine and Health, 47,* 46. https://doi.org/10.1186/s41182-019-0165-6

Terada, T., Saeki, N., Toh, K., Uwa, N., Sagawa, K., Takayasu, S., & Sakagami, M. (2007). Voice rehabilitation with Provox2 voice prosthesis following total laryngectomy for laryngeal and hypopharyngeal carcinoma. *Auris Nasus Larynx, 34*(1), 65–71. https://doi.org/10.1016/j.anl.2006.09.017

Thylur, D. S., Villegas, B. C., Fisher, L. M., Sinha, U. K., & Kokot, N. (2016). Device life of two generations of Provox voice prostheses. *Annals of Otology, Rhinology & Laryngology, 125*(6), 501–507. https://doi.org/10.1177/0003489415624701

Tonin, F. S., Rotta, I., Mendes, A. M., & Pontarolo, R. (2017). Network meta-analysis: A technique to gather evidence from direct and indirect comparisons. *Pharmacy Practice, 15*(1), 943–943. https://doi.org/10.18549/PharmPract.2017.01.943

Van Den Hoogen, F. J., Oudes, M. J., Hombergen, G., Nijdam, H. F., & Manni, J. J. (1996). The Groningen, Nijdam and Provox voice prostheses: A prospective clinical comparison based on 845 replacements. *Acta Oto-Laryngologica, 116*(1), 119–124. https://doi.org/10.3109/00016489609137724

Van Rijkom, H., Truin, G., & Van't Hof, M. (1998). A meta-analysis of clinical studies on the caries-inhibiting effect of fluoride gel treatment. *Caries Research, 32*(2), 83–92. https://doi.org/10.1159/000016436

Van Weissenbruch, R., Nelis, H. J., Bouckaert, S., Aerts, R., Remon, J.-P., & Albers, F. W. (1997). Chemoprophylaxis of fungal deterioration of the Provox silicone tracheoesophageal prosthesis in postlaryngectomy patients. *Annals of Otology, Rhinology & Laryngology, 106*(4), 329–337. https://doi.org/10.1177/000348949710600413

Wan, X., Wang, W., Liu, J., & Tong, T. (2014). Estimating the sample mean and standard deviation from the sample size, median, range and/or interquartile range. *BMC Medical Research Methodology, 14,* Article 135. https://doi.org/10.1186/1471-2288-14-135

Ward, E. C., Hancock, K., Lawson, N., & van As–Brooks, C. J. (2011). Perceptual characteristics of tracheoesophageal speech production using the new indwelling Provox Vega voice prosthesis: A randomized controlled crossover trial. *Head & Neck, 33*(1), 13–19. https://doi.org/10.1002/hed.21389

Zuur, J. K., Muller, S. H., Vincent, A., Sinaasappel, M., de Jongh, F. H. C., & Hilgers, F. J. M. (2008). Assessment of tracheal temperature and humidity in laryngectomized individuals and the influence of a heat and moisture exchanger on tracheal climate. *Head & Neck, 30*(8), 1072–1082. https://doi.org/10.1002/hed.20835

Zuur, J. K., Muller, S. H., Vincent, A., Sinaasappel, M., de Jongh, F. H., & Hilgers, F. J. M. (2009). The influence of a heat and moisture exchanger on tracheal climate in a cold environment. *Medical Engineering and Physics, 31*(7), 852–857. https://doi.org/10.1016/j.medengphy.2009.04.004