CADTH

CADTH RAPID RESPONSE REPORT:
SUMMARY WITH CRITICAL APPRAISAL

Indwelling Voice Prostheses for Adults Following Laryngectomy: A Review of Clinical Effectiveness, Cost-Effectiveness, and Guidelines

Service Line:	Rapid Response Service
Version:	1.0
Publication Date:	December 13, 2017
Report Length:	28 Pages

CADTH

Authors: Charlotte Wells, Lorna Adcock

Cite As: Indwelling voice prostheses for adults following laryngectomy: a review of clinical effectiveness, cost-effectiveness, and guidelines. Ottawa: CADTH; 2017 Dec. (CADTH rapid response report: summary with critical appraisal).

Acknowledgments:

ISSN: 1922-8147 (online)

About CADTH: CADTH is an independent, not-for-profit organization responsible for providing Canada's health care decision-makers with objective evidence to help make informed decisions about the optimal use of drugs, medical devices, diagnostics, and procedures in our health care system.

Context and Policy Issues

Total laryngectomies are a form of surgery which entirely removes the larynx and some adjacent tissues from a patient's throat, often as a treatment option for laryngeal and hypopharyngeal cancers.[1,2] Usually, in a total laryngectomy, the thyroid and cricoid cartilages, the arytenoid cartilage, the epiglottis, the hyoid bone, and the prelaryngeal muscles are all removed and the airway is separated from the esophagus.[1] A stoma connecting the remaining tracheal tube to the outside air is then created on the neck, through which the patient breathes.[1]

Total laryngectomies can be debilitating for patients, as it removes or hinders the ability for a patient to speak, smell, and communicate, and creates a hole in the neck that requires constant care.[1] The negative impact of total laryngectomy on quality of life and social relationships is well established.[3,4] Additionally, the negative impacts of a total laryngectomy extends to the spouses of patients, with psychological distress being prevalent in this group.[3]

There are three options for total laryngectomy patients to recover their speaking ability, with approximately 85% to 90% of laryngectomees regaining their ability to communicate verbally.[2] One option is to use an electropharynx, or an artificial larynx, which is an externally operated vibrator placed on the cheek or chin. It creates vibration in the throat that is then manipulated into speech by the patient's mouth.[2,5] Although this is the quickest restoration available, it can sound robotic.[5] A second option is esophageal speech, in which a patient creates vibration by insufflating the esophagus and then "belching" the air out. The patient then manipulates these vibrations into speech. This is usually the most difficult method to learn but circumvents the need to use medical devices or external instruments.[2,5]

The final option for laryngectomees to regain speech is through the use of indwelling or non-indwelling voice prostheses. A fistula which connects the trachea and the esophagus is surgically created (a tracheoesophageal puncture, or TEP) and a voice prosthesis with a one way artificial valve is placed to prevent aspiration of food and liquids.[4,5] The patient can then occlude the stoma and direct air into the esophagus through the valve to produce speech.[4,5] Patients will frequently wear a heat moisture exchange (HME), which covers the stoma and filters dust, as well as preserve moisture in the respiratory tract.[2] Whilst the voicing of TEP prostheses sounds the most natural and produces the best vocal quality, prostheses can malfunction, cost significantly more than other methods, and requires the physical ability of the patient to successfully occlude the stoma (often rendering patients with musculoskeletal issues, amputations or previous strokes ineligible for this treatment).[4] Patients must also have the manual dexterity to clean the prosthesis effectively.[6]

Within TEP prostheses, patients have the choice of an indwelling voice prostheses or non-indwelling voice prosthesis. Non-indwelling prostheses are removable by the patient and can be changed, creating a higher level of independence for the patient. However, non-indwelling prostheses require the stoma be easily accessible, and the patient to have adequate eyesight and dexterity to remove and reinsert the device.[2] Indwelling prostheses, prostheses that are exclusively changed by the physician or speech-language pathologist are another option, and often last longer than non-indwelling prostheses.[2] Generally there is

no difference in voicing between the two devices,[2] but some patients prefer indwelling prostheses as they have reduced maintenance and care, and do not require taping the attached safety strap to their neck.[7] There are many standard indwelling prostheses available for patients, and there are additionally specialty indwelling prostheses available, which are designed to be more durable and to have increased resistance to air flow.[8] Many indwelling prostheses are available as sterile or as non-sterile prostheses, and some prostheses have silver oxide incorporated into the device to prevent *Candida* growth.[1]

The aim of this review is to evaluate the comparative clinical evidence regarding the longevity of varying types and brands of indwelling voice prostheses, and to evaluate cost-effectiveness outcomes to support reimbursement decision making.

Research Questions

1. What is the clinical effectiveness of indwelling voice prostheses for adults following laryngectomy?
2. What is the cost-effectiveness of indwelling voice prostheses for adults following laryngectomy?
3. What are the evidence-based guidelines regarding the use of indwelling voice prostheses for adults following laryngectomy?

Key Findings

Five clinical studies were identified regarding the clinical effectiveness of indwelling voice prostheses for adults following laryngectomy. One of these identified studies was a randomized controlled trial (RCT), two were prospective non-randomized studies, and two were retrospective non-randomized studies. Four studies examined longevity outcomes for indwelling voice prostheses, and two studies examined patient preferences and perspectives regarding their prosthesis.

On average, indwelling prostheses appear to fail between 53 to 298 days after insertion. However, the length of time between insertion and failure varied between and within the included studies. The length of time between insertion and failure also varied between different patients within the same treatment group. Within the included studies there was variation in device lifespan, for example, in one study, Provox Vega® had a significantly longer lifespan than Provox 2®, but in a separate study, the opposite occurred. In one study, no significant differences in device life were found between some standard voice prostheses (Provox 2® compared with Blom-Singer Classic® and Blom-Singer Dual Valve®, and Provox Vega® compared with Blom-Singer Classic® and Blom-Singer Dual Valve®). In one study, more specialized prostheses (prostheses designed to be stronger and have increased resistance to airflow) appeared to have longer lifespans than more standard prostheses, but in another study there was no difference between specialty and standard prostheses.

Overall, patients appeared to prefer Provox Vega® prostheses to Blom-Singer Classic® prostheses in one study, mostly due to their perceptions of easier and more effective cleaning of Provox Vega®. However the patient responses were diverse, demonstrating that individual patients have differing opinions on prosthesis use.

These results should be interpreted with caution as the intervention and comparator groups were not always comparable. Most studies were of poor quality, with methodological concerns surrounding selection of patient groups, no randomization (with the exception of one RCT), inappropriate statistical methods, and self-reporting bias. The sample sizes in many of the included studies were also small, which may have resulted in the studies being underpowered. There was also a lack of controlling for confounding variables, for example, diet, socio-economic status and prosthesis hygiene habits. There are also limitations with regards to generalizability to the Canadian context, and generalizability to the general laryngectomee population, which is also a heterogeneous population in itself. Additionally, there were few or no direct comparisons between some prostheses types. Finally, only primary studies were available in the literature. No cost effectiveness studies, economic evaluations, or evidence-based guidelines were identified. Therefore, the body of evidence described in this report should be interpreted with caution.

Methods

A limited literature search was conducted on key resources including Ovid Medline, PubMed, The Cochrane Library, University of York Centre for Reviews and Dissemination (CRD) databases and a focused Internet search. No methodological filters were applied to limit retrieval by publication type. The search was limited to English language documents published between January 1, 2012 and November 15, 2017

Literature Search Methods

Rapid Response reports are organized so that the evidence for each research question is presented separately.

Selection Criteria and Methods

One reviewer screened citations and selected studies. In the first level of screening, titles and abstracts were reviewed and potentially relevant articles were retrieved and assessed for inclusion. The final selection of full-text articles was based on the inclusion criteria presented in Table 1.

Table 1: Selection Criteria

Population	Adult patients following total laryngectomy
Intervention	Indwelling voice prostheses (e.g., Provox Vega®, Blom-Singer Dual Valve®, Provox 2®, Blom-Singer Classic®)
Comparator	Q1-Q2: Compared to each other Q3: No comparator
Outcomes	Q1: Clinical effectiveness (e.g., quality of life, patient perceptions, mental health benefits [i.e., depression, anxiety, self-esteem changes], device lifespan) and safety (i.e., patient harms) Q2: Cost-effectiveness Q3: Evidence-based guidelines
Study Designs	Health technology assessments, systematic reviews, meta-analyses, randomized controlled trials, non-randomized studies, economic evaluations, evidence-based guidelines

Exclusion Criteria

Articles were excluded if they did not meet the selection criteria outlined in Table 1, they were duplicate publications, or were published prior to January 1, 2012. Articles were also excluded if the population of interest included less than 50% laryngectomy patients, with no separate subgroup analyses, or exclusively included pharyngolaryngectomy patients. Articles were excluded if the number of devices used with patients exceeded 50% non-indwelling devices, the data for indwelling devices was not separated, or if no comparative data between indwelling devices was presented. Studies were not excluded based on whether the included patients received primary or secondary tracheoesophageal puncture, and results for these populations were presented separately, where possible.

Critical Appraisal of Individual Studies

The included randomized and non-randomized studies were critically appraised using the Down's and Black Checklist.[9] Summary scores were not calculated for the included studies; rather, a review of the strengths and limitations of each included study were described. Details of the critical appraisal of the included studies are presented in Appendix 3.

Summary of Evidence

Quantity of Research Available

A total of 286 citations were identified in the literature search. Following screening of titles and abstracts, 264 citations were excluded and 22 potentially relevant reports from the electronic search were retrieved for full-text review. Three potentially relevant publications were retrieved from the grey literature search. Of these potentially relevant articles, 20 publications were excluded for various reasons, while 5 publications met the inclusion criteria and were included in this report. Appendix 1 describes the PRISMA flowchart of the study selection.

Additional references of potential interest are provided in Appendix 5.

Summary of Study Characteristics

Additional details regarding the characteristics of included publications are provided in Appendix 2.

Study Design

One study was a RCT with a crossover design.[7] The remaining studies were non-randomized, with two prospective studies[10,11] and two retrospective studies (both using data from electronic medical records and databases).[8,12] No relevant health technology assessments, systematic reviews, meta-analyses, economic evaluations, or evidence-based guidelines were identified.

Country of Origin

Three studies[8,10,12] were conducted in the United States (US), one study[7] was conducted in Australia, and one study[11] was conducted in Germany.

Patient Population

The patient population for all studies included adults who had undergone a total laryngectomy and received either a primary or secondary transesophageal puncture (TEP).[7,8,10-12] Two studies[8,11] also included individuals who had received a pharyngectomy (full or partial) or a pharyngolaryngectomy (which are not within the scope of this report), but these were fewer in total than the laryngectomy patients.

The age range of participants was 34 to 92 years and the mean age of participants in each study was between 62.0 and 64.4.[7,8,10-12] The study that reported the largest number of total patients was Lewin et al.[8] (390 patients, with 296 laryngectomy specific) and the smallest number of total patients was reported in Brownlee et al.[10] (14 patients). The remaining studies reported a total of 15 patients (Hancock et al.),[7] 102 patients (75 laryngectomy specific in Kress et al.),[11] and 21 patients (Thylur et al.).[12]

All included studies had at least twice as many male participants in comparison to female participants (ranging from a ratio of 2.5 to 8.2), likely due to the higher rates of laryngeal cancer in men compared with women.[1,7,8,10-12] Patients who had received primary TEP were more prevalent in the included studies in this report,[7,8,10] with the exception of Thylur et al., in which secondary TEP was more prevalent.[12] Kress et al.[11] did not record the type of TEP surgery that their participants received. The included populations were a mix of both individuals who had received radiation, and those who had not received radiation.[7,8,10-12]

Interventions and Comparators

All participants in all studies received indwelling prostheses, with the exception of one study, in which some participants had received non-indwelling prostheses.[8] Comparisons or device lifespans involving these non-indwelling prostheses are not included in this report.

Prostheses used in these studies included Provox 2® ,[8,11,12] Provox Vega® ,[7,8,11,12] Provox Activalve® ,[8,11] Blom-Singer Classic® ,[7,8,11] Blom-Singer Dual Valve® ,[10,11] Blom-Singer Single Valve,[10] Blom-Singer Increased Resistance® ,[8] Blom-Singer Standard Enlarged Flange® ,[8] and Blom-Singer Advantage® .[8]

Outcomes

Device lifespan was the main outcome in four of the five included studies, measured in days.[8,10-12] Lifespan was measured as time from insertion of the prostheses to time of failure of the prostheses (and subsequent removal or replacement). Failure of the prostheses was specifically device-related failure (e.g., leakage through the valve, high-pressure speech due to valve obstruction, or biofilm growth on the outside of the shaft) in two studies.[11,12] In one study,[10] device failure was defined as when the patient themselves determined that the valve has failed, and then scheduled an appointment to have it replaced. In Lewin et al.,[8] prostheses failure was not explicitly defined, but included routine replacement and leakage of the prosthesis.

Hancock et al.[7] recorded patient preferences and perceptions regarding two different devices as their primary outcome. These preferences were recorded in three structured questionnaires; the first was a voice prosthesis questionnaire regarding bloating, leaking, inadvertent valve opening, voicing, cleaning and maintenance issues, the second was an insertion questionnaire regarding the insertion process, and the final questionnaire was a comparative questionnaire. The voice prosthesis questionnaire and the insertion questionnaire were rated on a four-point scale (where 1 equated to 'none' and 4 equated to

'very much')., and the final comparative questionnaire was a rating of preference, in which participants were asked which device they preferred overall, and could answer the 'first device', the 'second device' or 'no preference'.[7]

Perceived speech quality and phonation effort was a secondary outcome in one study.[10]

No economic evaluations or evidence-based guidelines were identified; therefore, no outcomes related to research questions 2 or 3 were obtained. One study,[10] however, did an additional analysis on cost differences between the Blom-Singer Single Valve and the Blom-Singer Dual Valve® using the mean and median number of extrapolated annual visits per patient and the estimated average cost per device replacement.

Summary of Critical Appraisal

All of the included studies were appraised using the Down's and Black Checklist.[9]

Randomized Controlled Trial (RCT)

One RCT meeting the inclusion and exclusion criteria was identified and included in this report.[7] The authors of the study clearly stated the main hypothesis and aim of the study, outlined patient characteristics clearly in a table format, had proper randomization of initial treatment using permuted blocks of 16, and used comparable and clearly described prostheses for all patients. The trial was also a crossover design, allowing patients to act as their own controls, which helps to eliminate biases related to confounding. Additionally, the authors analysed whether there was a decision bias (i.e., patients would automatically prefer the first device they received) during the trial, and found no evidence of order effect. The authors also performed an analysis of patients lost to follow-up and found drop-out was unlikely to be related to the intervention.

In the RCT,[7] however, patients were automatically excluded if they lived remotely, therefore potential groups of patients who live in more remote communities were not represented in the study. Additionally, the clinicians inserting the prostheses had extensive experience related to one of the prostheses (the Blom-Singer Classic®) and very little or no experience in insertion of the comparator device (Provox Vega®). For insertion of the Blom-Singer device, dilation was mandatory (as per institutional protocol), but for the Provox Vega® device, it was solely the decision of the clinician whether to use dilation or not. This resulted in all of the Blom-Singer devices being inserted with dilation, but half of the Provox Vega® devices being inserted without dilation. Thus, it is unclear whether the questions that asked about insertion discomfort or pain may have been potentially skewed towards favouring one device over the other.

The questionnaires used in the study were not validated or standardized and therefore it is unknown whether questions effectively captured the correct outcomes. Additionally, in the results, some of the questions were collapsed into one measure instead of being presented separately. Out of these individual questions that were collapsed into one question, the most negative answer obtained on any question was taken as the overall score in the category, regardless of the other answers. This could potentially remove the nuances of some participants' answers. For example, if the patient thought the entrance of the prosthesis was very easy to find, but overall found the prosthesis difficult to clean, this would have been registered solely as a negative value (e.g., 'hard to clean'), regardless of the positive answer to the first question. With the admission by the authors that there was diversity in answers from participants, this may have been an issue. The questionnaires were conducted by a clinician who was separate from the insertion process and had no

knowledge of which prostheses was received, which eliminates potential of measurement bias in the outcomes.[7]

Non-Randomized Studies

Four non-randomized studies meeting the inclusion and exclusion criteria were identified and included in this report.[8,10-12]

All included studies adequately described the hypothesis or aim of the study.[8,10-12] The authors of each study clearly described the interventions used and what brand of prostheses were included in the study. No conflicts of interest were disclosed.[8,10-12]

Brownlee et al.[10] included a study design that allowed patients to be compared to themselves, but did not specify which type of t-test was performed, therefore it is unknown if the chosen statistical test was appropriate for the data collected. In Kress et al.,[11] the analysis method used was not appropriate for the data collected as they used a Mann-Whitney test that assumes independence of samples with non-independent samples.

The authors of Brownlee et al.[10] also did not restrict on the inclusion of brands or sizes of single-valve prosthesis, increasing the heterogeneity of interventions in the single valve group. The intervention and comparator groups were therefore not comparable to each other, as the dual valve comparator group was more restrictive to one specific brand, whilst the single valve group may have contained a variety of difference brands of prosthesis. The data collected from the single valve group was also a mean duration collected over three valve replacements, and the dual valve was a mean of one replacement or two replacements. Since there was clear variability in the lengths of the valve life, relying on one measurement for dual valve, and the means of three measurements for single valve may not properly reflect the correct valve lifetimes. There was also inherent self-reporting bias present in the analysis, as patients decided themselves if the valve had failed, and this was not objectively measured.

The patient selection method employed by Kress et al.[11] may have impacted the overall study findings. The study institution has an algorithm to determine which prostheses a patient receives. All patients receive a standard prosthesis (Provox 2®, Provox Vega® and Blom-Singer Classic®) to begin with, and if this device did not demonstrate a life span of less than six weeks, the patient then receives a different prosthesis (Blom-Singer Dual Valve®, hard valve assembly). If then, the prosthesis does not last long enough, they are fitted with Provox Activalve®. Because the authors followed this algorithm, patients in the standard prosthesis group represented a "general" laryngectomy population, whilst the Provox Activalve® patients consisted of individuals for whom the previous prostheses did not work. These differences between the populations in the intervention groups were not controlled for in the analysis. The inclusion of multiple devices in the same patients also created an overrepresentation of devices with a short lifespan in the standard prostheses groups, especially with Provox 2® (as this was the only prostheses available to patients in the beginning of the study).

Individuals included in Brownlee et al.[10] are also not comparable to the general laryngectomy patient population, as they only included patients who had short-term failure (device failure within 3 months of placement) of their prostheses. The generalizability of findings in Thylur et al.,[12] is also unclear as only patients who received devices with a diameter of 22.5 French units (Fr) were included. However, Thylur et al.[12] controlled for potential confounding variables such as diabetes, whilst none of the other studies controlled for confounders. Thylur et al.[12] were not able to adjust for confounders such as dietary

differences, hygiene habits, and differences in reimbursement as these factors were not available in their databases used for the retrospective study.[12]

Summary of Findings

A summary of findings of the included studies is presented in Appendix 4.

What is the clinical effectiveness of indwelling voice prostheses for adults following laryngectomy?

Four studies evaluated the longevity of various devices in laryngectomy patients.[8,10-12]

Dual Valve versus Single Valve Prostheses

Brownlee et al.[10] examined dual valve prostheses versus single valve prostheses in patients who have experienced short-term failure of their voice prosthesis, finding that the dual valve prostheses had a mean duration of improvement of 104 days (P = 0.0169) and a median duration of improvement of 33 days (P = 0.0131) when compared to the single valve prostheses. After switching to the dual valve prosthesis, 86% of patients saw improvements in valve life.[10]

Mean and Median Device Lifespan

Provox brand prostheses (Provox 2®, Provox Vega®, and Provox Activalve®) had a mean device lifespan ranging from 53 days[8] to 298 days.[11] When stratified by device type, Provox 2® had a mean device life ranging from 98 days[11] to 115.6 days,[12] Provox Vega® had a mean device life ranging from 53 days[8] to 107 days,[11] and Provox Activalve® had a mean device life ranging from 192 days[8] to 298 days.[11]

Blom-Singer brand prostheses (Classic, Dual Valve, Advantage, and Enlarged Flange) had a mean device lifespan ranging from 86 days[11] to 164 days.[10] When stratified by device type, Blom-Singer Classic® had a mean device life of 86 days in two studies,[8,11] Blom-Singer Dual Valve® had a mean device life ranging from 104 days[11] to 164 days,[10] and Blom-Singer Advantage® had a mean device life of 97 days in one study.[8] Also, Blom-Singer Standard Enlarged Flange® had a mean device lifespan of 71 days in one study.[8]

Many studies also reported median values for device lifespan. Provox brand prostheses had a median device lifespan ranging from 45 days[8] to 291 days.[11] When separated by device type, Provox 2® had a median device life ranging from 66 days[11] to 77 days,[8] Provox Vega® had a median device life ranging from 45 days[8] to 92 days,[11] and Provox Activalve® had a median device life ranging from 161 days[8] to 291 days.[11]

Blom-Singer brand prostheses (Classic, Dual Valve, Advantage, and Enlarged Flange) had a median device lifespan ranging from 42 days[8] to 89 days.[11] When separated by device type, Blom-Singer Classic® had a median device life ranging from 59 days[8] to 89 days,[11] Blom-Singer Dual Valve® had a median device life ranging from 75 days[11] to 84 days,[10] and Blom-Singer Advantage® had a median device life of 67 in one study.[8] Also, Blom-Singer Standard Enlarged Flange® had a median device lifespan of 42 days in one study.[8]

Comparison between Device Brands

The majority of studies performed a statistical analysis comparing the lifetimes of devices, but not all comparisons between all device types were available.

In Kress et al.,[11] Provox 2® was compared with Blom-Singer Classic®, Blom-Singer Dual Valve®, and Provox Vega®. There was no statistically significant difference in device lifespan between Provox 2® and Blom-Singer Classic® or Provox 2® and Blom-Singer Dual Valve®.

Provox 2® appeared to have significantly shorter lifespan than Provox Vega®, in both Mann-Whitney tests and Kaplan-Meier survival curve log rank tests at 6 months (P = 0.006 and P =0.024 respectively).[11] At one year of measurement, the Kaplan-Meier survival curve log rank test showed no statistically significant difference between Provox 2® and Provox Vega®.[11] In another study by Thylur et al,[12] Provox 2® was compared with Provox Vega®. Provox 2® had a longer time interval between prosthesis replacements (F = 31.9, P < 0.001) and Provox 2® had a longer device lifespan than Provox Vega® (P = 0.0001), which opposes the results found in Kress et al.[11]

Provox Vega® was also compared to Blom-Singer Classic® and Blom-Singer Dual Valve®, with no significant differences found between the Provox Vega® and Dual Valve prostheses.[11] However, Provox Vega® was found to have a longer device lifespan than Blom-Singer Classic® (P = 0.004), and at one year of measurement, the Kaplan-Meier log rank tests found significant differences in device lifespan between Provox Vega® and Blom-Singer Classic®, favouring Provox Vega® (P = 0.043). Blom-Singer Classic® and Blom-Singer Dual Valve® were found to not significantly differ from one another in device lifetime.[11] Provox Activalve® had the longest device lifespan in Kress et al.,[11] lasting significantly longer than all other groups combined (P < 0.001).

Device life was found to not significantly differ between devices that were regular indwelling prostheses (Blom-Singer Classic®, Provox 2®, and Provox Vega® grouped together) and those that were specialty indwelling prostheses (Blom-Singer Indwelling Increased Resistance, Blom-Singer Standard Enlarged Flange®, Blom-Singer Advantage®, and Provox Activalve® grouped together)(P = 0.45).[8] In this study,[8] the mean and median values for each device brand were provided, but no comparative statistical analyses were done between individual device brands.

Patient Preferences and Perspectives

Two studies[7,10] evaluated patient preferences or perspectives in regards to indwelling voice prostheses. The RCT authored by Hancock et al.[7] reported that patients had higher discomfort in insertion of the Provox Vega® prosthesis compared with the Blom-Singer Classic® (P = 0.003). However, during insertion there was no difference in pain levels or in extent of coughing between the two prostheses.[7]

For cleaning and care of the devices, there was no significant difference in cleaning frequency between the Provox Vega® and Blom-Singer Classic® prostheses but Provox Vega® was rated by participants as being significantly easier to clean (P = 0.001) and cleaning was significantly more effective (P = 0.008).[7]

Voicing effort was not perceived at being significantly different between Provox Vega® and Blom-Singer Classic®.[7] However, patients reported using less effort to speak with the Provox Vega®, and reported having a better overall voice with Provox Vega® (P = 0.05 and P = 0.001 respectively).

Patients had significantly higher perceptions of intermittent leakage with the Blom-Singer Classic® prosthesis (P = 0.011), but there were no significant differences in perceptions of

inadvertent valve opening.[7] Overall, a significantly higher proportion of patients in this study group preferred the Provox Vega® prosthesis over the Blom-Singer prosthesis (P = 0.019).[7]

In another study,[10] 86% of patients reported that phonation effort did not differ between single-valve and dual-valve prostheses, but 14% reported that phonation was more difficult with a dual-valve prosthesis. Speech quality did not change between the two prostheses in this group of patients.[10]

What is the cost-effectiveness of indwelling voice prostheses for adults following laryngectomy?

No cost effectiveness data was identified regarding indwelling voice prostheses for adults following laryngectomy. One study,[10] however, did an analysis in the discussion on cost differences between the Blom-Singer Single Valve® and the Blom-Singer Dual Valve® using the mean and median number of extrapolated annual visits per patient and the estimated average cost per device replacement (including professional fees). Both the mean and median costs per year for the dual valve prostheses ($6,301 and $3,564 USD) were lower than the mean and median costs for the single valve prostheses ($10,627.75 and $5867.19 USD). These costs, however, were from one specific institution, and therefore should be interpreted with caution.

What are the evidence-based guidelines regarding the use of indwelling voice prostheses for adults following laryngectomy?

No evidence-based guidelines were identified regarding the use of indwelling voice prostheses for adults following laryngectomy.

Limitations

One of the major limitations of the body of evidence identified by this report is the lack of Canadian based studies, and therefore, limited generalizability to the Canadian context. Variations in region-specific dietary habits, the microflora within a patient, and physician reimbursement are contributing factors to device replacement.[12] It is unclear whether regional differences would have a large impact on device lifespan results.

A second limitation of the evidence is the lack of economic evaluations that focus on cost-effectiveness outcomes and the lack of identified evidence-based guidelines for the use of indwelling voice prostheses.

There are some limitations with the quantity of evidence pertaining to newer voice prostheses, for example, the Blom-Singer Advantage® prosthesis or Provox Vega® XtraSeal. There was also no mention or differentiation between the sterile series and the non-sterile series of voice prostheses. There was also limited evidence with regards to heat-moisture exchange (HME) filters, as no studies recorded whether a patient used HME filters or not. Voice prostheses such as Provox 2® have been available in institutions for a much longer period of time (available since 1997) than newer prostheses, such as Provox Vega® (available since 2009[1]), and uptake of these newer prostheses is often slow (for example, in Thylur et al.,[12] where uptake of the Provox Vega® did not occur in their institution until 2014). Therefore the majority of the evidence is focused on these older prostheses, and the clinicians often have more experience inserting and evaluating these devices. There was also a lack of direct comparisons for some prostheses types, making decisions relating to the comparative effectiveness of these brands difficult.

The heterogeneous nature of the general population of laryngectomy patients also limits the generalizability of this report. None of the studies separated patients out by radiation treatment or reasoning for the laryngectomy. Radiation may have an effect on device lifespan,[2,8] and it appears that prosthesis choice is a very personal and individual decision. Not having specific information on these various groups limits the evidence for decision making, and forces a clinician to make a decision for one individual patient based on the laryngectomy population as a whole. Hancock et al.[7] found a very diverse selection of patient opinions and preferences for devices, and suggests that patient opinion and preferences should be part of the decision making process for prosthesis selection.

Conclusions and Implications for Decision or Policy Making

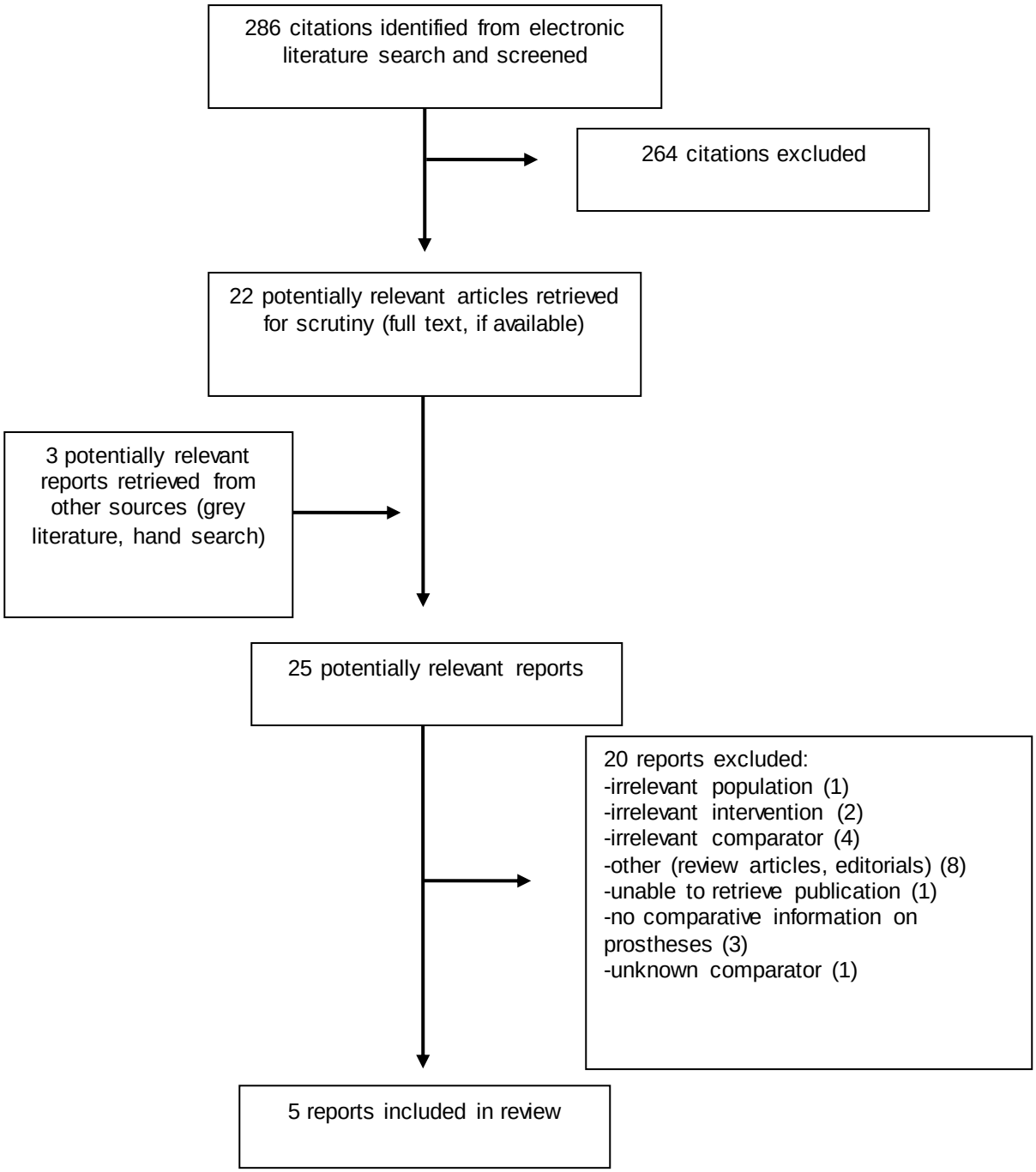
There are a variety of brands and types of indwelling prostheses available for laryngectomy patients. Device lifespan appears to vary between different study populations and different types of prostheses. For example, Provox 2® appears to have a longer lifespan than Provox Vega® in one population,[11] but in a different population, this pattern is reversed.[12] More specialized prostheses, such as Provox Activalve® appear to have a longer lifespan than standard prostheses, such as Provox 2®, Provox Vega® and Blom-Singer Classic®. However, the quality of this evidence is very limited, as there are many limitations to these comparative studies, including issues surrounding selection of patient groups, inappropriate statistical methods, lack of controlling for confounders, small sample sizes, and self-reporting bias. There are also many relevant comparisons that are not available in the current literature included in this report, therefore much of the overall comparative evidence regarding the lifespans of indwelling prostheses are still unknown. Finally, there are only primary research studies available in the literature, and no health technology assessments, systematic reviews or meta-analyses provide access to higher quality or larger patient populations. Therefore, the results of these studies should be interpreted with caution.

Patient preference data is also limited, and was exclusive to the Provox Vega® and the Blom-Singer Classic® voice prostheses. Patients appeared to prefer the Provox Vega® device overall, but the authors received a diverse selection of opinions regarding preferred prostheses types. The authors therefore concluded that patient perspectives should be included in the clinician decision making process and reiterate the importance of considering every appropriate device across all areas.[7] No cost-effectiveness studies or evidence-based guidelines were identified; therefore no conclusions can be made regarding these outcomes.

References

1. Lorenz KJ. Rehabilitation after total laryngectomy-a tribute to the pioneers of voice restoration in the last two centuries. Front Med (Lausanne) [Internet]. 2017 [cited 2017 Nov 16];4:81. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5483444/pdf/fmed-04-00081.pdf
2. Brook I. The laryngectomee guide [Internet]. Alexandria (VA): American Academy of Otolaryngology - Head and Neck Surgery; 2013. [cited 2017 Nov 17]. Available from: http://www.entnet.org/sites/default/files/LaryngectomeeGuide.pdf
3. Summers L. Social and quality of life impact using a voice prosthesis after laryngectomy. Curr Opin Otolaryngol Head Neck Surg. 2017 Jun;25(3):188-94.
4. Tang CG, Sinclair CF. Voice restoration after total laryngectomy. Otolaryngol Clin North Am. 2015 Aug;48(4):687-702.
5. Lombard LE. Laryngectomy rehabilitation. In: Medscape [Internet]. New York (NY): Medscape LLC; 2014 [cited 2017 Nov 17]. Available from: https://emedicine.medscape.com/article/883689-overview
6. Voice prosthesis for voice rehabilitation following total: policy. Number: 0560 [Internet]. Hartford (CT): Aetna Inc.; 2017. [cited 2017 Nov 17]. Available from: http://www.aetna.com/cpb/medical/data/500_599/0560.html
7. Hancock K, Ward E, Lawson N, van As-Brooks CJ. A prospective, randomized comparative study of patient perceptions and preferences of two types of indwelling voice prostheses. Int J Lang Commun Disord. 2012 May;47(3):300-9.
8. Lewin JS, Baumgart LM, Barrow MP, Hutcheson KA. Device life of the tracheoesophageal voice prosthesis revisited. JAMA Otolaryngol Head Neck Surg. 2017 Jan 1;143(1):65-71.
9. Downs SH, Black N. The feasibility of creating a checklist for the assessment of the methodological quality both of randomised and non-randomised studies of health care interventions. J Epidemiol Community Health [Internet]. 1998 Jun;52(6):377-84. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1756728/pdf/v052p00377.pdf
10. Brownlee B, Ahmad S, Grammer T, Krempl G. Selective patient experience with the Blom-Singer Dual Valve voice prosthesis. Laryngoscope. 2017 Aug 7.
11. Kress P, Schafer P, Schwerdtfeger FP, Rosler S. Are modern voice prostheses better? A lifetime comparison of 749 voice prostheses. Eur Arch Otorhinolaryngol [Internet]. 2014 Jan [cited 2017 Nov 16];271(1):133-40. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3889693/pdf/405_2013_Article_2611.pdf
12. Thylur DS, Villegas BC, Fisher LM, Sinha UK, Kokot N. Device life of two generations of Provox voice prostheses. Ann Otol Rhinol Laryngol. 2016 Jun;125(6):501-7.

Appendix 1: Selection of Included Studies



Appendix 2: Characteristics of Included Publications

Table 2A: Characteristics of Included Clinical Studies

First Author, Publication Year, Country	Study Design	Population	Intervention(s)*	Comparator(s)*	Outcome(s), Length of Follow-up
Hancock 2012[7] **Australia**	RCT Crossover design	31 laryngectomy patients (25 male; 6 female) aged 34 to 89 years (mean age 64.3) 17 patients had TL 14 patients had TPL 30 had primary TEP, 1 had secondary TEP	Provox Vega® (n=15)	Blom-Singer Classic® Indwelling (n=16)	Patient perceptions and preferences Length of follow-up was 7 months to 19 years (mean 5.1 years, SD 4.2 years)
Brownlee 2017[10] **US**	Prospective phase IV non-randomized clinical study	14 TL patients (10 male; 4 female) aged 45 to 81 years (mean age 64) who had previously experienced early (within 3 months of valve placement) valve failure when using a single-valve voice prosthesis 10 with a primary TEP, 2 with secondary TEP, 2 with pectoralis flap	20-Fr Blom-Singer Dual Valve® Voice prosthesis (n=17*) *3 patients received 2 dual valve prostheses due to valve failure	Blom-Singer Single Valve Voice prosthesis (n=40*) *3 single valves were used per patient prior to use of dual valve	Device lifespan (duration) and valve life Perceived speech quality (patient-rated) and phonation effort Length of follow-up was NR
Kress 2014[11] **Germany**	Non-randomized prospective cohort study	102 laryngectomy patients (91 male; 11 female) aged 42 to 86 years (median age of 64.4 for females and 61.2 for males) 75 had laryngectomy, 18 had pharyngolaryngectomy and 9 had unknown surgery	Provox Vega® (n=117) Provox 2® (n=424) Provox Activalve® (n=38)	Blom-Singer Classic® Indwelling 20 Fr (n=108) Blom-Singer Dual Valve 20 Fr (n=62)	Device lifespan Length of follow-up was 3 months to 23 years (median of 6.8 years)
Lewin 2017[8] **US**	Non-randomized retrospective study	390 laryngectomy patients (317 male; 73 female) aged 34 to 92 (mean age 62; median age 62) 296 had TL, 32 had TPL and 62 had TL plus a partial pharyngectomy	Provox Vega® (n=44) Provox 2® (n=1096) Provox Activalve® (n=40)	Blom-Singer Indwelling (n=1383) Blom-Singer Indwelling - Increased Resistance (n=NR) Blom-Singer	Device lifespan Length of follow-up was NR

Table 2A: Characteristics of Included Clinical Studies

First Author, Publication Year, Country	Study Design	Population	Intervention(s)*	Comparator(s)*	Outcome(s), Length of Follow-up
		248 patients had primary TEP and 142 patients had secondary TEP		Standard Enlarged Flange® (n=205) Blom-Singer Advantage® (n=251)	
Thylur 2016[12] **US**	Non-randomized retrospective observational study	21 patients (17 male; 4 female) laryngectomy patients aged 49 to 80 (mean age 64) 2 had primary TEP and 19 had secondary TEP	22.5 Fr Provox 2® (n=NR)	22.5 Fr Provox Vega® (n=NR*) *181 total device replacements (including Provox Vega® and Provox 2®)	Device lifespan Number of prostheses changes per patient Length of follow-up of 1 to 5 years (mean of 2.3 years)

Fr = French (French catheter scale); NR = not reported; RCT = randomized controlled trial; TEP = tracheoesophageal puncture; TL = total laryngectomy; TPL = total pharyngolaryngectomy; SD = standard deviation

*Some studies included multiple prostheses inserted into one patient, therefore the number of included prostheses do not match the number of patients included in the study

Appendix 3: Critical Appraisal of Included Publications

Table 3A: Strengths and Limitations of Randomized Controlled Trials using the Down's and Black Checklist[9]

Strengths	Limitations
Hancock, 2012[7]	
• The hypothesis and aim of the study is clearly stated • Ethics statement provided and protocol submitted prior to start of trial under the Clinical Trials Notification Scheme • The main outcomes are clearly stated, with the order in which they were asked outlined • Characteristics of the included patient group are clearly stated, with the demographic data displayed in a table format • Interventions are clearly described, and are comparable in size, length, and device type (indwelling prosthesis) • Patients require larger or smaller prostheses were given the equivalent prosthesis size in the crossover phase (i.e., Provox Vega® 17 Fr vs. Blom-Singer Classic® 16 Fr, and the Provox Vega® 20 Fr vs. Blom-Singer Classic® 16 20) • Crossover design allowed patients to be their own controls, eliminating some potential confounders • Randomization done using permuted blocks of 16 • Raw numbers for the main findings clearly stated • Reasoning for loss to follow-up described for the two patients lost in the Provox Vega® trial, and loss to follow-up analyzed using the Clinical Trials Notification Scheme (deemed to be unlikely to be related to the intervention) • Population likely to be representative of the general laryngectomy population (in dwelling prostheses are often inserted by clinicians in similar hospitals) • A separate clinician performed the interviews/administered the questionnaires, with no knowledge of the prosthesis received and no involvement in the insertion • Blinding was not possible for the patients, but brand names were removed and devices referred to as 'the current device' or 'the previous device' to prevent bias • Decision bias explored (i.e., patient would prefer the first device they received), with no major order effect found and prior device familiarity was not a factor • Analysis was intent-to-treat, maintaining the random assignment • Actual probability values were stated for all statistical analyses • Compliance likely not an issue as indwelling prostheses are non-removable and non-compliance results in inability to speak, discomfort, and coughing • No conflict of interest declared	• Characteristics of patients lost to follow-up in Provox Vega® trial not described • Patients were excluded if they lived remotely, (if it prohibited study attendance), although more feasible, this exclusion eliminated participants living further away or in remoter communities • No reasons for declination of participation for the seven eligible participants, which was a large proportion (18%) of the eligible study size. It is unknown if these participants were fundamentally different than individuals who did agree to participate • Clinicians inserting the prostheses had extensive experience in insertion of the Blom-Singer prosthesis, but comparatively less experience in the insertion of the Provox Vega® prosthesis • Dilation during insertion was optional for the Provox Vega® insertion, but required for the Blom-Singer prosthesis (as stated in the standard clinical practice), allowing the un-blinded clinician a choice to use dilation in one insertion but not the other insertion. All patients were previously used to dilation in prostheses changes • Questionnaires used were created for this study and were not previously validated or standardized for this population or intervention, limiting validity • Some groups of questions were collapsed into one measure instead of being presented separately (e.g. "ease of cleaning" was two separate questions combined into one measure), using the most negative score received on the individual questions. This may eliminate important information (e.g. patient may have found stoma entrance hard to find, but once found, the prosthesis was easy to clean, but the score was registered as "difficult to clean" altogether). • No power calculation or reference to sample size calculations

Table 3B: Strengths and Limitations of Non-Randomized Studies using the Down's and Black Checklist[9]

Strengths	Limitations
Brownlee, 2017[10]	
• The hypothesis and aim of the study is clearly stated • Study plan submitted prior to start of study to institutional review board for the University of Oklahoma Health Sciences Center • • Study design allowed patients to be compared to themselves (they were their own controls), eliminating some potential confounders • Valves were changed during routine clinical visits, which is representative of the experience of the majority of laryngectomy patients • Although some more demographic details could have been included, the major demographic details were displayed in table format • No conflict of interest declared	• There was no explanation as to why some eligible participants refused to participate or what proportion of the eligible population did participate • Previous valves used were all different brands and sizes and may not be directly comparable to the Dual valve prosthesis, nor to the single valves used in the other patients. Combining these brands into one group of single valve may create a very heterogeneous comparator group • There was no specific adverse event data gathered, and no explanation as to why some patients who had dual valve prosthesis purposefully switched back to single valve prostheses after the dual valve failed • Data from single valve prostheses was a mean duration from 3 valve replacements, whilst dual valve durations were based off of mostly one, and occasionally two, valve durations. There was considerable variability in the number of days a single valve lasted in one patient (e.g., one patient had one valve last 9 days, but another valve in the same patient last 74 days), so it's likely there may be variability in the dual valve prosthesis that is not captured in the results. • There is inherent self-reporting bias as valve failure was determined by the patients, and not by a clinician or an objective measurement • One confounder was mentioned, but it was not controlled for in the statistical analysis • No statistical analyses were performed on the speech quality or phonation effort outcomes, so statistical significance of this data is unknown • The population in this study may not be representative of the general laryngectomy population, as the sample was only individual with short-term single valve failure • Conclusions were made about alleviation of gastric symptoms when this was not a collected outcome *a priori* and was only volunteered as extra information by a very small number of patients • No information on what type of t-test (i.e., paired or independent) was used; therefore, it is unknown if the chosen test was appropriate for the analysis • No power calculation or reference to sample size calculations
Kress, 2014[11]	
• The hypothesis and aim of the study is clearly stated • Had comprehensive description of each intervention with pictures of each used device • References to an *a prioiri* protocol • The authors removed Provox Activalve® and did a separate analysis without this group	• Study institution used an algorithm to determine what prosthesis each patient receives (i.e., they receive a standard prosthesis [Provox 2 ®, Provox Vega® or Blom-Singer Classic®], and after specific failure on this device, receive a special prosthesis [Blom-Singer Dual Valve® prosthesis], and upon failure of this one, receive Provox Activalve®). This means that all individuals in the Provox

• Authors used a Kaplan-Meier survival analysis to account for differing follow-up times • All prostheses were of a similar size (20 Fr to 22.5 Fr) • Characteristics of the included patient group are clearly stated, with the demographic data displayed in a table format • Actual probability values were stated for all statistical analyses • All participants were recruited from one outpatient centre in Germany, which may not be fully generalizable, but allows for a more homogenous population • No conflict of interest declared	Activalve® group are patients for whom the previous prostheses did not work for (or who are more suited to the specialty prosthesis), and therefore are a different population group. They are not comparable to the patients who only received the standard prostheses, and this may introduce selection bias in the analysis • Because multiple devices in the same patients are included in the analysis, there is an overrepresentation of devices that had a shorter lifespan over the study period, and an overrepresentation of devices with a shorter lifespan in the standard groups (especially Provox 2® , as this was the only Provox prosthesis available at the beginning of the study) • Standard devices were chosen on basis of the clinician's opinion and on characteristics of each patient, which may have introduced selection bias (i.e., are there some characteristics of a patient that would preclude them to receive one over another, and would this then affect device lifespan), and these were not accounted for in the analysis. For example, Blom-Singer Classic® are often used for puncture complication management, so individuals with this prosthesis were more likely to have related issues, and this was not accounted for in the analysis • Analysis did not address confounders, or list them in the study design or methods • 17 different physicians with different levels of experience both decided on choice of prostheses, and inserted them. This could create operator related confounding • Compared device groups had different devices in the same patients counted in the analysis, but the analysis test used is a Mann-Whitney U test, which assumes independence of samples • No power calculation provided
Lewin, 2017[5]	
• The hypothesis and aim of the study is clearly stated • Study plan submitted prior to start of study to institutional review board • Clearly defined exclusion criteria and reasoning for these exclusions • Clear definition of what the intervention was and which devices were eligible • Study subjects in all groups were recruited over the same period of time (2003 to 2013) • Some confounders were listed and included in the analysis (patient sex and age, tumor location, stage of disease, extent of treatment, type of VP, and reason for VP removal) • Characteristics of the included patient group and device group are clearly stated, with the demographic data displayed in a table format • Performed a sensitivity analysis restricted to leakage and showed no change in results • Actual probability values were stated for all statistical analyses • No conflict of interest declared	• Using chart based/electronic record retrospective design misses any individuals who were not recorded in this database • Records with incomplete data were excluded from the analysis, including data incomplete because of non-MDACC replacement data • Clinician choice is a factor in the decision of which prosthesis to get, and may introduce clinician biases into the analysis, and make device groups not be completely comparable populations • Study was conducted in the US, where care is not universally provided, and economic barriers may force patients to choose cheaper prosthesis options (e.g., non-indwelling prostheses over indwelling prostheses), which was not controlled for • Some confounders were missing, including comorbidities such as GERD • No power calculation provided

Thylur, 2016[12]	
• The hypothesis and aim of the study is clearly stated • Study plan submitted prior to start of study to institutional review board Institutional Review Board at the University of Southern California • Clearly defined exclusion criteria and reasoning for these exclusions • All eligible patients decided to use the second prosthesis type, so no patients were lost in the analysis • Confounder data was collected, including for demographics, surgical approach, postoperative complications, treatment, and time between device placement and removal, as well as comorbidities such as GERD and diabetes • Kaplan-Meier survival curves were used to account for differences in follow-up time between patients • Reason for device replacement was controlled for in the analysis • Most of the laryngectomies were performed in the location of the study, so differences in surgery procedures between patients may be lessened • Main findings displayed clearly with actual probability values were stated for all statistical analyses • No conflict of interest declared	• Only patients who received 22.5Fr diameter devices, so the population is limited in generalizability to the general laryngectomy population • Data collection for PPI use, antifungal therapy, and probiotic beverages were based on recall at the end of the trial data collection, which could be subject to recall bias from the patients • Could not adjust for some confounders, such as dietary differences, differences in physician reimbursement, or difference in hygiene habits • No power calculation provided

GERD = gastroesophageal reflux disease; MDACC = MD Anderson Cancer Center; PPI = proton pump inhibitor; RCT = randomized controlled trial; TEP = tracheoesophageal puncture; TL = total laryngectomy; TPL = total pharyngolaryngectomy;

Appendix 4: Main Study Findings and Author's Conclusions

Table 4A: Summary of Findings of Included Studies

Main Study Findings	Author's Conclusion
Hancock, 2012[7]	
Insertion and Removal - Patients had higher degrees of discomfort during Provox Vega® insertion than Blom-Singer Classic® insertion (Z = –2.95, P = 0.003) - Majority of Provox Vega® insertions associated with "a little discomfort" (15 out of 31, 48%) followed by no discomfort (8 out of 31, 26%) - Majority of Blom-Singer insertions associated with no discomfort (16 out of 31, 52%), followed by "a little discomfort" (13 out of 31, 42%) - There were no significant differences between the two prostheses in pain levels (Z = –0.250, P = 0.803), or extent of coughing (Z = –1.6, P = 0.109) **Cleaning and Care** - No significant difference in cleaning frequency for the device (P = 0.281) - Provox Vega® rated significantly easier to clean than Blom-Singer Classic® (P = 0.001) and cleaning rated more effective in Provox Vega® compared with Blom-Singer (P = 0.008) **Voicing** - There was no significant difference in patient perception of voicing effort (z = –1.668, P = 0.095) - Patients reported using less effort to speak when using Provox Vega® (P = 0.05), and reported having a better "overall voice" with Provox Vega® (P = 0.001) **Bloating, Leakage, Valve Opening** - Significantly higher perceptions of intermittent leakage with Blom-Singer Classic® (P = 0.011) - No significant difference in perceptions of inadvertent valve opening (P = 0.225) - Significantly higher perceptions of less bloating with Provox Vega® (P = 0.011) **Overall Preference** - A significantly higher proportion preferred the Provox Vega® prosthesis compared with the Blom-Singer prosthesis (P = 0.019) - Appears to be no significant bias for device familiarity or device order	*"While distinct patterns of patient preference for one device over the other were observed across a number of parameters, a high degree of diversity was observed in the patient responses. This highlights that there were individuals whose experience was either different to, or the complete opposite of, others within the group. Hence, while the current data can be used to some extent to point out certain device-specific features to patients which they may find beneficial, the data equally reveal the heterogeneous nature of this clinical population. The results of this study support that laryngectomized patients should be given the opportunity to trial different devices to help identify the one that works optimally for them." Page 306* *"Participants in this study were able to perceive differences between the two indwelling devices and demonstrated distinct preferences and reasons for using a particular system." Page 308*
Brownlee, 2017[10]	
Device Lifespan *Dual Valve Prosthesis* Mean = 164 days Median (range) = 84 (11 to 431) days	*"For patients whose single valves regularly fail in less than 3 months, our results suggest that the Dual Valve may increase valve life. Complications were expected to be less or the same as those devices previously in use, because nonclinical tests*

Table 4A: Summary of Findings of Included Studies

Main Study Findings	Author's Conclusion
Mean number of replacements per year = 7 Median number of replacements per year (range) = 4 (1 to 33) *Single Valve Prosthesis* Mean = 60 days Median (range) = 51 days (4 to 233 days) Mean number of replacements per year = 13 Median number of replacements per year (range) = 7 (2 to 52) After switching to dual valve, 86% of patients experienced an increase in valve life (43% had an increase > 150 days). Mean duration of improvement = 104 days (P = 0.0169) Median duration of improvement = 33 days (P = 0.0131) **Patient Perceptions** - 86% of patients reported phonation effort was the same with the dual valve compared to single valve, and 14% reported it was more difficult with the dual valve prostheses compared with the single valve - No patients reported changes in speech quality - Three patients noted an alleviation of gastric distention, but this was not solicited information by the authors **Cost information*** Mean extrapolated costs of dual valve prosthesis = $6,301 Median extrapolated costs of dual valve prosthesis = $3,564 Mean extrapolated costs of single valve prosthesis = $10,627.75 Median extrapolated costs of single valve prosthesis = $5,867.19 *cost information was based on the mean and median number of annualized visits per patient, the retail valve cost per patient, and the University of Oklahoma Health Sciences center 2017 professional fees.	*have shown that it is both safe and as effective as other devices" Page 4* *"The Dual Valve is a more expensive device ($388.00) than the single-valve prosthesis, which can cost anywhere from $207.00 to $336.00. In our select population, the cost of single valves ranged from $301.80 to $335.17 (Table IV)…. If valve life is consistent, annualized cost may be decreased in 79% of patients in this select population, with a mean charge reduction of $4,326.75 per year and a median charge reduction of $2,303.19 per year. Reducing the number of annual clinic visits a patient is required to make can also lead to cost savings in other areas such as travel to and from the clinic and lost wages." Page 4*
Kress, 2014[11]	
Device Lifespan, in days Average of all devices: 108 days Median of all devices: 74 days *Provox Vega®* - Median (range) = 92 (3 to 478) - Mean (SD) = 107 (80.6) *Provox 2®* - Median (range) = 66 (1 to 1974) - Mean (SD) = 98 (127.4) *Provox Activalve®* - Median (range) = 291 (5 to 786) - Mean (SD) = 298 (155.8) *Blom-Singer Classic®* - Median (range) = 89 (7 to 397)	*"The prosthesis with the longest dwell time was the Provox ActiValve; this device appeared to have at least three times longer lifetimes compared to the other devices, and its device life time was significantly longer than any of the other standard voice prostheses (P < 0.0001)." Page 137* *"Prostheses with a defined valve opening pressure (Blom-Singer Dual Valve, Provox Vega and ActiValve) had longer lifetimes than prostheses without a defined opening pressure (Blom-Singer Classic and Provox 2)." Page 137*

Table 4A: Summary of Findings of Included Studies

Main Study Findings	Author's Conclusion
- Mean (SD) = 86 (70.6) *Blom-Singer Dual Valve®* - Median (range) = 75 (6 to 387) - Mean (SD) = 104 (88.2) Provox Activalve® vs. all other groups: P < 0.001 Provox 2® + Blom-Singer Classic® vs. Provox Vega® : P < 0.05 Provox 2® vs. Blom-Singer Classic® : NS (P = 0.604) Provox 2® vs. Blom-Singer Dual Valve® : NS (P = 0.233) Provox Vega® vs. Provox 2® : P = 0.006 Provox Vega® vs. Blom-Singer Classic® : P = 0.004 Provox Vega® vs. Blom-Singer Dual Valve® : NS (P = 0.159) Blom-Singer Dual Valve® vs. Blom-Singer Classic® : NS (P = 0.202) **Kaplan-Meier Survival Curves** Logrank test: P <0.001 After removal of Provox Activalve® : - Logrank 1 year: P = 0.181 - Logrank 6 months: P = 0.088 - Logrank 3 months: P = 0.024 Provox Vega® vs. Provox 2® - 1 year: P = 0.133 - 6 months: P = 0.024 - 3 months: P = 0.005 Provox Vega® vs. Blom-Singer Classic® - 1 year: P = 0.043 - 6 months: P = 0.022 - 3 months: P = 0.006	
Lewin, 2017[8]	
Device Lifespan (in days) *Provox Vega®* - Median (range) = 45 (3 to 138) - Mean (SD) = 53 (32) *Provox 2®* - Median (range) = 77 (1 to 764) - Mean (SD) = 100 (84) *Provox Activalve®* - Median (range) = 161 (7 to 567) - Mean (SD) = 192 (166) *Blom-Singer Indwelling* - Median (range) = 59 (1 to 816) - Mean (SD) = 86 (87) *Blom-Singer Indwelling (increased resistance)* - NR *Blom-Singer Standard Enlarged Flange®* - Median (range) = 42 (1 to 469) - Mean (SD) = 71 (83) *Blom-Singer Advantage®* - Median (range) = 67 (1 to 760)	*"Extent of surgery did not significantly affect device life; patients treated with total laryngectomy, total laryngectomy with partial pharyngectomy, and total laryngopharyngectomy had median lifetimes of 62, 57, and 56 days, respectively (P = .22)" Page 67* *"In this large, contemporary laryngectomy cohort of TE prosthesis users, the average device life was roughly 2 months, with minimal effects of treatment history and device type observed." Page 68* *"However, when data were examined based on individual performance, the ActiValve as an individual device offered the longest longevity relative to other VPs, representing a roughly3-month longer duration beyond standard device life. This outcome is not unexpected but rather seems likely given the advanced design of a magnet-driven valve coupled with biofilm-resistant biomaterials." Page 69*

Table 4A: Summary of Findings of Included Studies

Main Study Findings	Author's Conclusion
- Mean (SD) = 97 (112) *Indwelling vs. Specialty Indwelling* - Median (range) = 70 (1 to 816) vs. 61 (1 to 760) - Mean (SD) = 94 (86) vs. 92 (98) - Device life did not significantly differ between specialty and standard prostheses (P = 0.45) **Reason for Removal (median [range]; mean[SD])** *Complications*: 28 (1 to 816); 57 (78) *Leakage*: 64 (1 to 672); 89 (82) *Other*: 61 (1 to 764); 92 (103) Surgery type did not affect device life (P = 0.22) - Total laryngectomy = 62 days - Total laryngectomy with partial pharyngectomy = 57 days - Total laryngopharyngectomy = 56 days	
Thylur, 2016[12]	
Device Lifespan Provox 2® vs. Provox Vega® - Mean device lifespan (SE) ○ 115.6 days (5.8) vs. 65.1 days (7.5) ○ P = 0.0001 - Mean prosthesis changes per patient (SE) ○ 4.5 (0.43) vs. 4.1 (0.75) ○ P = 0.63 - Provox 2® had a longer time interval between replacements than Provox Vega® (F = 31.9, P < 0.001) - Provox Vega® had fewer non-device related reasons for replacement (15%) compared with Provox 2® (27%; X^2 = 4.0, P = 0.046) **Device Replacement** There were 181 indications for device replacement, 95 in Provox 2® and 86 in Provox Vega® - Failure of prosthesis was the most common reason for replacement - Leakage of valve (n=144; 79%) ○ Provox 2® (n=72) ○ Provox Vega® (n=72) - High pressure speech (n=7) ○ Provox 2® (n=5) ○ Provox Vega® (n=2) - Periprosthetic leakage (n=10) ○ Provox 2® (n=8) ○ Provox Vega® (n=2) - Granulation tissue (n=11) ○ Provox 2® (n=5) ○ Provox Vega® (n=6) - No longer *in situ* (n=1) ○ Provox 2® (n=1)	*"Although the Provox Vega may offer superior speech quality and ease of insertion compared to its predecessor… Our study demonstrates that the Provox 2 offers longer device life compared with the Provox Vega, even when controlling for reason for device replacement." Page 505* *"In our study, we found that the Provox Vega was more frequently replaced for prosthesis-related reasons compared with the Provox 2 (P = .046). However, replacement reason did not have a statistically significant effect on device life in mixed regression in our sample." Page 506*

Table 4A: Summary of Findings of Included Studies

Main Study Findings	Author's Conclusion
○ Provox Vega® (n=0) - Size change (n=4) ○ Provox 2® (n=3) ○ Provox Vega® (n=1) - Inflammation/infection (n=1) ○ Provox 2® (n=1) ○ Provox Vega® (n=0) - No effect of replacement reason on the amount of time between replacements (F = 0.86, P = 0.36)	

NR = not reported; NS = not significant; RCT = randomized controlled trial; TEP = tracheoesophageal puncture; TL = total laryngectomy; TPL = total pharyngolaryngectomy; SD = standard deviation; SE = standard error

Appendix 5: Specialty and Standard Indwelling Voice Prostheses

Table 5A: Types of TEP Voice Prostheses

Voice Prosthesis Type or Brand	Specialty or Standard
Blom-Singer Indwelling	Standard
Provox Vega® Indwelling	Standard
Provox 2® Indwelling	Standard
Blom-Singer Indwelling - Increased Resistance Indwelling®	Specialty
Blom-Singer Standard Enlarged Flange® Indwelling	Specialty
Blom-Singer Advantage® Indwelling	Specialty
Provox Activalve® Indwelling	Specialty
Blom-Singer Dual Valve® Indwelling	Specialty

Appendix 6: Additional References of Potential Interest

Qualitative Studies

Impact on Social Function and Quality of Life

1. Summers L. Social and quality of life impact using a voice prosthesis after laryngectomy. Curr Opin Otolaryngol Head Neck Surg. 2017 Jun;25(3):188-94.

Unable to Retrieve Publication

2. Serra A, Spinato G, Spinato R, Conti A, Licciardello L, Di LM, et al. Multicenter prospective crossover study on new prosthetic opportunities in post-laryngectomy voice rehabilitation. J Biol Regul Homeost Agents. 2017 Jul;31(3):803-9.