

**An Economic Analysis of Implantable Doppler Technology in Head and Neck  
Reconstruction**

**Michael Gupta**

Thesis submitted to the Faculty of Graduate and Postdoctoral Studies in partial  
fulfillments of the requirements for the MSc. Degree in Epidemiology

Department of Epidemiology and Community Medicine

Faculty of Medicine

University of Ottawa

Ottawa, Ontario, Canada

© Michael Gupta, Ottawa, Canada, 2012

## **Abstract**

The goal of this thesis was to evaluate the cost-effectiveness of implantable Doppler technology (IDT) used to monitor free tissue transfer (FTT) procedures in the treatment of cancer of the upper aerodigestive tract (UADT).

First, a systematic review of the literature on the effectiveness of traditional and IDT monitoring techniques was performed. Second, a utility survey using a time trade-off technique was created and administered. The results from this survey were used to establish utility values for health states common in patients undergoing FTT procedures. Third, a cost study using the microcosting data available through the Ottawa Hospital was performed. Finally, a decision analytic model was created and an economic evaluation from the payer perspective was completed. A probabilistic sensitivity analysis (PSA) and a value of information analysis (VOI) were performed.

The thesis found that the currently available evidence supports IDT as a cost-effective intervention. Further research should be directed towards determining the effectiveness of both traditional and IDT monitoring.

## **Table of Contents**

<b>1. Introduction.....</b>	<b>1</b>
<b>2. Systematic Review.....</b>	<b>7</b>
2.1. Objective.....	7
2.2. Methods.....	7
2.2.1. Literature Search.....	7
2.2.2. Eligibility Criteria.....	8
2.2.3. Data Extraction.....	9
2.2.4. Analysis.....	9
2.3. Results.....	12
2.3.1. Summary of Literature Search.....	12
2.3.2. Quality Assessment.....	12
2.3.3. Descriptive Analysis.....	12
2.3.4. Pooled Estimate of Proportions for OSR, TR, SR.....	16
2.4. Discussion.....	18
<b>3. Utility Survey.....</b>	<b>27</b>
3.1. Introduction.....	27
3.2. Health States in Free Tissue Transfer in Head and Neck Cancer.....	31
3.3. Creation of a Time Trade-Off Questionnaire.....	34
3.3.1. Utility Survey Population.....	35
3.4. Results of Time Trade-Off Utility Questionnaire.....	35
3.4.1 Response Rates.....	35
3.4.2 Utility Values.....	36

3.5. Discussion.....	37
<b>4. Costing Study.....</b>	<b>40</b>
4.1. Introduction.....	40
4.2. Methods.....	41
4.3. Results.....	43
4.4. Discussion.....	45
<b>5. Health Economic Evaluation.....</b>	<b>49</b>
5.1. Introduction.....	49
5.1.1. Rationale for Health Economic Evaluation.....	49
5.1.2. Decision Models in Health Economic Evaluation.....	51
5.1.3. Discount Rates.....	54
5.1.4. Sensitivity Analysis.....	54
5.1.5. Cost Effectiveness Acceptability Curves and Expected Value of Perfect Information Analysis.....	58
5.2. Methods.....	60
5.2.1. Model Design.....	60
5.2.2. Assignment of Transition Probabilities in the Model.....	63
5.2.3. Assignment of Costs in the Model.....	66
5.2.4. Assignment of Length of Stay in the Model.....	67
5.2.5. Assignment of Utilities in the Model.....	68
5.2.6. Time Horizon.....	68
5.2.7. Discount Rate in the Model.....	68
5.2.8. Assumptions.....	68

5.2.9. Sensitivity Analysis.....	69
5.2.10. Cost-Effectiveness Acceptability Curve.....	71
5.2.11. Estimate of Expected Value of Perfect Information.....	72
5.3. Results.....	72
5.3.1. Results of Deterministic Model.....	72
5.3.2. Results of Model with Probabilistic Sensitivity Analysis.....	73
5.3.3. Cost Effectiveness Acceptability.....	76
5.3.4. Value of Information Analysis.....	78
5.4. Discussion.....	79
5.4.1. Model Assumptions.....	79
5.4.2. Sensitivity Analysis.....	80
5.5. Conclusion.....	81
<b>6. Discussion.....</b>	<b>83</b>
6.1. Strengths of Project.....	83
6.2. Limitations.....	83
6.2.1. Time Horizon.....	83
6.2.2. The Use of Local Cost Data.....	85
6.2.3. The Use of Evidence from Case Series and Expert Opinion.....	86
6.3. Implications for Decision Makers.....	87
6.4. Future Research.....	87
6.4.1. Design of a Randomized Control Trial to Assess Implantable Doppler Technology.....	88
6.4.2. Utility Values of Head and Neck Health States.....	90

**7. Conclusion.....92**  
**References.....93**

**List of Figures**

**Systematic Review**

**Figure 1:** Data Extraction Sheet for Systematic Review.....11

**Figure 2:** Flow Chart of Systematic Review.....14

**Utility Questionnaire**

**Figure 3:** Health States Associated with Free Tissue Transfer in Head and Neck  
Cancer.....33

**Health Economic Evaluation**

**Figure 4:** Model Format.....62

**Figure 5:** Scatterplot Derived from PSA.....74

**Figure 6:** Probability of Cost Effectiveness Acceptability.....77

## **List of Tables**

### **Systematic Review**

**Table 1:** Newcastle-Ottawa Quality Assessment Scale Case Control Studies.....14

**Table 2:** Summary of Free Tissue Transfer in Systematic Review.....17

**Table 3:** Calculated Pooled Proportions of Primary Endpoints.....18

**Table 4:** Placement of Implantable Doppler Probes in Systematic Review.....26

### **Utility Questionnaire**

**Table 5:** Utility Values of Health States Associated with Free Tissue Transfer to Treat Head and Neck Cancer as Determined by Time Trade-Off Exercise.....36

**Table 6:** Pearson Correlation Coefficients between health States.....37

### **Cost Study**

**Table 7:** Average Length of Stay and Associated Costs.....45

**Table 8:** Procedural Costs Associated with Free Tissue Transfer.....45

**Table 9:** Billing Codes Associated with Each Procedure.....45

**Table 10:** Comparison to Median Literature Values.....47

**Table 11:** Comparison to Mean Literature Values.....47

### **Health Economic Evaluation**

**Table 12:** Summary of Inputs.....63

**Table 13:** Parameter Distributions used in Probabilistic Sensitivity Analysis.....71

**Table 14:** Results of Deterministic Model.....73

**Table 15:** Results of Probabilistic Analysis.....75

**Table 16:** Results of Probabilistic Analysis with Two Probes.....76

**Table 17:** Results of Probabilistic Analysis with Utility Values Derived from the Masters of Epidemiology Group.....75

**Table 18:** Results of Probabilistic Analysis with Utility Values Derived from the Department of Otolaryngology Group.....76

**Table 19:** Estimate of Expected Value of Perfect Information.....78

**Table 20:** Estimates of Expected Value of Partial Perfect Information.....78

**Table 21:** Estimates of Expected Value of Partial Perfect Information for Salvage Rate and Take Back Rate.....79

**List of Appendices**

**Appendix A:** Excluded Full Papers from Systematic Review.....111

**Appendix B:** Screen Captures from Online Utility Survey.....115

**Appendix C:** Derivation of Revision Proportion in Decision Analytic Model.....121

**Appendix D:** Sample Calculation of Transition Probabilities.....122

**Appendix E:** Results of the Decision Analytic Model as a Function of the Number of Replications of the Model.....125

**Appendix F:** Clinical Pathway of Free Tissue Transfer in the Head and Neck at the Ottawa Hospital.....127

## **List of Acronyms**

CEAC	cost acceptability curve
DSA	deterministic sensitivity analysis
EVPI	expected value of perfect information
EVPII	expected value of perfect partial information
FTT	free tissue transfer
HNC	head and neck cancer
HS1	health state 1
HS2	health state 2
HS3	health state 3
HS4	health state 4
ICU	intensive care unit
INB	incremental net benefit
MOH	Ministry of Health
OSR	overall survival rate
QALYs	quality adjusted life years
RCT	randomized controlled trial
SDU	step down unit
SG	standard gamble
TP1	transition probability 1
TP2	transition probability 2

TP3	transition probability 3
TP4	transition probability 4
TP5	transition probability 5
TP6	transition probability 6
TR	take-back rate
TTO	time trade-off
VOI	value of information
VAS	visual analog scale



## **1. Introduction**

Comprehensive surgical treatment of cancer includes excision of both the cancer and a margin of normal tissue. For patients to live a meaningful life after cancer surgery the function of this lost tissue must be appropriately reconstituted. Cancer of the upper aerodigestive tract (UADT), often called head and neck cancer (HNC), is an excellent example of how crucial it is to replace tissue lost during cancer-treating surgery. Functions of the UADT include breathing, speaking and swallowing. When tissue is removed from the UADT during cancer surgery, these crucial processes are often compromised. The quality of life of survivors of HNC depends on how well speech, respiration, and swallowing are restored after treatment.

Fortunately, over the last three decades the ability to reconstruct defects of the UADT has improved dramatically. This improvement is largely the result of the application of free tissue transfer (FTT) to the reconstruction of UADT defects. FTT is a technique that was developed first in the 1950s and first applied to HNC in the late 1970s (Tamai, 2009). FTT involves harvesting tissue from one part of the body and using this tissue to reconstruct a defect elsewhere in the body. The FTT reconstruction most often is performed immediately after the cancer and its surrounding tissue have been removed.

FTT can transfer various different types of tissue to better reconstitute the function of the lost tissue. Thin forearm skin and connective tissue is use to replace the lining of the mouth or throat. Thicker tissue from the thigh can be used to replace large defects of the tongue. Leg bones can be transferred to replace missing segments of the jaw.

For the transferred tissue to survive in its new environment it must have an adequate blood supply. To ensure the tissue has an adequate blood supply, the tissue is harvested with its constituent artery and vein. This artery and vein are then connected to an artery and vein in the area of the defect. This connection, called the anastomosis, provides the blood supply necessary for the transferred tissue to survive.

While FTT has a success rate of close to 95% when used to treat HNC, failure of the anastomosis is a disaster (Blackwell, 1999; Blackwell, Brown, Gonzalez, 1997; Haughey et al, 2001). Failure of the anastomosis not only can kill the transplanted tissue, but also can cause infection and other life threatening complications in the neck. Patients experiencing FTT failure have prolonged hospital stays, often require subsequent operative procedures, and suffer from a diminished quality of life.

For this reason, surgeons are ever vigilant for the possibility of FTT failure. Early recognition of a failing FTT is extremely important (May et al. 1978, Kerrigan et al., 1984). If an FTT failure is recognized early enough it is possible to return the patient to the operating room and revise the anastomosis. Using this type of timely intervention it is possible to salvage a failing FTT (Furnas and Rosen, 1991). Successful salvage of a failing FTT prevents the long hospital stays, secondary operative procedures, and poor quality of life associated with FTT failure.

Given the importance of recognizing FTT failure early, surgeons have developed many techniques to monitor the tissue transferred during FTT (Furnas and Rosen, 1991; Spiegel and Polat, 2007). These monitoring techniques aim to recognize a failing FTT early. Conventional monitoring techniques usually require intermittent checks by health care professionals. Such techniques include monitoring the colour and temperature of

transferred tissue, pricking transferred tissue with a needle to see if it bleeds, or examining the capillary refill of the transferred tissue. Other methods involve using a transcutaneous Doppler probe to assess flow in the transferred tissues artery (Luu and Farwell, 2009; Smit et al., 2010).

A newer method of FTT monitoring is the Cook-Swartz Doppler flow monitoring system (Cook Vascular Incorporated, Vandergrift, Pennsylvania). This was introduced by Swartz in 1988 (Swartz et al, 1988). The Cook-Swartz Doppler flow monitoring system uses a small Doppler flow monitor that is placed directly on the vessels of the anastomosis. The device works as an implantable Doppler technology (IDT). The Doppler flow monitor consists of a small piezoelectric crystal that is embedded in a soft silicone sheath. Initially the silicone sheath was sutured to the blood vessel being monitored. More advanced methods of attaching the sheath to the vessel have now been developed. It may be placed directly on the artery or the vein.

Unlike conventional techniques which provide only intermitted data The IDT gives continuous information on the status of blood flow across the anastomosis (Swartz, Izquierdo, Miller, 1994). The IDT relays information about blood flow via a constant audible signal. Specific audible signals are associated with both a functioning vein and a functioning artery. Absence of these signals or change in these signals is an indication of reduced blood flow to the transferred tissue. As IDT does not rely on visualizing the transferred tissue, it also can monitor tissue that has been transferred deep within the body.

While IDT may offer advantages over conventional techniques, surgeons are in agreement that a detailed clinical exam of a patient by a surgeon trained in FTT

techniques followed by examination of the anastomosis in the operating room is the gold standard for determining if a FTT (often called a free flap or simply a flap) is failing or not. Consequently, IDT is used as an adjunct to intermittent clinical checks by surgeons. The benefit of using IDT as an adjunct is that it may provide information between checks by a surgeon or nursing staff and thus allow earlier detection of a failing flap. Earlier detection has been shown to improve salvage rates (SR) of FTT (Kerrigan, Zelt, and Daniel, 1984; May et al. 1978). Thus, IDT offers the potential advantage of reducing the catastrophic complication of FTT failure by improved SR.

There have been no direct comparisons in SR and overall success rates (OSR) between FTT monitored with IDT and those monitored with conventional techniques. However, there are published case series of FTT monitored with IDT showing salvage rates and success rates higher than case series of FTT using conventional monitoring. These same authors have also found that the rate at which FTT procedures are taken back to the operating room to be reassessed, the take-back rate (TR), is higher in cases using IDT technology (Guillemaud et al., 2008; Paydar et al, 2010). As yet, there has been no systematic review of the subject in the literature.

The benefits of a higher FTT salvage rate are obvious. FTT failure significantly reduces quality of life, leads to prolonged hospitalization, and often requires multiple surgical procedures to fix. Costs associated with FTT failure could be expected to be substantial. If IDT does significantly improve SR and reduce FTT failure, one would expect both reduced costs associated with FTT procedures and an improved quality of life for patients undergoing FTT procedures.

However, the benefit of IDT remains unclear. Although FTT failure is disastrous, it is quite rare. As previously noted, FTT is a success 95% of the time. Consequently, any new intervention has little room to improve the overall success rate of FTT. The case for the use of IDT technology further suffers from the fact that most protocols call for its use on all FTT procedures. This means that given the high success rate of FTT, the vast majority of cases will gain no benefit from the use of IDT technology. In addition, IDT probes are designed to be single use. The result is that IDT costs become significant with many patients deriving no benefit from the intervention.

A health region or hospital considering whether or not to adopt IDT technology has no strong basis to guide its decision. Certainly the benefits of avoiding FTT failure are obvious; but, the relationship between cost and benefit with IDT technology is not clear.

The objective of this thesis was to provide data and analysis which could aid the decision to adopt IDT technology. To achieve this goal, the thesis incorporates a cost-utility analysis of IDT technology through development of a decision analytic model.

The theoretical basis of cost-utility analysis is utility theory. Utility theory as it applies to health economics is used to quantify preferences for different health outcomes, also called health states (Drummond et al., 2005). More desirable health states have higher utility values. Cost-utility analyses are designed to measure the improvement in utility associated with the cost of an intervention.

The cost-utility analysis of IDT technology included several parts. First a systematic review of the literature was performed to establish the efficacy of IDT in monitoring FTT procedures performed for HNC. This systematic review sought to firmly

establish literature values for the OSR and the SR of FTT in the head and neck region when conventional monitoring techniques, not IDT, were used. Then, the systematic review aimed to establish the literature values for OSR and SR in FTT procedures monitored with IDT.

The second component of the cost-utility analysis was a determination of the utility values of health states associated with FTT in the head and neck. As there are no published values on this subject, the values were determined by a utility questionnaire.

Finally, to determine the costs of FTT, a costing study was performed. The costs were obtained using a cost analysis tool known as microcosting.

Once the efficacy of IDT was estimated using the systematic review, the utility values of FTT in the Head and Neck were obtained, and the costs associated with FTT were established, all the inputs for economic evaluation were set. A Markov decision model was then created and a decision analysis was carried out. A probabilistic sensitivity analysis was then performed on the results of the decision analysis.

## **2. Systematic Review**

### **2.1 Objective**

A systematic review was performed to answer the question of whether the use of Cook-Swartz Doppler (IDT) technology in patients undergoing free flap surgery to repair head and neck defects improves free tissue transfer (FTT) salvage rates (SR) and overall success rates (OSR) compared to conventional clinical monitoring techniques.

The project's population, intervention, comparison, and outcome, (PICO) statement is the following:

Population: Patients undergoing FTT surgery to repair defects created by surgery for cancer of the upper aerodigestive tract (UADT).

Intervention: Use of IDT technology to monitor FTT procedures for failure

Comparison: Standard clinical techniques of FTT monitoring, including pinprick, colour assessment, capillary refill, transcutaneous Doppler.

Outcome: The primary outcome was the SR of FTT procedures. Secondary outcomes included the take-back rate (TR), and the OSR.

### **2.2 Methods**

#### ***2.2.1 Literature Search***

MEDLINE (Ovid) was the primary search engine used. Medline records from 1950 to the second week of April 2010 were searched with the keywords "head and neck neoplasms" and "surgical flaps" were used for the search. The keywords were searched together using the "and" conjunction. Both keywords were exploded. Results were limited to those published in English with only human subjects. Only publications in English were used as these were felt to represent the vast majority of publications and

because accurate and timely translation of other languages was not available. All citations were screened for eligibility.

In addition to the Medline search, supplemental searches were performed using Pubmed, Scopus, and Embase as search engines. The same keywords were used in these searches and the same exclusion criteria were applied. Furthermore, the grey literature was searched using Google Scholar and Scirus as search engines.

A single reviewer was used (MG). It would have certainly preferable to have multiple reviewers; however, the limitations of thesis work led to only a single reviewer being available. An academic librarian was consulted to assist with the literature search.

### *2.2.2 Eligibility Criteria*

Articles were only eligible if they included a consecutive series of all of the FTT procedures performed for head and neck defects at a given institution for a given period of time. Many articles constituted a case series of a specific type of FTT procedure. These articles do not represent true clinical practice, which involves a variety of FTT procedures. Consequently, these articles were excluded.

Articles that did not clearly state their monitoring technique were excluded. An attempt was made to contact the corresponding author of all articles that met other criteria but did not state their monitoring technique. Articles whose authors responded with their monitoring technique were included.

Articles that contained series of FTT cases monitored by a combination of IDT and clinical techniques were included only if the authors specifically stated the breakdown of cases receiving IDT monitoring and those receiving clinical monitoring and the respective SR and OSR for both groups. Similarly, articles that included FTT

used to treat defects outside the Head and Neck region were used only if the key outcomes of OSR, SR, and TR for cases involving the Head and Neck was clearly stated.

When there were multiple published reports from the same institution during overlapping time periods, the larger of the case series was included in our review and the other publications were excluded.

### *2.2.3 Data Extraction*

Three *a priori* outcome measures were identified. These outcomes were OSR, SR, and the take-back rate (TR). The TR was the proportion of flaps that were returned to the operating room for revision of the FTT anastomosis. The TR is a crucial variable with respect to the effectiveness of the alternative procedures as it is likely to be one prime difference between the two treatment alternatives. The three outcome measures were chosen as they were the most clinically relevant. Other outcome measures such as the sensitivity, specificity, negative predicative value, and positive predicative value are not as directly related to patient experiences and outcomes nor are they relevant for incorporation into the model.

A standardized data extraction sheet was developed and used for all data extraction (Figure 1). In addition to the three primary outcome measures, data on age, sex, flap types, type of malignancy, prior radiation or chemotherapy, location of tumour, and use of special techniques such as vein grafts were collected.

### *2.2.4 Analysis*

#### *Quality Assessment*

The Newcastle-Ottawa Scale for Case Control Studies (available at [http://www.ohri.ca/programs/clinical\\_epidemiology/oxford.htm](http://www.ohri.ca/programs/clinical_epidemiology/oxford.htm)) was used to rate the quality of all the studies.

*Pooled Estimate of Proportions for OSR, SR, and TR*

A pooled estimate of proportions was used to calculate each primary endpoint. The pooled estimate was calculated using both a random effects model and a fixed effects model. The software used to calculate the pooled proportions transformed the proportions first into a quantity so that they would be suitable for fixed effects and random effects models. This transformation was according to the Freeman-Tukey variant described by Stuart and Ord in 1994 (Stuart and Ord 1994). The pooled proportion was then back-transformed to a pooled proportion.

To calculate the heterogeneity in the data the  $I^2$  statistic was used. An attempt to define heterogeneity was made using sensitivity analysis and subgroup analysis. The confidence interval around the  $I^2$  statistic was calculated using the non-central chi-squared method (Hedges and Piggott, 2001). The fixed effects model was weighted using the inverse arcsine variant weights. The random effects model was weighted using the Dersimonian-Laird method (Dersimonian and Laird, 1986). The StatsDirect (Cheshire, UK) software program was used for all statistical calculations.

**Figure 1: Data Extraction Sheet for Systematic Review**

Title and Bibliographic Information:

Authors:

Number of patients:                      Gender ratio                      Age (median or mean)

Number of procedures:

Number of flaps:

Breakdown of indications for flaps (% or absolute number):

Breakdown of sites reconstructed (% or absolute number):

Breakdown of types of flaps used (% or absolute number):

Breakdown of preoperative treatment (none vs. radiation vs. chemoradiation vs. surgery, % or absolute number)

Monitoring technique:

Take backs (absolute number):

Salvage (absolute number):

Partial fails (absolute number):

Total fails (absolute number):

Vein graft use:

Other:

## 2.3 Results

### *2.3.1 Summary of Literature Search*

The primary Medline search returned a result 3993 articles. Once eligibility criteria were applied the number of articles was reduced to 20 (see Table 1 for list of references). The use of supplementary search engines did not identify any additional articles. While several relevant posters and oral presentations were identified using Scirus, none contained information on the key endpoints of OSR, TR, or SR. As a result, a search of the gray literature yielded no additional articles (Figure 2). The excluded full articles are listed in Appendix A.

Of the 20 articles that were left after the exclusion criteria were applied, three were case series using IDT technology. The other seventeen of these studies reported cases with conventional clinical monitoring as the monitoring technique. All of these studies were consecutive case series. All of these articles collected data retrospectively. None of these papers included a control arm.

### *2.3.2 Quality Assessment*

Table 1 summarizes the Newcastle-Ottawa scores for all the included studies.

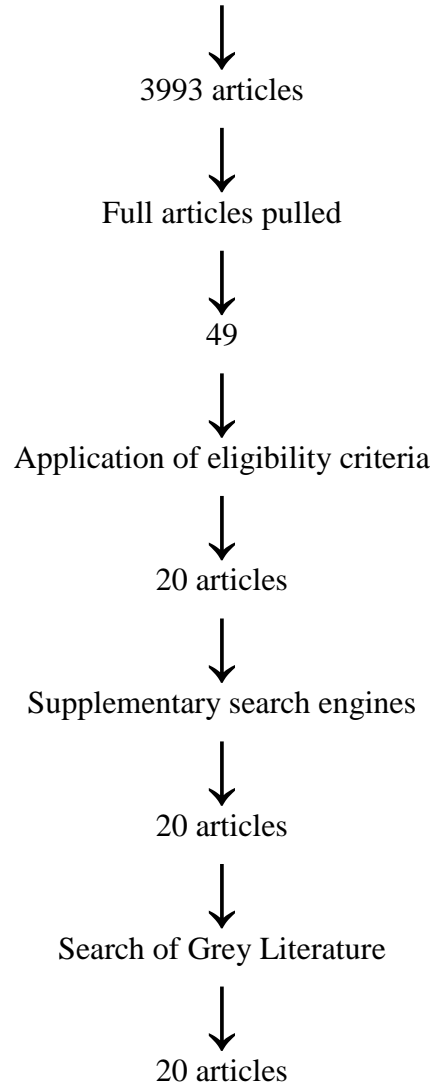
### *2.3.3 Descriptive Analysis*

Prior to the calculation of the OSR, SR, and TR, a descriptive analysis of the study populations and clinical techniques was performed. The presence of risk factors for FTT failure was analyzed. Few risk factors for FTT failure have been clearly identified in the literature. Podrecca et al. have shown that the use of vein grafts to lengthen the blood vessels of transferred tissue increases the FTT failure rate (Podrecca et al., 2006). Though it is worth noting that not all authors have found that interposition vein grafts

have been associated with higher rates of FTT failure (Fur, Canady, Wax, 2011). Unfortunately, very few of the studies in this review reported on the use of vein grafts. Most studies that did mention vein grafts stated no vein grafts were used. Vein graft use is generally avoided as it complicates FTT surgery (Fur, Canady, Wax, 2011). Given the lack of reporting on vein graft use and the overall reticence of FTT surgeons to perform vein grafting it was assumed that rates of vein grafting between studies were not significant enough to alter the results of the systematic review.

**Figure 2: Flow Chart of Systematic Review**

Medline Search (1950-2<sup>nd</sup> week of April 2010) keywords “surgical flaps” and “head and neck neoplasms”



**Table 1: Newcastle-Ottawa Quality Assessment Scale Case-Control Studies**

Study	Selection	Comparability	Outcome
<b>Clinical Monitoring Group</b>			
Eckhardt et al. 2007	*		*
Suh et al. 2004	*		*
Pohlenz et al. 2007	*		*
Nakamizoo et al. 2004	*		*

O'Brien et al. 1998	*		*
Tabah et al. 1984	*		*
Dassonville et al. 2008	*		*
Simpson et al. 1996	*		*
Hoffmann et al. 1998	*		*
Ross et al. 2003	*		*
Ross et al. 2008	*		*
Singh et al. 1999	*		*
Urken et al. 1994	*		*
Nakatsuka et al. 2003	*		*
Hyodo et al. 2007	*		*
Podrecca et al. 2006	*		*
Yoshimoto et al. 2010	*		*
<b>Implantable Doppler Group</b>			
Guillemaud et al. 2008	*		*
Pryor et al. 2006	*		*
Paydar et al. 2010	*		*

No study has proven that any diagnosis increases the chance of FTT failure; however, some surgeons believe that HNC, which often occurs in smokers and patients with poor functional status, carries a higher risk of FTT failure.

Although, all case series populations primarily consisted of patients being treated for HNC, most case series also included a small number of non-malignant conditions that were treated with FTT. Not all series reported the breakdown of diagnoses in a similar manner and as a consequence it was not possible to directly compare the studies. Despite the inability to perform a direct comparison, the breakdown of diagnoses was very similar in both the IDT and clinical monitor groups. Only one of the case series had less than 80% of its cases with a diagnosis of malignancy. This case series used the clinical

monitoring technique and contributed only 2.9% of the total number of cases in the entire clinical monitoring group.

In the 17 case series using conventional monitoring techniques, a total of 7750 FTT procedures were performed. In the three case-series using IDT, 577 FTT were performed. In the conventional monitoring group, 390 FTTs were returned to the operating room for revision of the FTT anastomosis. In the IDT group, 58 FTTs were returned to the operating room. Of those returned to the operating room in the conventional group, 186 FTTs were successfully salvaged. In the IDT group, 49 FTTs were salvaged. In the conventional monitoring group, 7415 FTTs succeeded. In the IDT group, 568 FTTs succeeded (See Table 2).

#### *2.3.4 Pooled Estimate of Proportions for OSR, SR, and TR*

The pooled estimates for OSR, SR, TR along with the respective  $I^2$  statistics are summarized in Table 3. The pooled estimate of the OSR for the clinical monitor group was 95.8% (95% confidence interval: 94.9%-96.6%). The pooled estimate of the OSR for the IDT group was 98.3% (95% confidence interval: 96.8%-99.3%). The pooled estimate of the SR for the clinical monitoring group was 51.0% (95% confidence interval: 38.9-62.7%). The pooled estimate of the SR for the IDT group was calculated from only two studies. This was because one of the studies had only a single attempt at salvaging an FTT (Pryor et al. 2006). Consequently, a decision was made that this study's SR would not be meaningful. Based on the two studies the pooled estimate of SR was 85.2% (95% confidence interval: 75.3-96.2%).

The pooled estimate of TR for the clinical monitoring alone group was 6.6% (95% confidence interval: 4.8-8.7%). For the IDT group the pooled estimate of the TR was 10.2% (95% confidence interval: 7.9-12.8%).

These pooled estimates of SR and TR were used to as inputs in the decision model. The confidence intervals were used in the sensitivity analysis portion of the decision model. Although the estimate of OSR was not used in the model, its calculation was felt to be informative to future work.

Table 2: Summary of Free Tissue Transfer in Systematic Review

	<b>Total</b>	<b>Take Back (% of total)</b>	<b>Salvage (% of take back)</b>	<b>Success Rate (% of total)</b>
<b>Clinical Monitoring Alone Group</b>	7750	390 (5.0)	186 (47.7)	7415 (95.7)
<b>Eckhardt et al. 2007</b>	534	37 (6.9)	10 (27.0)	507 (94.9)
<b>Suh et al. 2004</b>	400	7 (1.8)	4 (57.1)	397 (99.3)
<b>Pohlenz et al. 2007</b>	540	22 (4.0)	15 (68.2)	505 (93.5)
<b>Nakamizo et al. 2004</b>	187	7 (3.7)	2 (28.6)	182 (97.3)
<b>O'brien et al. 1998</b>	250	21 (8.4)	11 (52.4)	240 (96.0)
<b>Tabah et al. 1984</b>	75	5 (6.6)	3 (60.0)	70 (93.3)
<b>Dassonville et al. 2008</b>	213	32 (15.0)	21 (65.6)	199 (93.4)
<b>Simpson et al. 1996</b>	150	31 (20.7)	24 (77.4)	143 (95.3)
<b>Hoffmann et al. 1998</b>	227	9 (4.0)	5 (55.5)	223 (98.2)
<b>Ross et al. 2003</b>	151	9 (6.0)	5 (55.5)	147 (97.4)
<b>Ross et al. 2008</b>	492	32 (6.5)	22 (68.8)	468 (95.1)
<b>Singh et al. 1999</b>	201	18 (9.0)	14 (77.8)	197 (98.0)
<b>Urken et al. 1994</b>	200	18 (9.0)	5 (27.7)	187 (93.5)
<b>Nakatsuka et al. 2003</b>	2372	50 (2.1)	21 (42.0)	2272 (95.8)
<b>Hyodo et al. 2007</b>	513	21 (4.0)	7 (33.3)	499 (97.3)

<b>Podrecca et al. 2006</b>	347	29 (8.4)	10 (34.5)	328 (94.5)
<b>Yoshimoto et al. 2010</b>	898	54 (6.0)	7 (13.0)	851 (94.8)
<b>Implantable Doppler Group</b>	577	58 (10.1)	49 (84.5)	568 (98.4)
Guillemaud et al. 2008	384	38 (10.0)	31 (81.6)	377 (98.2)
Pryor et al. 2006	24	1 (4.2)	0 (0)	23 (95.8)
Paydar et al. 2010	169	19 (11.2)	18 (94.7)	168 (99.4)

**Table 3: Calculated Pooled Proportions of Primary Endpoints from the Systematic Review**

	<b>Fixed Effects (95% CI)</b>	<b>Random Effects (95% CI)</b>	<b><i>I</i><sup>2</sup> (95% CI)</b>
<b>Clinical Monitoring Group</b>			
Take back rate	0.048 (0.043-0.053)	0.066 (0.048-0.087)	90.6% (86.9%-92.8%)
Salvage rate	0.459 (0.412-0.507)	0.510 (0.389-0.627)	80% (67.8%-86.2%)
Success rate	0.957 (0.953-0.962)	0.958 (0.949-0.967)	67.9% (47.3%-79.4%)
<b>Implantable Doppler Group</b>			
Take back rate	0.102 (0.0789-0.128)	0.102 (0.789-12.8)	0% (0%-72.9%)
Salvage Rate	0.854 (0.753-0.931)	0.852(0.753-0.962)	Only two studies
Success Rate	0.983 (0.971-0.992)	0.983 (0.968-0.993)	16% (0%-77.1%)

## 2.4 Discussion

This systematic review calculated the pooled proportion of take back, salvage and overall success in patients with HNC undergoing FTT procedures with and without the assistance of IDT. While there is not much literature to guide the results of the IDT group, the finding in the clinical monitoring group of an overall success rate in excess of 95% is consistent with reviews on FTT procedures in the head and neck (Smith et al., 2007; Genden et al., 2004).

The quality of the conclusions of this review is limited by the fact that all included studies were retrospective case series. At the time of the literature review there were no studies that had a comparison group. Recently, a few studies examining the use of IDT for FTT procedures of the lower limbs and in breast reconstruction following mastectomy have included comparison groups (Rozen et al., 2010; Rozen, Chubb et al., 2010). These studies were also retrospective analyses. The comparison groups were matched controls and the decision to use IDT was arbitrary. These studies have found conflicting results in the efficacy of IDT monitoring.

The use of non-randomized studies in systematic reviews does raise the possibility of bias and confounding. The Cochrane Collaboration has authored specific recommendations regarding the use of non-randomized studies in systematic reviews (Reeves et al., 2008). They note that the use of non-randomized studies in systematic reviews was legitimate in cases where randomized trials are very unlikely to be performed. They also identified several pitfalls in performing a systematic review with non-randomized studies. Key among these pitfalls was the increased chance of confounding in the study. Subsequently, studies may suffer from systematic bias or increased heterogeneity in the measured effect size.

Unfortunately, the recommendations of the Cochrane Collaboration are largely related to the inclusion of studies that have a comparison group and do not give much guidance in the case of systematic reviews of studies without control groups. The public health literature does include some evidence on the use of systematic reviews of studies without control groups. The National Collaborating Centre for Methods and Tools published a review of methods for performing systematic reviews of studies without

comparison groups (Fitzpatrick-Lewis et al., 2009). This comprehensive document tackles many of the issues of performing systematic reviews using studies without comparison groups, including rating quality of studies and synthesizing data from studies without control groups. The document also refers to work by Dalziel et al. specifically looking at the use of case series data for health technology assessment in the surgical literature.

Dalziel et al. assessed the results of case-series by methodology and where possible compared results of case series to the intervention arm of randomized controlled studies (Dalziel et al., 2005). The authors examined several factors that could create bias in case series including, retrospective design, small sample sizes, the length of follow-up, the use of blinded observers, the date of publication, and multi-centre trials. No factor was consistently associated with bias. Some analyses showed multi-centre trials and date of publication were associated with a positive bias. Others that the use of blinded observers and length of follow-up were associated with a negative bias. The authors cautioned that these results could be spurious given the large number of null hypotheses tested.

Furthermore, the authors found that results of case series when compared to those of randomized controlled trials were not significantly different. Dalziel et al. also reviewed the health technology assessments of the National Institute for Health and Clinical Excellence and found that 14 included evidence from case series.

Another issue with the use of non-randomized studies is the assessment of the quality of data. This review used the Newcastle-Ottawa Scale. All studies in this review were rated the same. All earned a star for the representativeness of cases and all earned a star for the ascertainment of exposure. This scale was not adapted for use with case-

series. To the author's knowledge, no published method is available to guide the modification of the Newcastle-Ottawa Scale for use in case series.

Fitzpatrick-Lewis et al. address the issue of quality assessment. They mention several publications, though the most relevant to this study is a work by Steuten et al (Steuten et al., 2004). Steuten et al. created a scale specifically designed for Health Technology Assessments that accommodated both randomized and observational designs. They validated the score for inter and intra-observer variability and in terms of hierarchical ranking. This scale is another option for systematic reviews of health technology.

Publication bias is another threat to the results of this review. While efforts were made to include all studies available including those from the gray literature, publication bias likely remains a significant factor in systematic reviews of surgical case series.

It has long been recognized in the surgical literature that when high-volume centres publish data relating to complication rates, these rates may be significantly different than those of lower volume centres and may not represent true population rates (Syin et al., 2007). Furthermore, series that show good patient outcomes are more likely to be published than those showing poor outcomes (Yoshimoto, 2003; Dickersin, 1990). Some biostatisticians have identified methods that attempt to correct for this publication bias (Preston, Ashby, Smyth, 2004).

Unfortunately, a second reviewer was not used for this review. Second reviewers are helpful to ensure completeness and reproducibility of the review. Inter-reviewer agreement can be measured and used to ensure the methods of the review were valid. As there was only a single reviewer (MG) that reviewer completed the review at each stage

twice to ensure completeness. One study was removed from the list of included studies during this second review as it became clear it represented a specialized population of free flaps (Ferguson and Yu, 2009).

The methods of this paper dictated that where two series from the same institution over the same time period were published the larger series was used. This approach did not consider that some smaller studies may have had improved methodology. Fortunately, this was not much of an issue in our review. The quality of the studies in general was similar between larger and smaller studies. There was a case where a smaller study was used when a larger one was in the literature but did not fulfill our inclusion criteria. This study was the work by Singh et al. (Singh et al., 1999).

One advantage of the studies used in this systematic review was that as they were all observational studies of consecutive cases, they are pragmatic and represent the entire experience of head and neck FTT procedures. Reporting related to the TR, SR, and OSR was all relatively good and consistent. However, reporting on other factors that may affect FTT success was poor. Few studies mentioned the use of vein grafts. No studies commented on nicotine use, which is known to constrict blood vessels, and theoretically would be associated with a higher rate of FTT failure. There was also heterogeneity in both how the clinical monitoring and the IDT monitoring was performed.

The clinical monitoring group was monitored by a variety of techniques including temperature, pinprick, colour and transcutaneous Doppler. Teasing out which technique was used on which patient was impossible. The general approach is to use simple non-invasive means of measuring FTT if initial indicators such as colour suggest good blood flow. More invasive tests such as transcutaneous Doppler and pinprick tend to be used if

the initial assessment raises concern. The specific regiment including frequency of tests and types of tests used varied between studies.

While this heterogeneity is troublesome from an analytic point of view, it is likely unavoidable. Although it is very likely that some clinical protocols are superior to others, surgeons will always be somewhat arbitrary in how they apply their specific brand of clinical FTT monitoring. It is highly unlikely a randomized control trial will be performed comparing different regiments of clinical monitoring. This type of variety in clinical monitoring technique is likely going to be the reality in any comparison of case series.

In the IDT group, the use of the IDT was not uniform. The largest study used a combination of techniques (Guillemaud et al., 2008). The majority, 77.8%, of the patients in their study were monitored using a single IDT probe on the artery. Only 1.1% was monitored using a single IDT probe on the vein. The rest, 21.1%, were monitored using a multiple probes. The second largest study used a single IDT probe on the vein for all patients (Paydar et al., 2010). The smallest study, Pryor et al., used a single IDT probe on the artery in 48%, a single venous probe in 48% and multiple probes in 4% (Pryor et al., 2006). This is summarized in Table 4.

Guillemaud et al. found no significant difference in outcomes between patients monitored with a single probe and those monitored with multiple probes. However, this comparison may be biased. The authors do state that later in the case series they began to place the probes on both the artery and vein but otherwise do not explain how they chose whether to place probes on the artery, vein, or both. They did not compare the outcomes

of FTT monitored with single probes on the artery to those monitored with single probes on the vein (Guillemaud et al. 2008).

Theoretically, the use of probes on the vein should make IDT technology more efficient. Venous failure may precede arterial failure and thus only monitoring the artery may increase the time from FTT failure to recognition of failure by medical staff. If arterial failure is the inciting cause, venous failure should occur nearly simultaneously (Olivas et al., 2001, Kubo et al., 2002). Swartz, in a follow-up to his initial study, used an animal model to address the question of whether placement on the vein was preferable to placement on the artery. He found that he was able to detect vessel failure earlier and salvage failing vessels more successfully if the IDT probe was placed on the vein (Swartz et al., 1994). Perhaps as a consequence of this animal model as well as theoretical considerations, the manufacturer recommends single placement on the vein.

However, interpretation of the signal from a probe on the artery is easier than interpreting the signal from a probe on the vein. The signal from functioning artery is associated with a very obvious sound. Venous sounds can be more subtle and require more experience to interpret. Consequently, many clinicians prefer probe placement on the artery.

The significance of different uses, vein versus artery placement, of IDT is unclear. However, this clinical diversity should be considered when analyzing the proportions calculated from this systematic review.

The result of this clinical diversity between the included studies was significant heterogeneity between studies. The  $I^2$  statistic gives information on the degree to which this heterogeneity affects the analysis (Deeks et al., 2008). There was significant

evidence that in the large group of studies that used clinical monitoring there was significant inconsistency. The results for the IDT group were less convincing given the small number of included studies. However, one could expect similar inconsistency if the sample became larger given the results of the clinical monitoring group.

Investigations of this heterogeneity are limited because most studies did not provide the requisite information to perform either a subgroup analysis or a meta-regression.

Sensitivity analysis could be performed on the clinical monitoring group, but further analysis of the IDT group was not possible as so few studies were found. Performing a sensitivity analysis with the addition of excluded trials is likely not useful. The inclusion criteria were fairly broad and the exclusion criteria are straightforward and not contentious. The methodology of the included studies does not differ significantly. There was not consistently enough information provided in the papers to perform a sensitivity analysis on the basis of flap type, site of defect, age of patients, tobacco use, vein graft use.

Given the heterogeneity found, a decision had to be made about which analysis, a random-effects analysis or a fixed-effects analysis, to use. The fixed-effects analysis assumes that variability between the results of studies is based on chance alone, and that the true result is the same in all studies. A random effects model assumes that the true result varies between studies and does so in a normal distribution (Deeks et al., 2008).

The difference between these two assumptions is significant. As Deeks et al. explain, in the setting of heterogeneity due clinical diversity, the fixed effects model is a best estimate of the true value whereas the random effects model is a representation of

the average value. Deeks et al. also point out that the decision to choose one analysis over another should never be based on a test for heterogeneity alone, such as the  $I^2$  statistic.

The random effects estimates were chosen as model inputs for several reasons. First, the assumption that there will be a distribution of true proportions is reasonable. This was especially true for the proportions of take back and salvage. The decision to return to the operating room and revise the vascular anastomosis is a very surgeon-dependent decision. Thus, the assumption that the true proportions would be institution dependent is likely correct. With this in mind, an economic evaluation would likely benefit most from an attempt to define the average proportion rather than trying to define a best estimate. Second, the fixed effects model generates a smaller confidence interval, for the purposes of this study, a larger confidence interval is likely more appropriate.

The proportions from this systematic review will be used as key inputs in this project's decision analytic model.

**Table 4: Placement of Implantable Doppler Probes in Systematic Review**

<b>Paper</b>	<b>Number of Flaps in Series</b>	<b>Single arterial probe (%)</b>	<b>Single venous probe (%)</b>	<b>Multiple</b>
<b>Guillemaud et al.</b>	384	77.8	1.1	21.1
<b>Paydar et al.</b>	169	0	100	0
<b>Pryor et al.</b>	24	50	45.8	4.2

### **3. Utility Survey**

#### **3.1 Introduction**

As the goal of this project was to perform a decision analysis of implantable Doppler technology (IDT), it was necessary to quantify the benefit of IDT.

Several techniques can be used to measure the benefit of an intervention. One of the most intuitive and simplest ways for physicians to measure benefit is in clinical terms. In our study, the most relevant clinical measurement would be to measure the number of successful free tissue transfer (FTT) procedures in both groups. Measuring the number of successful FTT procedures is both easy to understand and relevant to clinicians. However, clinical outcome measurements were avoided for two main reasons.

First, the results of the decision analysis are less useful to decision makers who are responsible for decisions across disease areas if the benefit is reported in clinical terms which are difficult to compare (Drummond et al., 2005). For example, the value of successful FTT procedures is difficult to compare to the value of successful procedures in cardiac surgery. To maximize the usefulness of the study, the benefit of IDT had to be expressed in a way that could be more universally applied.

Second, clinical measurements such as failure rates belie the vast gulf in quality of life between an FTT failure and FTT success. Patients whose upper aerodigestive tract (UADT) is successfully reconstructed are able to perform many of the basic functions necessary to ensure adequate quality of life, namely speaking and eating. Failure of FTT procedures is disastrous. Patients are unable to perform basic functions of the UADT and are often hospitalized or are discharged home with a significantly reduced quality of life.

Simply measuring the number of FTT successes, would ignore this large difference and handicap this project.

To address the problem of comparability and ensure that differences in quality of life were captured in this project, the benefit of IDT was quantified in terms of utility. Utility is a measurement of benefit anchored in the utility theory developed by Von Neumann and Morganstern (Drummond et al. 2005; von Neumann and Morganstern, 1944). Different diseases, side-effects, disabilities, and other health states each have a degree of desirability associated with them. Von Neumann and Morganstern utilities can be considered measurements of this desirability. More desirable health states have higher utility values and less desirable health states have lower utility values.

Quantifying the benefit of IDT using utility overcomes many of the pitfalls of using a clinical outcome to quantify benefit. First, utility values can be compared across health fields and are thus more useful to decision makers. Second, utility values are increasingly becoming the standard method of reporting outcomes in health economic evaluations. Using utility values in our study thus allows comparison with the current literature in health economics. Third, utility values should capture the significant difference in quality of life between free flap failure and free flap success.

One of the primary challenges of performing a health economic analysis of IDT in head and neck cancer (HNC) is that the literature contains very little information on the utility values associated with HNC. Furthermore, there is virtually no literature on the utility values associated with FTT (Thoma, Jansen, Sprague, 2009). Consequently, to perform health economic analysis of FTT, utility values associated with FTT to treat HNC needed to be determined.

Health utility values can be determined in several different ways. Torrance et al., classify the different methods of measuring health utility as either direct measurement or the use of multi-attribute systems (Torrance, Furlong, Feeny. 2002).

Multi-attribute systems extend the assumptions of traditional von Neumann and Morganstern utility theory. These systems consist of a questionnaire divided into multiple domains. The questionnaires are completed by subjects currently living in the health state under investigation. Participants rank their level of function in each of these domains. Using a predetermined scoring method, the ranks in each domain are converted to a utility measurement. Several multi-attribute systems exist. The most commonly used include the EQ-5D, the SF-6D, the HUI2 and the HUI3 (Drummond et al., 2005).

A multi-attribute system was not used in this project. As multi-attribute systems require the participation of patients currently in the health states in question, this method of determining utility values works best in diseases that are relatively common. Head and neck cancer is relatively uncommon. Furthermore, the health states associated with FTT failure are even less common. Consequently, accruing an adequate number of participants would take an inordinate amount of time. For this reason, direct measurements were used to determine the utility values of the relevant health states.

Direct measurements consist of various exercises that use healthy respondents to determine the utility value of various health states. The most commonly used direct measurements are standard gamble (SG), time trade-off (TTO), and the visual analog scale (VAS).

The standard gamble is based on the axioms of von Neumann and Morganstern Utility Theory. Standard gamble is an exercise that requires participants to choose

between two alternatives under conditions of uncertainty. One alternative is the health state in question. The second alternative is to undergo a hypothetical treatment which can restore the individual to perfect health with a chance of death. The chance of death is varied until the participant has no preference between the health state in question and the second alternative.

How the participants choose determines the desirability of each health state. Strictly speaking, because the choice is made under uncertain conditions, SG is the only method that measures utility as defined by von Neumann – Morganstern normative decision theory.

Time trade-off is similar to SG in that it is an exercise that requires healthy participants to make a choice. However, the choice is not made under conditions of uncertainty. Instead, participants are asked to determine the number of years of life in a perfect health they would consider equivalent to a fixed number of years in the health state in question. Perhaps as a consequence of this lack of uncertainty, scores obtained by TTO tend to be lower than scores obtained by SG (Krabbe et al, 1997).

Visual analog scales do not require participants to make a choice. Instead, participants rank health states on a scale. Sometimes called a feeling thermometer, the method essentially requires participants to mark the desirability of a health state on a line. One side of the line represents perfect health while the other represents death. The distance along the line is measured and the desirability of the health state is determined. The VAS method suffers from several biases most notably an end-of-scale bias and a context bias (Torrance, Feeny, Furlong, 2001).

The TTO was felt to offer the best method of determining the utility values of health states in FTT surgery in HNC. The TTO method, although not as theoretically sound as SG, remains theoretically robust and true to von Neumann and Morganstern utility theory (Drummond et al., 2005). As previously noted, TTO does not suffer from many of the biases that plague the VAS methods. Time trade-off is often preferred to SG as it is more easily applied and better understood by research participants (Green, Brazier, Deverill, 2000).

### 3.2 Health States in Free Tissue Transfer in Head and Neck Cancer

To determine the utility values associated with IDT technology in Head and Neck FTT it was first necessary to identify the health states associated with Head and Neck FTT. Once these health states were identified, the utility values of each of these health states could be determined.

To identify the health states associated with FTT, expert opinion was obtained. Both the author of the thesis (MG) and the clinical thesis supervisor, Martin Corsten (MC), had significant expertise in the clinical course of FTT procedures. As IDT technology is primarily relevant for the first few weeks after an FTT procedure, a decision was made to examine the health states that occur in the first 6 weeks after an FTT procedure.

Four key health states were identified in these first 6 weeks by the clinical experts (MC and MG). The first health state, HS1, was the health state that all patients experience immediately following an FTT procedure in the UADT. Patients in HS1 were admitted in hospital. They were feed through a feeding tube that passed through their nose into their stomach. They were not able to receive nutrition by mouth. Patients in

HS1 had a tracheotomy tube through which they breathed and which limited their ability to speak. They also had mildly limited mobility (See Figure 3: Health States).

The second health state, HS2, occurred in patients whose FTT procedure failed and who were awaiting revision surgery for a failed FTT. This health state was quite similar to HS1 except that they have regular dressing changes to their neck and are fed through a feeding tube that passes through an incision in their stomach (See Figure 3: Health States).

In the third health state, HS3, patients have recovered from surgery and are at home. They have had a good surgical result and are able to speak intelligibly, they are able to eat most food by mouth, and have resumed a normal level of activity.

In the final health state, HS4, patients have a failed free flap surgery and are healing at home. This state is very similar to HS2 with the only difference being that they are healing at home rather than in hospital.

**Figure 3: Health States associated with Free Tissue Transfer in Head and Neck  
Cancer**

**Health State 1:**

- You are in hospital.
- You have a tracheotomy tube (a breathing tube in your neck) and are not capable of speech.
- You are being fed through a feeding tube in your nose and are not able to take either liquids or solids by mouth.
- Your activity is very limited. You are capable of only very short walks and sitting in a chair.

**Health State 2:**

- You have had surgery for cancer of the head and neck.
- You are admitted to hospital.
- You have a tracheotomy tube (a breathing tube in your neck) and are not capable of speech.
- You are being fed through a feeding tube that enters your stomach through an incision in your abdomen.
- You are not able to take either liquids or solids by mouth.
- You require twice daily dressing changes for an open wound on your neck.
- You are capable of most non-strenuous normal activities.

**Health State 3:**

- You have had surgery for cancer of the head and neck.
- You are at home.
- You are capable of slightly altered, but intelligible speech that can be understood on the telephone.
- You are able to eat most food.
- You are capable of most non-strenuous normal activities.

**Health State 4:**

- You have had surgery for cancer of the head and neck.
- You are at home.
- You have a tracheotomy tube (a breathing tube in your neck) and are capable of only severely altered speech.
- You are being fed through a feeding tube that enters your stomach through an incision in your abdomen.
- You are not able to take either liquids or solids by mouth.
- You are capable of most non-strenuous normal activities.
- You require twice daily dressing changes for a wound on your neck.

### 3.3 Creation of a Time Trade-Off Questionnaire

Once these states were identified, it was necessary to determine the utility values for each state. As previously mentioned, the preferred method was to use TTO techniques. A web-based assessment tool was created to use TTO techniques. Previous studies have used web-based assessments of utility (Chang, Collins, and Kerrigan 2001; Sinno et al. 2010). These studies have shown that web-based assessments of utility are well understood by participants, inexpensive, and simple to administer.

The TTO utility tool in this study consisted of three parts: an introduction and general information section, a sample question with explanation, and the TTO utility assessment of four health states.

The introduction and general information section included information about the survey and a description of the TTO technique. The introduction and general information section also included a contact email and phone number should any of the participants have difficulty understanding the concept of TTO. The sample question allowed participants to see an example of how TTO techniques determine health utility values.

Finally, participants were asked to perform TTO exercises to determine the health utility values of the four FTT associated health states. Following standard TTO techniques, participants were asked how many months of perfect health they would be willing to trade for 10 months of life in each of the health states (ten months was chosen because it was an easy number for participants to consider and because it was closer to the clinically relevant times in each health state than was 10 years). Participants input a number between 0 and 10 months. The web-based tool was hosted by SurveyMonkey

(SurveyMonkey, Palo Alto, CA). The screen captures from the survey are included in Appendix B.

### *3.3.1 Utility Survey Population*

Once created the survey was sent to two groups. The first group was a convenience sample of students in the Masters of Epidemiology program at the University of Ottawa. Although this group did contain a few health professionals, this group was considered to primarily be made up of non-experts in the field. The second group consisted of members of the Department of Otolaryngology at the University of Ottawa. This group included both fully trained Otolaryngologists, including specialists in the treatment of HNC, as well as residents in training. All members of this group had experience with the care and clinical course of patients undergoing FTT for the treatment of HNC. These groups were chosen because it was felt they would have a higher response rate than a larger group of Otolaryngologists or of the general population.

Both groups were contacted via email. To ensure a maximal response rate each group was contacted three times. The group of students in the Master's program was also contacted via email by their Program Director to encourage participation. The group comprised from the Department of Otolaryngology was additionally contacted at a weekly Grand Rounds meeting which most of the Department attended.

## 3.4 Results of TTO Utility Questionnaire

### *3.4.1 Response Rates*

The survey was emailed to 32 members of the Department of Otolaryngology. Fourteen members of the Department of Otolaryngology completed the survey for a response rate of 43.8%. The survey was emailed to 51 members of the Masters of

Clinical Epidemiology program. There were 37 respondents. Three respondents were excluded because of incomplete responses. This gave a response rate for the Masters of Clinical Epidemiology program of 73% and an overall response rate of 61%.

### 3.4.2 Utility Values

The utility values determined by the TTO questionnaire are summarized in Table 5. The mean utility value was calculated for each health state in each group as well as in the combination of both groups. Differences in the mean utility values between groups were not significant in any health state. However, there was a trend for ENT students to consider the health states HS1 and HS2 to be less appealing than the MSc students.

**Table 5: Utility values of health states associated with free tissue transfer to treat head and neck cancer as determined by time trade-off exercises**

	Health State 1		Health State 2		Health State 3		Health State 4	
	value	95% confidence interval	value	95% confidence interval	value	95% confidence interval	value	95% confidence interval
<b>Group 1 (Masters Program)</b>	0.48	0.38-0.57	0.52	0.44-0.61	0.76	0.71-0.81	0.59	0.51-0.67
<b>Group 2 (Department of ENT)</b>	0.32	0.20-0.44	0.32	0.20-0.45	0.78	0.68-0.88	0.52	0.37-0.70
<b>Group 1 and 2</b>	0.43	0.36-0.51	0.47	0.39-0.54	0.76	0.72-0.81	0.57	0.50-0.64

Table 6 provides the correlation matrix between the four health states for each of the groups using Pearson correlation coefficients. This is necessary for the conduct of the probabilistic sensitivity analysis within the economic evaluation.

**Table 6: Pearson Correlation Coefficients between Health States**

	Health State 1	Health State 2	Health State 3	Health State 4
Health State 1	1	0.865	0.189	0.558
Health State 2	0.865	1	0.222	0.695
Health State 3	0.189	0.222	1	0.482
Health State 4	0.558	0.695	0.482	1.000

### 3.5 Discussion

Our assumption was that HS3, the state in which patients have had a good surgical result and have good function of the organs of the UADT, would be rated the most desirable state. The benefit from IDT would be if it increased the number of patients in HS3 and reduced the number of patients in HS2 and HS4 the health states associated with free flap failure. Less IDT failures would also mean fewer patients receiving second operations and thus fewer patients in HS1.

The results of the survey suggest that, as expected, HS3 was the most desirable state. The least desirable state was HS1, the state that exists immediately after surgery. This result also met our expectations.

The health states HS2 and HS4 were similar except for the fact that HS2 was a hospitalized state and HS4 was an outpatient state. Again the survey results met our expectations. The outpatient state was assigned a higher utility value than the inpatient state.

That the observed utility ratings matched our expectations suggests that participants in this utility survey understood the concept of utility and responded appropriately.

Two experts determined the health states used in this survey. The health states in the care of patients undergoing FTT procedures are fairly standardized across institutions but importantly the experts in this study were from the institution from which all costing data was gathered. As a consequence, local expertise on health states was likely the most relevant. Further study could include survey techniques or a Delphi method to better assess health states following FTT.

The two surveyed groups, the group of Otolaryngologists and the group of Master's candidates, ranked the health states similarly. While there were no significantly different results there was a trend for the otolaryngologists to value the inpatient states worse than the MSc students. The reason for this trend is unclear. It is possible that it results from the familiarity that otolaryngologists have with the course of FTT patients in hospital. Hence for the sensitivity analysis, decision analytic model was rerun using the values from each of the samples.

The response rate from the otolaryngologists was poor at 43.8%. As the survey was mentioned at several meetings of the Department of Otolaryngology, it is possible that the attendees of these meetings were more likely to fill out the survey. This subgroup represents physicians more familiar with the care of patients undergoing FTT procedures. While there is the possibility of bias if this subgroup was overrepresented, it would be difficult to guess the direction or the strength of the bias. Certainly, the degree

of bias was not so strong that the results of the survey from the group of otolaryngologists were significantly different than that of the MSc. group.

The utility questionnaire determined the utility values of each of the relevant health states. These utility values were used as inputs in the decision analytic model to determine the benefit of the IDT as a health intervention.

## **4. Costing Study**

### **4.1 Introduction**

A costing study was performed to determine the costs associated with FTT success and FTT failure.

Costing studies can be performed in several different ways and multiple subtleties arise with different costing techniques. Several guidelines exist on the ascertainment of cost data (Drummond et al., 2005). These guidelines agree that as much as possible all relevant costs should be included in cost analysis and that maximum precision should be used for those costs most affecting the outcome of cost-benefit analyses (Barnett, 2009).

The costs considered in health economic analyses are economic costs, sometimes called opportunity costs. This considers the value of the input cost if it were used for another purpose. In most cases, this is the market price of the goods. However, the health care market does not function as perfect market. Consequently, analyses that consider costs from a broad societal perspective must consider the possibility that costs in the health care system are not easily transferable to cost in other sectors. This study adopted the payer perspective of the Ministry of Health. As a result, the opportunity cost is the portion of the payer budget that cannot be spent on other health interventions.

Traditionally, a technique known as microcosting has been used to obtain cost data in studies analyzing the effect of new technologies (Frick, 2009). Microcosting is a recording of all the costs associated with the treatment of a particular patient during a given time horizon. Microcosting is preferred to other costing methods such as gross costing and activity based costing which do provide the same level of cost detail, but may

not resolve cost differences associated with the adoption of novel technology (Barnett, 2009).

Microcosting technique is not standardized in the literature (Frick, 2009). Multiple different sources of data have been used to perform microcosting studies. Frick in his study from 2009 reviews some of the different sources of data including direct observation, insurer databases, item-by-item databases, and time and motion studies (Frick, 2009).

Costing studies using microcosting technique also differ in how fixed costs are treated. Fixed costs can pose daunting challenges to cost studies. Fixed resources are often shared among multiple health programmes and are used to generate multiple health products. Determining the portion of fixed cost associated with each output can be challenging (Barnett, 2009). Some microcosting analyses ignore fixed costs entirely while some may include these costs (Frick, 2009).

#### 4.2 Methods

This costing study was performed using the payer point of view. In this case, the payer was the Ontario Ministry of Health. This perspective was chosen because it was felt that this would be the perspective of decision makers in Canada considering the adoption of IDT technology.

A microcosting technique was used. The microcosting data was captured from the Ottawa Hospital costing database. The Ottawa Hospital is a tertiary care hospital that collects microcosting data. This data includes nursing costs, laboratory costs, food services costs, special care unit costs, pharmacy costs, allied health costs, endoscopy

costs, operating room, and imaging costs for each patient. These costs are available for each day of admission.

The microcosting technique included both fixed and variable costs. Variable costs such as nursing, pharmacy, allied health care costs, were assigned based on the amount of resources consumed by each patient. Fixed costs were then assigned based on the proportion of variable resources each patient consumed.

A consecutive series of patients undergoing free tissue transfer (FTT) procedures from August 2008 until January 2010 was identified using a retrospective review of surgeon case logs. The case logs are maintained for billing purposes and identify all cases performed in a given time period. Patients undergoing free tissue transfers were identified from these lists.

Costing data was then requested from the Ottawa Hospital on this series of patients. Prior to receipt of this data our request was approved by the Ottawa Hospital Research Ethics Board. The cost of the initial operation was not included in our costing analysis as it was not relevant to the costs associated with FTT failure and salvage. Subsequent operations were included in the cost analysis.

An average cost per day was calculated both for successful FTT procedures and failed FTT procedures. This average excluded all costs associated with the first FTT procedure. An average length of stay was calculated for both the FTT successes and the FTT failures.

Physician costs were obtained from the Ontario Ministry of Health Physician Schedule of Benefits as of June, 2010 (Ministry of Health and Long-Term Care). The costs of the implantable Doppler technology (IDT) were obtained from the manufacturer,

Cook Medical. For the purposes of this analysis, the cost of using a single IDT probe was considered.

Costs of operative procedures were calculated by adding the fixed and variable costs of the hospital to the fees billed by the treating surgeons. To determine the fees billed by the treating surgeons the three surgeons trained to perform FTT procedures at the Ottawa Hospital were surveyed.

Home care and outpatient costs were not considered as both successful and unsuccessful FTT procedures require home care assistance often with a negligible difference in resource use over a six-week time horizon. This is because most patients upon discharge, whether they have had successful surgery or not require home care for a period of weeks. Furthermore, it's expected that these outpatient costs would be small in comparison to the inpatient hospital costs.

As the payer perspective was adopted, out-of-pocket expenses and lost-wages were not considered.

All costs were expressed in 2010 Canadian dollars.

#### 4.3 Results

A total of 72 FTT procedures were identified between August 2008 and January 2010. One patient had two FTT procedures. This list of procedures was generated from the case log of the author (MKG) during his fellowship training. The case log was a consecutive list of all procedures performed during this period.

Cost data was available through the Ottawa Hospital on 38 procedures, corresponding to 37 different patients. Five of these patients had failed FTT procedures with one or more revision procedures.

The lack of data on the other 34 procedures was due to the fact that the costing initiative through the Ottawa Hospital is delayed many months after discharge and data on the most recent patients was not available.

The results of the costing study are summarized in Table 7.

The average length of stay for FTT successes was 18.8 days. Average length of stay for FTT failures was 34.5 days and the average length of stay after a successful revision of an FTT procedure was 23.2 days. In this case, a revision procedure referred to a subsequent reconstructive procedure and not simply the salvage of the initial FTT procedure. The average cost per day for FTT successes was \$1423.31. The average cost per day for FTT failures was \$1515.52.

The costs of each procedure are summarized in Table 8. The billing codes used from the Ontario Ministry of Health Schedule of Benefits for Physicians are summarized for each procedure in Table 9. The billing practices of the surgeons performing FTT at the Ottawa Hospital were identical. The assistant fee codes were predicated on eight hours for a fellow and four hours for a staff surgeon who billed one of the major codes. The average cost of a FTT re-exploration for salvage was \$4247.70. The average cost of performing a subsequent FTT procedure to revise a failed FTT was \$9840.07. The average cost of performing a different type of subsequent reconstructive procedure, known as a pedicled flap, to revise a failed FTT was \$3765.95.

The cost of each implantable Doppler probe was \$600. The IDT probes are designed for single-use and thus a new probe is necessary for each patient. The vendors stated that the costs of the IDT monitors and wires could be included in the \$600 cost per probe.

**Table 7: Average Length of Stay and Associated Costs**

	<b>Average Length of Stay (days)</b>	<b>Average Cost per day (2010 Cdn \$)</b>
<b>Successful free tissue transfer</b>	18.8	1423.31
<b>Failed Free tissue transfer</b>	34.5	1515.52

**Table 8: Procedural Costs Associated with Free Tissue Transfer**

<b>Procedure</b>	<b>Average Cost (2010 Cdn \$)</b>
<b>Salvage of Free Tissue Transfer</b>	4247.70
<b>Pedicled Flap Revision of Failed FTT</b>	9840.07
<b>Free Flap Revision of Failed FTT</b>	3765.95
<b>Implantable Doppler</b>	600

**Table 9: Billing Codes Associated with Each Procedure**

<b>Procedure</b>	<b>Ontario Ministry of Health Billing Codes as of June 2010</b>
<b>Salvage of Free Tissue Transfer</b>	R815, R820
<b>Pedicled Flap Revision of Failed FTT</b>	R006
<b>Free Flap Revision of Failed FTT</b>	R055, R066, R064, R055B (8 hours), R066B (4 hours)

#### 4.4 Discussion

This costing study benefitted from the practices of the Ottawa Hospital which performs microcosting on all admissions. Without this information, the costing study would have been very difficult. Unfortunately, as only information from the Ottawa Hospital was used, the external validity of the costing data is limited. While the costs should be similar to other institutions in Ontario, costs may differ across Canada depending on physician practices, fee schedules, and institutional costs. Costs may be significantly different in other countries.

Physician fees represented another threat to the generalizability of the study's findings. Physician compensation varies both within Canada and throughout the world. The codes determining how physicians bill the Ontario Health Insurance Plan are fairly well defined and there should not be significant variation of physician compensation within Ontario.

Another limitation of the costing study was the small numbers of FTT failures included in the study. As FTT failure is a relatively uncommon event, these small numbers were not unexpected. To overcome this limitation when using these data in the decision model, a sensitivity analysis will be performed and a large degree of uncertainty will be attributed to the cost data surrounding failed FTT procedures.

The exclusion of the most recent group of FTT procedures from the costing study was unfortunate as it creates the possibility of a sampling bias. However, the team treating these patients at the beginning of the surveyed time period and at the end was similar. Furthermore, the team existed far along the learning curve making it unlikely the group for which costing information existed was significantly different from the group for which there was no costing information.

Several other authors have performed cost-analyses on patients undergoing FTT for HNC. Most authors report costs and length of stay as mean values despite the skewed distribution of these variables. Two studies were found in the literature that reported median values (Tsue et al., 1997; de Bree et al., 2007). This thesis reported cost and length of stay as a mean value to meet the requirements for economic evaluation through modeling. However, median values from this study are shown in Table 10 with comparison to the other published report. Table 11 compares the mean values of this

study to those published in the literature. Where possible, values for cost of stay were converted to 2012 Canadian dollars using the Bank of Canada inflation calculator (found at <http://www.bankofcanada.ca/rates/related/inflation-calculator>) and the historic currency rate converter at GoCurrency.com (found at <http://www.gocurrency.com>). It is worth noting that there are geographical differences and methodological differences between these studies in addition to differences in currency and time-period.

The cost study determined the costs of hospital admissions and procedures in FTT procedures. These costs will be used as inputs in the economic model along with the previously determined utility values and the data on salvage rates and take-back rates.

**Table 10: Comparison to Median Literature Values**

<b>Study</b>	<b>Median cost per admission</b>	<b>Median cost in 2012 CDN\$</b>	<b>Median Length of Stay</b>	<b>Costing Method</b>	<b>Location</b>
<b>The Ottawa Hospital Cost Study</b>	31 665 (2010 CDN\$)	33 244	15	microcosting	Ottawa, Ontario
<b>Tsue et al., 1997</b>	41 112 (1997 US\$)	Not possible with info from paper	13	Microcosting (hospital charges)	Seattle, Washington
<b>de Bree et al., 2007</b>	23 266 (2003 Euros)	45 728	Not reported	microcosting	Amsterdam, Holland

**Table 11: Comparison to Mean Literature Values**

<b>Study</b>	<b>Mean cost per admission</b>	<b>Cost in 2012 CDN\$</b>	<b>Mean Length of Stay</b>	<b>Costing Method</b>	<b>Location</b>
<b>The Ottawa Hospital Cost Study</b>	44 034 (2010 CDN\$)	46 230	18.8	microcosting	Ottawa, Ontario
<b>Jones et al., 2007</b>	34 300 (US\$ no date given)	Not possible with	16.4	microcosting	Los Angeles, California

		info from paper			
<b>Miller et al., 1991</b>	27 000 (US\$ no date given)	Not possible with info from paper	16.0	Microcosting (hospital charges)	New Orleans, Louisiana
<b>Kroll et al., 1997</b>	28 460 (US\$ no date given)	Not possible with info from paper	13.2	Microcosting and case-mix costing	Houston, Texas

## **5. Health Economic Evaluation**

### 5.1 Introduction

#### 5.1.1 Rationale for Health Economic Evaluation

Decisions on the adoption of new health care interventions and programs usually rest on three types of evaluations. The first type of evaluation is one of efficacy. In the epidemiologic sense, efficacy refers to the effect of an intervention within an experimental setting. An efficacious intervention may or may not be useful outside the confines of an experimental setting. The second type of evaluation is one of effectiveness. In this case, effectiveness is a measure of the interventions effect within the day-to-day clinical world. Most clinical trials provide a combination of efficacy and effectiveness data. Trials that are performed in highly artificial situations are obviously more likely to provide efficacy type data, whereas those that are performed in the standard clinical realm tend to provide more effectiveness data.

Ultimately, adoption of technology requires both efficacy and effectiveness data. However, increasingly decision-makers are requiring economic data in addition to efficacy and effectiveness data prior to instituting new health technologies. In Canada, Provincial Health Ministries face ageing populations and rising health care costs (Canadian Medical Association, 2010). Governments are re-examining how their limited health care resources are spent. Although economic data is still rarely used in the creation of clinical guidelines, it is increasingly being used to guide provincial reimbursement of health programs and interventions (Rocchi et al., 2008)

This thesis aimed to provide the economic data and analysis necessary to guide the decision on adoption of IDT in monitoring FTT procedures to reconstruct defects of

the head and neck. However, any economic evaluation of IDT use in the head and neck suffers from the limited available efficacy and effectiveness data. The author of this thesis decided to perform an economic evaluation of IDT use in spite of this limited data for two main reasons.

First, the available efficacy and effectiveness data is unlikely to improve. The incidence of FTT failure is quite low and even large institutions rarely perform significantly more than one hundred FTT procedures in the head and neck a year. The result is that executing a clinical trial to evaluate the efficacy or effectiveness of IDT is very difficult. A case cohort study has recently been published addressing this question; however, the study sample of this case-cohort included FTT procedures outside the head and neck (Rozen, Chubb, et al., 2010). While new case series may emerge to refine pooled estimates, a systematic review of the available case series with a pooling of failure and salvage rates is likely to represent the best available data for the foreseeable future.

Second, IDT technology has been available for thirty years. Decision makers cannot continue to avoid the question of IDT adoption and wait for better efficacy and effectiveness data. By avoiding making a decision, they may be both denying patients a helpful new health care technology and restricting the use of a cost-saving device.

Techniques exist within health economic evaluation to incorporate the uncertainty that is inherent in the current efficacy and effectiveness data. Such techniques include both deterministic and probabilistic sensitivity analysis. Furthermore, by calculating the expected value of perfect information (EVPI), health economic evaluation can provide information on which parameters within the model contributes most to the inherent

uncertainty over cost effectiveness and therefore will have more value with respect to data generation through further research.

Economic evaluations not only rely on efficacy and effectiveness data, they also rely on cost data and utility data. Uncertainty in these data may be treated in a similar fashion as uncertainty in other data used for economic evaluations.

This project has gathered cost data and utility data through primary research. Utilities associated with FTT procedures were determined via a time-trade off survey. Costs were determined by a microcosting analysis using data from the Ottawa Hospital. The available data on the effect of the intervention, the utility of relevant health states and the cost of the intervention formed the basis for the health economic evaluation

#### *5.1.2 Decision Models in Health Economic Evaluation*

Once the requisite data had been gathered, a decision model was created. Decision models form the backbone of analysis in many health economic evaluations.

According to Drummond et al., a model achieves five main goals: (Drummond et al, 2005). First, the model provides structure to the analysis. As much as possible, the structure of a model is designed to replicate the real-life clinical scenarios. Second, by providing the structure of a real-life clinical scenario to the analysis, the model determines how the available evidence is interpreted. Inputs such as cost data, utility data, and effect data flow in and out of the model in a logical and realistic fashion. Third, once the data is in the model, the model allows evaluation of health interventions in terms of cost and benefit. Fourth, the model also allows for uncertainty in the data. Thus, the results of evaluation of health interventions will include parameters of uncertainty.

Finally, by identifying the greatest sources of uncertainty, a decision model can direct future research.

Not all health economic evaluations use models. Occasionally, a health economic evaluation is combined with a clinical trial. In these cases, the design of the clinical trial forms the structure for health economic evaluation. Critics of decision analytic models have claimed that this method of health economic evaluation is preferable. These critics note that piggybacking an economic evaluation onto a clinical trial avoids the subjectivity that can plague models.

Kassirer and Angell describe these concerns in an editorial for the *New England Journal of Medicine*'s, stating that "the discretionary nature of model building" creates a significant chance for bias in studies. They go on to emphasize the need for models to be "lucid," fearing that the average reader may not grasp intricate models unless they are appropriately described (Kassirer and Angell, 1994).

Kassirer and Angell's points may be valid. Models can certainly be opaque and are vulnerable to significant bias. However, setting an economic evaluation within a clinical trial also has significant disadvantages. Buxton et al., offer a convincing defence of modeling (Buxton et al. 1997). They point out that the data provided by a single clinical trial is often limited. These limitations may take several forms. First, the data associated with a clinical trial may be limited by a relatively short follow-up period. Health economic evaluations may require a much longer time horizon to adequately consider the necessary costs and benefits. Second, clinical trials may gather data in the form of surrogate endpoints, rather than the final endpoints that are of interest to decision

makers determining whether or not new technology should be adopted (Buxton et al 1997).

In addition to the concerns proposed by Buxton et al., the setting of clinical trials is often highly artificial and does not represent the clinical experience. Clinical trials may include additional tests, additional physician or nurse visits, even additional procedures that would not be included in normal practice.

Buxton et al., also identify the most significant benefit of using decision models, namely, that models allow for increased external validity. Models may be used to allow the results of trial to be generalized to a different setting. Models may allow head-to-head comparisons of health interventions that have not been examined by a trial. However, models are still limited by available data. It is difficult to determine results precisely if few studies exist. Yet the use of decision analytic models may be a more rational approach to decision making in the face of limited evidence. As Weinstein et al. claim, the use of models for data extrapolation is less fraught of potential error than extrapolating blindly without the aid of models (Weinstein et al., 1996)

Decision analytic models typically take one of two forms in health economic evaluation. The first form is a decision tree. The second form is a Markov model (Drummond et al., 2005). These models are designed to replicate how patients move from one health state to another over a period of time. The movement from one health state to another is governed by transition probabilities. These probabilities may be based on primary research or data from secondary sources.

The decision tree form is designed to model forward, stepwise progressions. In a decision tree model, multiple pathways or branches are created to represent possible

outcomes. The decision tree model does not easily provide for movement backwards to previous health states. Furthermore, time is not a variable in decision tree models. Instead, the events of the model occur instantaneously. The Markov model is more fluid. It allows for movement between health states in any direction and also allows for the possibility of remaining stagnant within a health state.

The Markov model is able to overcome the limitations of a decision tree model in part because it considers time as an explicit variable (Drummond et al., 2005). Time exists in discrete units, or cycles in a Markov model. With each cycle, there is movement between health states. This movement is determined by probabilities which determine the chance of patients moving between health care states. These are often referred to as transition probabilities.

### *5.1.3 Discount Rate*

Another consideration of economic models is the discount rate. A discount rate is the factor by which future costs are reduced. This factor represents a time-preference, specifically the preference to defer costs to the future. The discount rate must of course be applied to any benefits as well as any costs, otherwise decision makers would defer spending on health programs indefinitely. Studies commonly choose discount rates close to the going interest rate. However, some authors have argued that the logic behind both discounting and not discounting costs and benefits is shaky. These authors recommend all studies contain a sensitivity analysis that considers multiple different discount rates including rates equal to zero (Coyle and Tolley, 1992).

### *5.1.4 Sensitivity Analysis*

Health economic evaluation is unique among epidemiologic studies as uncertainty arises in multiple ways and is managed differently. Among the first to characterize the different sources of uncertainty and devise an approach to handling such uncertainty were Briggs, Sculpher, and Buxton in 1994 (Adronis, Barton, Bryan, 2009).

They identified four main types of uncertainty in economic evaluations carried out alongside clinical trials. The four sources of uncertainty were: (1) uncertainty in sample data; (2) uncertainty in generalization of results beyond the trial; (3) uncertainty in extrapolating from intermediate end-points to final end-points; and (4) uncertainty related to the analytic methods used to analyze data (Briggs, Sculpher, Buxton, 1994).

A different classification of uncertainty has been proposed for uncertainty in model-based economic evaluation (Adronis, Barton, Bryan, 2009). Manning et al. identified three types of uncertainty in model-based analyses. These three types were: (1) parameter uncertainty; (2) methodologic uncertainty; and (3) structural Uncertainty (Manning, Fryback, Weinstein, 2002).

Parameter uncertainty is the uncertainty associated with the inputs in the model. Usually, the inputs are the effectiveness of interventions, utility of health states, and other data that are used to populate the model. Uncertainty in this case arises typically from variability in the sample data. Methodologic uncertainty refers to the uncertainty surrounding the economic methods used to create the decision analytic model. Adronis, Barton, and Bryan refer to the method of adjusting for time preference as an example of methodological uncertainty (Adronis, Barton, Bryan, 2009). Structural uncertainty refers to uncertainty in how to best design a decision analytic model (Briggs, 2000).

Sensitivity analyses are the primary method used by analysts to account for the uncertainty in economic evaluations (Drummond, et al., 2005). Sensitivity analyses may take two forms, deterministic or probabilistic. The deterministic form varies model inputs across a group of values that is chosen by the analyst. The probabilistic form uses various probability distributions to create a range of possible values for each model input.

The probabilistic approach has several steps. First, each of the input variables is assigned a distribution. This distribution is chosen such that it is consistent with the uncertainty surrounding the parameter. Values that are illogical, such as negative should not be part of the chosen distribution. Furthermore, the distribution should be representative of the expected range of values for a given parameter. As an example, an input based on length of hospital stay would be expected to have an asymmetrical distribution with a long tail. For such an input, a gamma distribution would be chosen. Once the specific type of distribution has been chosen, the mean and variance of sample data can be used to create the distribution representing the population.

Once the distribution is formed, the probabilistic approach then randomly samples from this distribution and uses this sample as an input parameter in the model. A sensitivity analyses may randomly sample from a single parameter distribution while leaving the values for other input parameters fixed or it may sample from multiple parameter distributions concurrently.

The results of a decision analytic model are recalculated multiple times, usually thousands of times, based on different samples from the distributions. The results of these replications are then averaged to refine the initial result of the decision analytic model. The replications also define a range of possible outcomes that are useful in

further analyses such as the expected value of perfect information analysis and the cost effectiveness acceptability curve.

The number of replications of the model is somewhat arbitrary however most authors have used several thousand replications.

Authors have tended to encourage the use of probabilistic sensitivity analyses (PSA) over deterministic sensitivity analyses (DSA) on the basis that the PSA method may be more comprehensive and less arbitrary (Drummond et al., 2005).

The most recent Guidelines for the Economic Evaluation of Health Technologies released by the Canadian Agency for Drugs and Technology in health, recommended that at minimum a DSA is performed as part of the analysis of uncertainty. However, they encourage the use of PSA with a Monte-Carlo simulation. The guidelines encourage the use of value of information analysis to identify key areas of uncertainty in the model inputs (Canadian Agency for Drugs and Technology in Health, 2006). In a recent review of methods for economic evaluation, Claxton and others also encourage the use of PSA methods (Claxton et al, 2005).

However, detractors of the PSA method have stated that PSA only provides the illusion of objectivity. They claim that ultimately the choice of which probability distribution to employ is arbitrary. Briggs notes that such a criticism fails to recognize that the choice of the probability distribution is constrained by the type of data that is being represented; and that ultimately, an analyst will have little choice on the distribution if he wishes it to be relevant to the data (Briggs, 2005).

Often times, several strategies are used to confront the various types of uncertainty inherent in an economic evaluation. Probabilistic sensitivity analysis has

become the primary method used to handle parameter uncertainty. Methodologic uncertainty is usually treated with deterministic approaches. In the most common case, several different discount rates are arbitrarily chosen and the analysis is run with each of these discount rates.

Structural uncertainty is much more difficult to handle. Some authors have suggested that several different models should be created, assigned different probabilities of being true, and then analyzed, to account for structural uncertainty (Manning 2002; Adronis, Barton, Bryan, 2009). However, such analyses are extremely cumbersome and deciding on the relative probabilities of different models is practically impossible. Ultimately, structural uncertainty may be best resolved by comparing the results of different investigators who have researched the same economic question (Briggs, 2000).

#### *5.1.5 Cost Effectiveness Acceptability Curves and Expected Value of Perfect Information Analysis*

Once the decision analysis model has yielded a result, two additional forms of analysis help decision makers interpret this result.

The first is the creation of a cost effectiveness acceptability curve (CEAC). A CEAC is an analysis designed to help decision makers determine whether alternative interventions are cost effective (van Hout, et al., 1994). It plots the probability of an intervention being cost-acceptable as a function of the threshold willingness-to-pay (the threshold willingness-to-pay is the maximum amount a decision maker would be willing to pay for a single unit of benefit, usually the amount one would be willing to pay for one quality-adjusted life-year). A CEAC gives a visual representation of the probability of each alternative intervention being optimal for a given threshold willingness-to-pay. The

curve also depicts how the optimal intervention may vary with the threshold willingness-to-pay.

A CEAC is calculated using the replications created by the PSA. For each possible threshold value for a quality-adjusted life-year (QALY), the number of replications in which an intervention is most cost effective is calculated. From this, the probability that an intervention is optimal is derived. These probabilities are plotted for the corresponding threshold willingness-to-pay, creating the curve.

The second is a value of information (VOI) analysis. The VOI analysis is designed to achieve two primary goals: (1) identify the consequences of choosing the incorrect intervention; (2) identify the areas of greatest uncertainty in an economic evaluation (Coyle and Oakely 2008; Felli and Hazen, 1998).

Ultimately, decision makers must choose an intervention to adopt from imperfect information. The results of a probabilistic analysis along with a cost-effectiveness acceptability curve determine which intervention is most likely to be cost-effective. However, decision makers may wish to know the consequences of choosing the wrong alternative. The VOI analysis quantifies the costs of an incorrect decision. By identifying the costs of an incorrect decision, the value of perfect information can be obtained.

The VOI analysis, like the CEAC, uses the replications created by the PSA to estimate the Expected Value of Perfect Information (EVPI).

The steps taken to estimate the EVPI are as follows: (Coyle and Oakley, 2008)

1. Determine the optimal intervention based on the results of all replications.
2. For each replication estimate the net benefit (NB) for each intervention

$$\text{NB} = \text{QALYs} * \text{value of a QALY} - \text{Cost}$$

3. For each replication estimate the difference between the maximal NB and the NB of the optimal treatment
4. EVPI is the average of the values in step 3

A value of information analysis can be performed on all input parameters or it can be performed on a single input parameter. Performing this analysis on a subset of the input parameters is called an expected value of partial perfect information analysis, or EVPPI analysis (Coyle and Oakely, 2007).

The significance of the EVPPI analysis is that it identifies input parameters that contribute most significantly to the uncertainty in the model. Consequently, the EVPPI may direct future research to areas that are most likely to aid decisions on adoption of technology (Briggs, Claxton, Sculpher, 2006).

## 5.2 Methods

### *5.2.1 Model Design*

Several different models were attempted to replicate the clinical course of free tissue transfer (FTT) procedures for head and neck cancer (HNC). Initially, a decision tree-type model was tried as it was felt that a decision tree model would offer a simple, yet accurate representation of patients' progression following FTT procedures. However, given the possibility of repeated FTT failure, the decision tree-model became unworkable. A Markov-type model offered a better solution to the modeling problem. Microsoft Excel Vista (Microsoft Corporation, Redmond WA) was used to create the model and for the conduct of all relevant analyses of the model.

In the model, patients existed in four health states:

1. Admitted to hospital, recovering from an FTT procedure that is currently successful.
2. Admitted to hospital, recovering from an FTT procedure that is a failure and awaiting treatment for the failure.
3. At home, recovering from an FTT procedure that was successful.
4. At home, recovering from an FTT procedure that was a failure.

These health states were identical to the health states used in the utility survey.

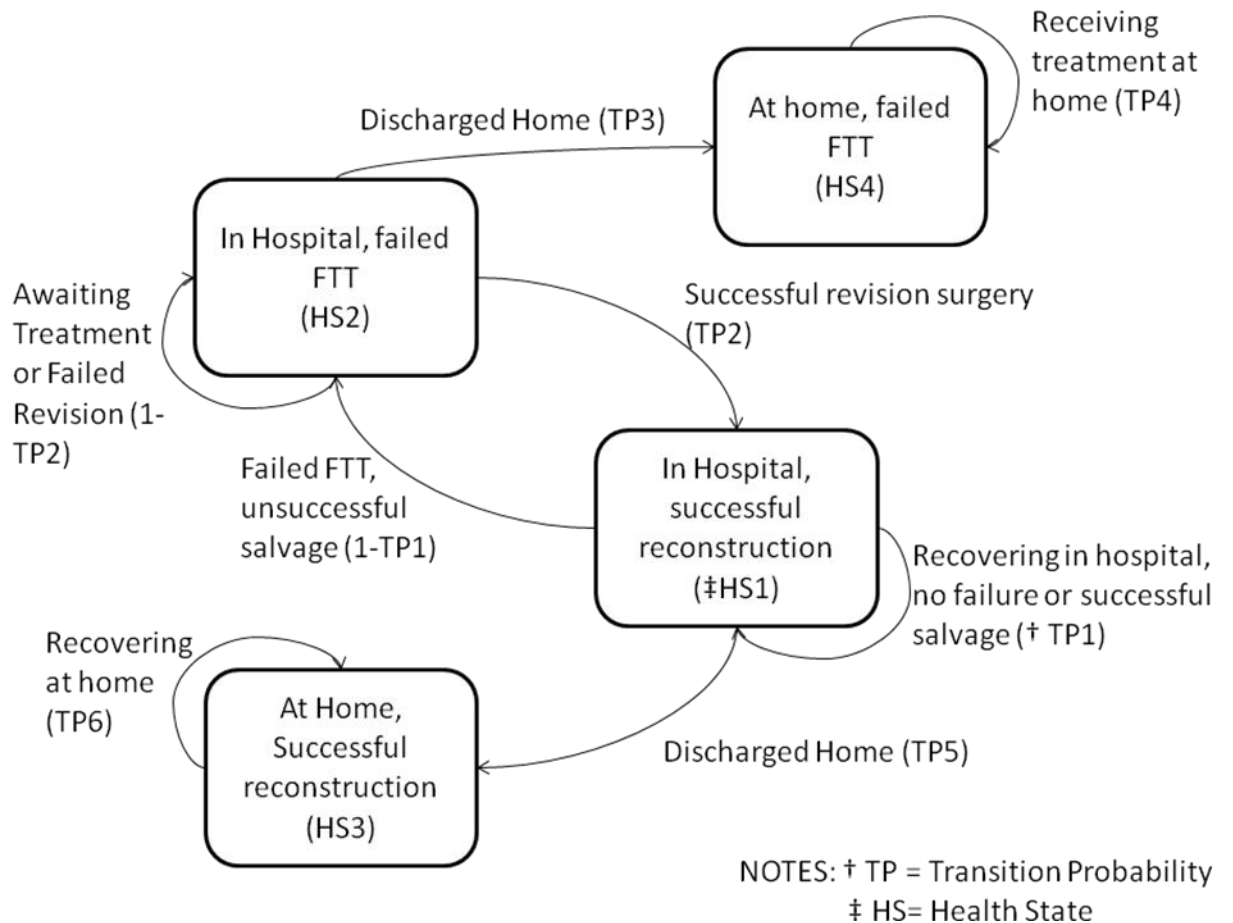
All patients started, in the first state, in hospital recovering from an FTT procedure that is currently successful. Patients who showed signs of FTT failure were returned to the operating room for re-exploration of the FTT. Those that were successfully salvaged were returned to the initial state. Patients that did not develop failure were eventually discharged home from hospital and landed in the third state, at home recovering from an FTT procedure.

Patients whose FTT procedure could not be salvaged slipped into the second state. They are admitted to hospital recovering from an FTT procedure that is a failure and awaiting treatment. Patients who developed failure and were in the second state had the potential to remain in that state awaiting a revision surgery to treat the failure. At this point is important to note the difference between a revision procedure and a salvage procedure. A salvage procedure refers to a procedure done in an attempt to save a failing FTT. A revision procedure refers to procedure performed after the FTT has completely failed. The goal of a revision procedure is to remove the failed FTT and reconstruct the defect initially created by extirpation of the cancer. These patients could undergo a successful revision surgery and return to the first state. These patients could also be





















discharged home and receive non-operative treatment for the FTT failure thus entering the last state, at home recovering from a failed FTT procedure.

The model is represented in pictorial form in Figure 4. The parameter inputs in the model are summarized in Table 12.






**Figure 4: Model Format**



**Table 12: Summary of Parameter Inputs**

Parameter	Value
<b>Conventional take back rate</b>	0.065 
<b>Conventional salvage rate</b>	0.51 
<b>Implantable Doppler take back rate</b>	0.10 
<b>Implantable Doppler salvage rate</b>	0.85 
<b>Proportion of take back in first 3 days</b>	0.95 
<b>Proportion of failures treated with granulation</b>	0.25 
<b>Proportion of revised failures treated with pedicled flaps</b>	0.67 
<b>Proportion of revised failures treated with FTT</b>	0.33 
<b>Length of Stay Success</b>	18.8 days 
<b>Length of Stay Failure</b>	34.5 days 
<b>Length of Stay after Revision</b>	23.2 days 
<b>Costs per Day Success</b>	C\$1423.31 
<b>Costs per Day Failure</b>	C\$1515.52 
<b>Cost of Pedicled flap revision</b>	C\$3765.95 
<b>Costs of FTT revision</b>	C\$9840.07 
<b>Cost of Salvage</b>	C\$4247.70 
<b>Utility of Health State 1</b>	0.46 
<b>Utility of Health State 2</b>	0.50 
<b>Utility of Health State 3</b>	0.74 
<b>Utility of Health State 4</b>	0.56 

**Legend:**

-  Derived from Table 3
-  Please see Appendix C
-  Derived from Table 7
-  Derived from Table 8
-  Derived from Table 5

*5.2.2 Assignment of Transition Probabilities in the Model*

The transition probabilities were assigned as follows (please note the numbering system refers to that used in Figure 4 above):

**TP1:** This transition probability refers to the probability of FTT success. It is a combination of two probabilities, both determined from the systematic review performed

as part of this project. The first is the probability of re-exploring a failing FTT and the second is the probability of successful salvage of a failing FTT.

The TP1 transition probability was varied with the post-operative day. Most authors recognize that FTT failure happens predominantly in the first few days after FTT procedures (Novakovic et al., 2009). Failure after more than a week is rare and when it occurs is uniformly unsalvageable (Yu et al, 2009). A separate TP1 was determined for the first three days after FTT, for the following seven days after FTT, and for the remaining days in hospital after FTT. Ninety five percent of re-explorations and failures occurred in the first three post-operative days. The remaining five percent of failures occurred in the following seven days. No failures occurred more than ten days after the procedure.

This division into three time periods was based partially on the expert opinion of surgeons at the Ottawa Hospital and partially on published case series. There has been no publication in the literature that has thoroughly examined failure rates of FTT as a function of time after surgery. Given the lack of experimental data for this division, it was given very wide confidence intervals in the sensitivity analysis.

**TP2:** This transition probability represented the probability of successful revision of a failed FTT. Like TP1, it was also a combination of other probabilities. The components of TP2 were the average wait time for revision, the relative use of different treatments for failed FTT and the success rates of those treatments. As the revision of a failed FTT is rare, there is very little evidence on which to base TP2. Instead, expert opinion was sought. The lack of hard data meant that during the sensitivity analysis wide confidence intervals were used.

The average wait time for revision is difficult to estimate. It depends greatly on the surgeons and the institution performing the FTT procedures. Most Canadian centres have limited operating room resources; consequently, a wait of several days is usually required before a FTT reconstruction can be revised. These revisions are essentially secondary reconstructions where the tissue lost to cancer surgery is reconstituted a second time. Secondary reconstructions are relatively rare as FTT failure is relatively rare. Most of these cases are booked as less urgent emergency cases. Expert opinion from FTT surgeons was sought to determine the average wait. The consensus of free flap surgeons at the Ottawa Hospital was that the average wait was three days.

Three different options are available to surgeons seeking to revise a failed FTT. These include a repeat FTT, another type of repair known as a regional flap, or removing the free tissue and allowing the area to heal on its own. Each institution and each surgeon has their own method of deciding between these options. At the Ottawa Hospital most secondary reconstructions are achieved with regional flaps. Less often, a second FTT is used or the patient is allowed to self-heal the area. Again there is very little data to breakdown the relative numbers in each category. A survey of the local surgeons determined that 50 percent of patients underwent reconstruction with a regional flap, 25 percent were reconstructed with a FTT, and 25 percent were allowed to heal on their own.

Failure of secondary reconstructions is another possibility that needed to be considered in calculating TP2. The failure rate of regional flaps is essential zero. The failure of secondary FTT procedures has not been well described in the literature. Most

experts agree that it is higher than the failure rate of primary reconstruction. Again local expert opinion was used to choose a secondary failure rate of 20%.

**TP3:** This transition probability refers to the probability of discharging a patient with a failed FTT home. For patients in the model to face this transition probability they must have not received a revision procedure. This is incorporated in the model by employing transition probabilities which allow replication of the average length of stay. Sensitivity analysis adopts alternative values for the length of stay.

**TP4:** This transition probability refers to the probability of discharged patients who had a failed FTT staying in their current health state. TP4 always has a value of 1 reflecting that this is an absorbing state, once a patient enters it there are no further transitions within the time horizons of the model.

**TP5:** This is the transition probability of discharging a patient home who received a successful FTT home. Similar to TP3, this is incorporated in the model by employing transition probabilities which allow replication of the average length of stay. Sensitivity analysis adopts alternative values for the length of stay.

**TP6:** This transition probability is the probability that patients in the model who are home after a successful FTT procedure stay in their health state. This transition probability is always equal to 1, again reflecting that this is an absorbing state.

Appendix D displays sample calculation of transition probabilities TP1, TP2, TP3, and TP5 using the actual values.

### *5.2.3 Assignment of Costs in the Model*

The two in-hospital health states had costs assigned. These costs were determined by the costing study carried out as part of this project. These costs are summarized in

Table 7. No costs were assigned to the health states that existed outside the hospital. These costs, primarily home-care costs, were negligible relative to other costs in the model and as they are similar in both successful and failed FTT procedures they are unlikely to change the outcome of the evaluation.

Procedural costs were assigned to the transitions between health states. The procedural costs were summarized in Table 8. Revision procedure costs were assigned to the revision procedures (see the assumptions section for which revision procedures were used). Costs from re-exploration of FTT were assigned when patients were returned to the operating room for exploration of the FTT. All patients in the IDT group had the additional cost of the IDT probe.

Data on costs and hospital admissions typically do not have a normal distribution. Instead, they typically have long tails. Certainly, this was the case in the data gathered from the Ottawa hospital on costs and length of stay. The mean of this type of data is usually not as representative of the data as a median is. However, means were reported, as medians cannot be used in the creation of economic models. To adjust for the non-normal distribution of the data in the model, a probabilistic sensitivity analysis was devised that used gamma distributions for data on costs and hospital stays.

#### *5.2.4 Assignment of Length of Stay in the Model*

The length of stay in hospital after both successful FTT procedures and unsuccessful FTT procedures was determined from the averages in the costing study done at the University of Ottawa. After unsuccessful FTT procedures two values were considered. The first was the average length of stay after a successful revision procedure. The second was the total average length of patients after failed FTT. Patients in the

model who had a successful revision stayed the average number of post-operative days after a successful revision and were then discharged home. Patients in the model who did not receive a successful revision procedure were discharged home once they reached the average number of days a patient with a failed FTT stayed in hospital.

#### *5.2.5 Assignment of Utilities in the Model*

The four health states in the model (HS1, HS2, HS3, HS4) corresponded to the four health states in the utility survey. The results of the utility survey were used to determine the utilities each of the health states in the model. However, the value for HS1 was only used for the first post-operative day, when patients most mirrored the state described in HS1. As the patients continued in HS1 their utility increased daily towards the value in HS3. This incremental increase was determined by taking the difference of utility between HS3 and HS1 and divided by the average number days spent in hospital after a successful FTT procedure.

#### *5.2.6 Time Horizon*

The time horizon in this economic model was six weeks.

#### *5.2.7 Discount Rate in the Model*

As the time horizon was very short in our study, no discount rate was used.

#### *5.2.8 Assumptions*

The assumptions used to design this project's decision analytic model are listed.

1. The model assumed that any FTT showing signs of failure were returned to the operating room for re-exploration of FTT vessels and attempted salvage.

2. The model assumed that as patients healed from surgery and prepared for discharge home, the utility value of their health state improved in an additive fashion.
3. The model assumed that patients whose FTT procedures were successfully salvaged returned to the same post-operative day they were re-explored. They did not return to the first post-operative day as if they had had the entire procedure to remove their cancer and reconstruct lost tissue with an FTT procedure.
4. The model assumed that patients who were monitored by IDT techniques were also monitored by conventional techniques.
5. The model assumed that a single Doppler probe was used for each patient.

#### *5.2.9 Sensitivity Analysis*

A PSA was performed as part of this project to account for parameter uncertainty. All parameters in the model were subject to this analysis. Appropriate distributions were chosen for each parameter.

The transition probabilities in the model were varied in the sensitivity analysis according to beta distributions. Cost and length of stay inputs were subjected to gamma distributions. Uncertainty around utilities was characterized by a distribution taking the form of 1-lognormal distribution. Utilities were first converted to disutilities ( $1 - \text{utility}$ ). The distributions of disutilities were then subjected to lognormal distributions which were then converted back to utility values ( $1 - \text{disutility}$ ). The parameters and the associated distributions are summarized in Table 13.

The choice of parameter distribution was based on the expected distributions and the limits of the parameter. The framework provided by Briggs, Claxton and Sculpher was also used as a guide to the choice of parameters (Briggs, Claxton, Sculpher, 2011).

To create the beta distributions for the take back rate and the salvage rate, the point estimate and the confidence limits for the rates from the systematic review were used. The methods of moments estimation technique was then used to create the distribution (Pratt, Raiffa, Schlaiffer, 1995). For other transition probabilities where standard errors were unknown, a conservative sample size ( $n=25$ ) was adopted. For a summary of these distributions please see Table 13.

For costs and length of stay parameters, a gamma distribution was used. In this case point estimates and standard error were determined from the costing project at the Ottawa Hospital. When there were relative few data on the costs, for example the cost of free flap salvage, a larger standard error was adopted which was twenty-five percent of the point estimate.

All utility parameters were varied according to a 1-lognormal distribution in the PSA. The point estimates and standard errors were taken from the combined results of the utility survey. As individuals provided estimates for utility values for all health states in the model, estimates are clearly correlated. To incorporate the correlation between the point estimations for utility values the Cholesky decomposition was used (Stollenwerk et al., 2010; Teerawattananon et al. 2007, Daniels and Zhao, 2003).

The Cholesky decomposition method allows recognition that a proportion of the uncertainty around the utility values is explained by the uncertainty around other health states' values. New estimates for the utility values are derived as a function of the

original estimate of utility and an adjustment factor which is a function of the covariance matrix and random values from a standard normal distribution.

**Table 13: Parameter Distributions used in the Probabilistic Sensitivity Analysis**

<b>Beta Distributions of Model Parameters</b>				
<b>Parameter</b>	<b>Point Estimate</b>	<b>Standard Error</b>	<b>Alpha</b>	<b>Beta</b>
<b>Conventional take back rate</b>	0.065	0.010	37.9	544.8
<b>Conventional salvage rate</b>	0.51	0.061	33.5	32.2
<b>Implantable Doppler take back rate</b>	0.10	0.013	57.5	517.5
<b>Implantable Doppler salvage rate</b>	0.85	0.045	51.7	9.13
<b>Proportion of take back in first 3 days</b>	0.95	0.043	23.75	1.25
<b>Proportion of failures treated with granulation</b>	0.25	0.085	6.25	18.75
<b>Proportion of revised failures treated with pedicled flaps</b>	0.67	0.092	16.75	8.25
<b>Proportion of revised failures treated with FTT</b>	0.33	0.092	8.25	16.75
<b>Gamma Distributions of Model Parameters</b>				
<b>Parameter</b>	<b>Point Estimate</b>	<b>Standard Error</b>	<b>Shape</b>	<b>Scale</b>
<b>Length of Stay Success</b>	18.8	2.29	67.16	0.280
<b>Length of Stay Failure</b>	34.5	8.23	17.55	1.97
<b>Length of Stay after Revision</b>	23.2	3.53	43.32	0.536
<b>Costs per Day Success</b>	1423.31	86.32	271.90	5.23
<b>Costs per Day Failure</b>	1515.52	332.97	20.71	73.16
<b>Cost of Pedicled flap revision</b>	3765.95	851.89	19.54	192.7
<b>Costs of FTT revision</b>	9840.07	2460.02	16	615.0
<b>Cost of Salvage</b>	4247.70	1061.93	16	265.5
<b>1-Lognormal Distributions of Model Parameters</b>				
<b>Parameter</b>	<b>Point Estimate</b>	<b>Standard Error</b>	<b>Exp Mean</b>	<b>Exp SE</b>
<b>Utility of Health State 1</b>	0.46	0.039	-0.59	0.071
<b>Utility of Health State 2</b>	0.50	0.038	-0.63	0.071
<b>Utility of Health State 3</b>	0.74	0.023	-1.47	0.099
<b>Utility of Health State 4</b>	0.56	0.036	-0.85	0.084

In addition to the PSA, a DSA was performed by increasing the number of Doppler probes used per patient from one to two. The PSA was then repeated with this changed parameter. The utility values from each of the two survey groups from the utility questionnaire were also used and the analysis was repeated.

#### *5.2.10 Cost-Effectiveness Acceptability Curve*

As part of the analysis of the results of this model a CEAC was calculated using Microsoft Excel for Vista (Microsoft Corporation, Redmond, WA) using the methods detailed in Section 5.1.5.

#### *5.2.11 Estimate of Expected Value of Perfect Information*

In this project, VOI analysis was performed using a threshold willingness-to-pay of C\$50 000. The expected value of partial perfect information analysis was performed using the one-staged Monte Carlo simulation method detailed in section 5.1.5 (Coyle and Oakley, 2008).

### 5.3 Results

#### *5.3.1 Results of Deterministic Model*

The decision analytic model was run first without a probabilistic sensitivity analysis. The deterministic model considered 1000 patients. The results were based on the parameter inputs without any confidence intervals. These results are summarized in Table 14. In the deterministic model, the addition of IDT to the monitoring of FTT procedures post-operatively reduced cost and improved patient benefit. Use of IDT had a savings of C\$300.05 and an increase in utility of 0.0004 quality-adjusted life-years (QALYs). This resulted in the IDT intervention being dominant – less costly and more effective.

**Table 14: Results of Deterministic Model**

	Clinical Monitoring	IDT	Incremental Analysis
Cost (Canadian Dollars)	28701.23	28401.18	-300.05
Quality-Adjusted Life-Years Gained (QALY)	0.0800	0.0804	0.0004
Cost Utility Ratio (Canadian Dollars/QALY)			IDT Monitoring Dominant

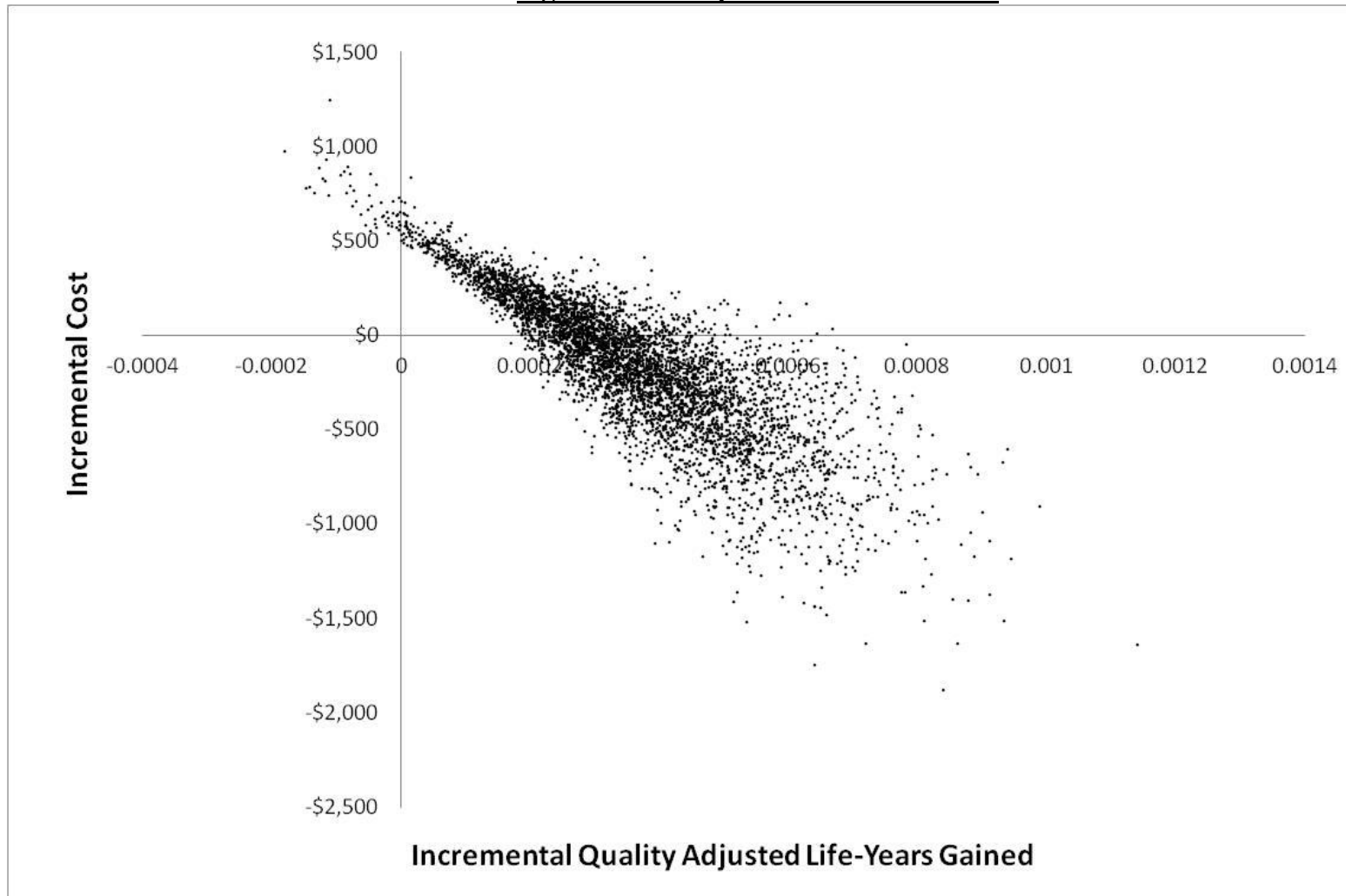
*5.3.2 Results of Model with Probabilistic Sensitivity Analysis*

The decision analytic model was then run 5000 times using the PSA method. The average results for difference in cost, difference in quality-adjusted life-years and incremental cost-effectiveness ratio are report in Table 15. Again the IDT intervention was dominant. With the use of IDT, an average cost savings of C\$160.55 was found per patient. Use of IDT was also associated with an average benefit of 0.0003 QALYs. A scatter plot of the results of the 5000 replications from the probabilistic analysis is presented in Figure 5. In 99% of replications, the IDT intervention generated more QALYs. In 66% of replications, it was less costly. In 66% of replications it was both less costly and generated more QALYs.

**Table 15: Results of Probabilistic Analysis**

Averages from 5000 cycles of Probabilistic Model	Clinical Monitoring	IDT	Incremental Analysis
Cost (Canadian Dollars)	28304.46	28143.91	-160.55
Quality-Adjusted Life-Years	0.0801	0.0804	0.0003
Cost Utility Ratio (Canadian Dollars/Quality-Adjusted Life-Years)			IDT Monitoring Dominant

**Figure 5: Scatterplot Derived From PSA**



The results of performing the analysis assuming two implantable probes per patient are summarized in Table 16. In this analysis, IDT use was no longer dominant. While an average benefit of 0.0003 QALYs was associated with the IDT use, there was an average increased cost of C\$449.98 per patient. This resulted in a large ICER of C\$1249000.

**Table 16: Results of Probabilistic Analysis with Two Probes**

Averages from 5000 replications within the PSA	Clinical Monitoring	IDT	Incremental Analysis Results
Cost (Canadian Dollars)	28 291.92	28 741.90	449.98
Quality-Adjusted Life-Years	0.0801	0.0805	0.0003
Cost Utility Ratio (Canadian Dollars/Quality-Adjusted Life-Years)			1249000

The results of repeating the analysis using the utility values derived from each of the separate groups in the utility survey are presented in Tables 17 and 18. In both analyses, the IDT group remains dominant.

**Table 17: Results of Probabilistic Analysis using Utility Values Derived from Masters in Epidemiology Group**

Averages from 5000 replications within the PSA	Clinical Monitoring	IDT	Probabilistic Model Results
Cost (Canadian Dollars)	28 195.08	28 031.73	-163.36
Quality-Adjusted Life-Years	0.0801	0.0805	0.0004
Cost Utility Ratio (Canadian Dollars/Quality-Adjusted Life-Years)			IDT Dominant

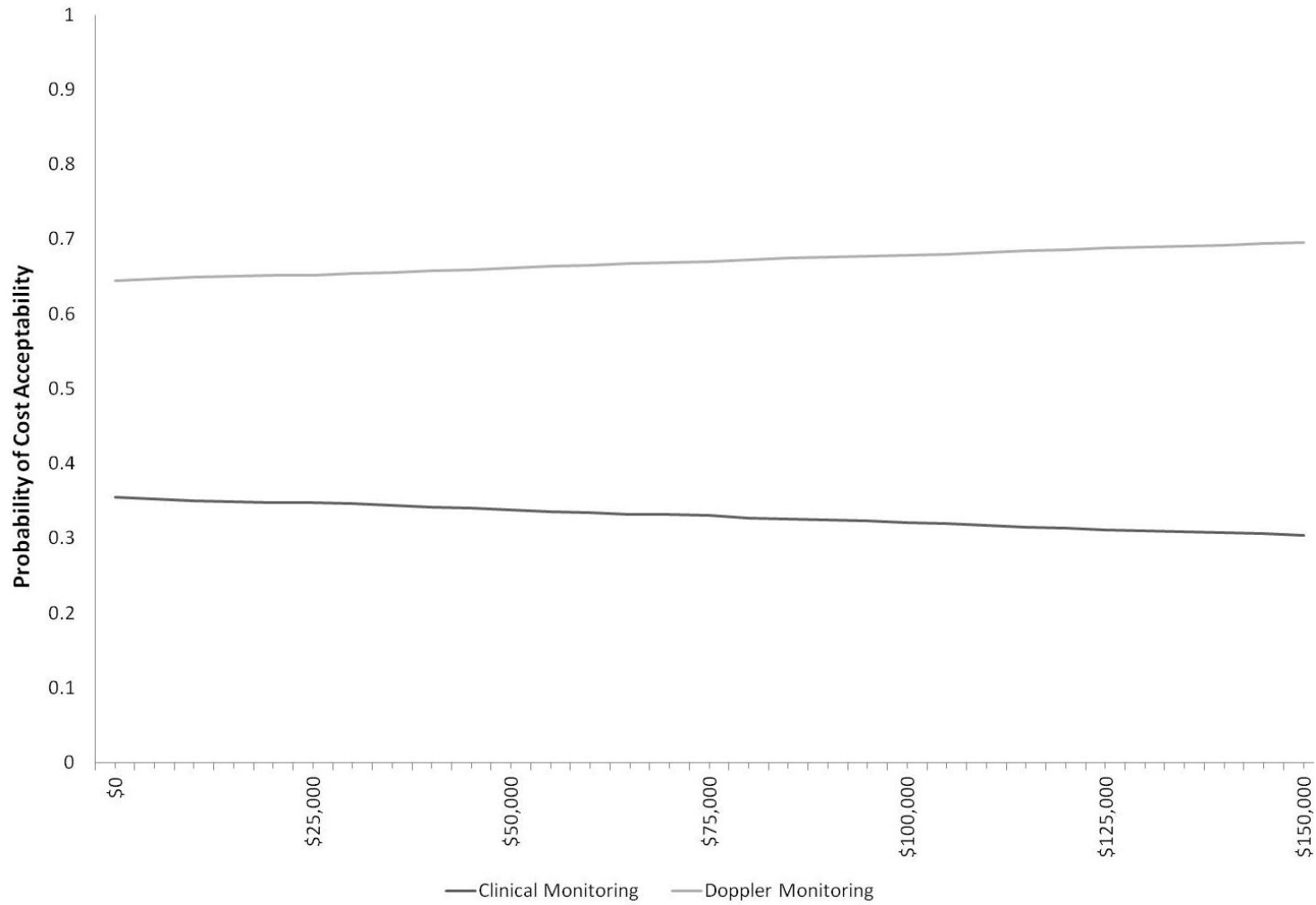
**Table 18: Results of Probabilistic Analysis using Utility Values Derived from  
Department of Otolaryngology Group**

Averages from 5000 replications within the PSA	Clinical Monitoring	IDT	Probabilistic Model Results
Cost (Canadian Dollars)	28 195.09	28 031.73	-166.85
Quality-Adjusted Life-Years	0.0801	0.0805	0.0004
Cost Utility Ratio (Canadian Dollars/Quality-Adjusted Life-Years)			IDT Dominant

*5.3.3 Cost-Effectiveness Acceptability*

A CEAC is displayed in Figure 6. At all threshold levels of willingness-to-pay the monitoring arm employing IDT has a greater probability of being cost-effective than the arm without IDT monitoring. For all threshold values between \$0 and \$150,000, the probability of cost-effectiveness is more than 60% but less than 70%. This highlights the degree of uncertainty over the results.

**Figure 6: Probability of Cost-Effectiveness Acceptability**



### 5.3.4 Value of Information Analysis

The estimated EVPI is displayed in Table 19. The estimates of EVPPI are displayed in Table 20. The majority of the uncertainty in the analysis was found in the input parameters derived from the systematic review, the take back rates and the salvage rates. While the value of perfect information was C\$70.98/patient, the value of partial perfect information for the take back rates and salvage rates was C\$60.61/patient.

The take back rates and salvage rates of each group were further analyzed. These are displayed in Table 21. The results suggest there is nearly equal value to obtaining further information for both IDT and conventional monitoring.

**Table 19: Estimate of Expected Value of Perfect Information**

Input Parameters Analyzed	Value of Perfect Information (C\$/patient)
All Inputs	70.98

**Table 20: Estimates of Expected Value of Partial Perfect Information**

Input Parameters Analyzed	Value of Partial Perfect Information (C\$/patient)
Inputs from Systematic Review: take back rates and salvage rates	60.61
Expert Opinion Values: failure rate in first 3 days, failure rate of revision FTT, proportion of revision procedures, wait time for revision	0.06
Cost and Length of Stay Inputs	8.43
Utility Inputs	0.00

**Table 21: Estimates of Expected Value of Partial Perfect Information of Take Back Rates and Salvage Rates**

Input Parameter Analyzed	Value of Partial Perfect Information (C\$/patient)
Take Back Rate and Salvage Rate in Clinical Monitoring Group	34.28
Take Back Rate and Salvage Rate in IDT Group	30.02

5.4 Discussion

*5.4.1 Model Assumptions*

The model assumptions were based on the clinical practices at the Ottawa Hospital; however, the assumptions may be generalized to other centres.

The first assumption was that FTT procedures showing signs of failure are returned to the operating room for re-exploration. Not all surgeons will attempt to salvage all failing FTT procedures. However, if surgeons are unwilling to attempt to salvage failing FTT procedures, the question of how best to monitor FTT procedures is moot.

The second assumption was that patients recovering from FTT have a utility that improves in an additive manner each day. This assumption was based on the clinical experience that as a patient recovers from surgery and slowly requires less pain medication, and no longer has the need for breathing or feeding tubes, they are slowly more pleased with their health state.

While the assumption that the quality of life of patients recovering from FTT surgery improves each day is certainly reasonable, whether this occurs in an additive manner is difficult to know. Although it is unlikely that the function by which patients'

utility scores improve will significantly affect the outcome, as the EVPI analysis indicated that utility scores are not a significant cause of uncertainty in the model. Future analysts may wish to include a sensitivity analysis with a different functions modeling post-operative recovery.

The third assumption held that patients whose FTT procedure is successfully salvaged return to their pre-failure state and not post-operative day 1. This was based on the expert opinion that most post-surgical healing is not significantly delayed by a few hours of tenuous blood supply to a FTT.

The fourth assumption was that FTT procedures in the IDT arm were simultaneously monitored by both clinical and IDT methods. This is a fair assumption as long as the FTT has been placed in a location which renders standard clinical methods unfeasible. Free tissue transfer reconstruction of the cervical esophagus after laryngectomy, is an example of a situation in which clinical monitoring may not be possible depending on the clinical practices of surgeons. Such FTT procedures would likely benefit more from IDT monitoring than other FTT procedures where standard clinical monitoring is possible.

The final assumption was that a single Doppler probe was used in each FTT procedure. This assumption is in accordance with the recommendations of the manufacturer and the originator of the probes. The sensitivity analysis included a scenario in which two Doppler probes were used per patient.

#### *5.4.2 Sensitivity Analysis*

No formal sensitivity analysis was performed to address methodologic uncertainty. The typical sensitivity analysis performed to address methodological

uncertainty revolves around the recalculation of the model given different discount rates. As the time course for this model was only six weeks, time preference was very unlikely to be an important factor.

Structural uncertainty was addressed; albeit, in a perhaps a limited fashion. The model's assumption of a single Doppler probe was subjected to deterministic sensitivity analysis. This DSA found the use of two Doppler probes not to be cost-effective as the calculated ICER was over one million dollars. Interestingly, the largest series of IDT use in this thesis frequently used multiple probes.

Different centres, with their differing care pathways and subsequently differing models, could reach differing conclusions from this project. Decision makers must consider this possibility when interpreting the results of this project. Prior to making any decision regarding the funding of IDT on the basis of this project, decision makers should have a solid understanding of the clinical pathways at work in their institution.

Finally, as stated in the methods section, the decision to use 5000 replications was arbitrary. Guideline statements on the methodology of economic evaluation do not explicitly recommend a number of replications. However, it is likely that doing more than 5000 replications would yield little in terms of enhanced precision. To illustrate this point Appendix E contains two scatterplots showing the incremental difference in QALYs and the incremental difference in costs as a function of the number of replications in the model.

## 5.5 Conclusion

The results of the model suggest that use of the IDT is dominant over use of clinical monitoring alone. However, these results are relatively fragile. With the PSA

model, the difference between arms in cost was C\$160.55. The difference in QALYs was 0.0003. The PSA was designed to account for a great deal of uncertainty in the parameter inputs. Despite this degree of uncertainty, the PSA findings were the same as those using deterministic techniques, namely that IDT use dominant over the use of clinical monitoring alone.

When two Doppler probes were used per patient instead of one, IDT ceased to be dominant. While two Doppler probes did offer a benefit over clinical monitoring alone, this benefit came at a cost, an ICER of C\$1 249 000 per QALY, which would be prohibitive for most decision makers.

## **6. Discussion**

### **6.1 Strengths of Project**

This project represents the first attempt to apply the techniques of economic evaluation to free tissue transfer (FTT) procedures in patients with head and neck cancer (HNC). It also represents several other firsts. The project is the first systematic review of the use of implantable Doppler technology (IDT) in FTT for treatment of surgical defects in the head and neck. It is the first to quantify utility states in patients undergoing who have received FTT procedures for HNC.

Many of the methods used to create the parameter inputs in the Markov simulation model were robust. The costs were generated using microcosting data. Utility scores were determined using the theoretically sound time trade-off method.

Uncertainty was handled using probabilistic sensitivity analysis in accordance with recommendations from national agencies such the Canadian Agency for Drugs and Technology in Health and the National Institute for Health and Clinical Excellence.

Given these strengths, the findings of this project are an important contribution to the literature and will guide decision-makers considering adopting IDT.

### **6.2 Limitations**

As with all economic evaluations, limitations must be addressed.

#### *6.2.1 Time-Horizon*

Ideally in economic analysis, the time-horizon should be as long as possible. However, in this project the time horizon was limited to 6 weeks. The reason for this was discussed earlier. The complicating factor with carrying the time horizon past 6 weeks is that beyond this period, adjuvant therapy, usually radiation or chemotherapy combined

with radiation, begins. After six weeks, patients who have successful reconstructive surgery will undergo adjuvant therapy whereas those patients whose reconstructions fail will not undergo adjuvant therapy until they are healed. Consequently, the payer will incur the additional cost of adjuvant therapy and the patients will suffer the short term decrease in their health state secondary to side-effects from the adjuvant therapy.

However, multiple studies have shown a survival benefit of adjuvant therapy in patients with cancer of the upper aerodigestive tract (UADT) (Furness et al, 2011; Bernier and Cooper, 2005; Corv, 2007). It can therefore be expected that the increase in cost and decrease in utility associated with the period of adjuvant therapy is associated with a longer term increase in survival. Unfortunately, modeling how delays in adjuvant therapy secondary due to failures in reconstructive surgery effect long-term survival is very difficult. There is no published data on this subject and any attempt to design a model to represent this decrease in long-term survival would require a model with significantly more assumptions.

The result of the shorter time-horizon is that our analysis likely underestimates the benefit of a successful FTT procedure. Accordingly, the analysis likely underestimates the benefit of technology such as implantable Doppler technology (IDT) which is associated with a higher rate of FTT success.

Another consequence of this study is that the costs considered are primarily those associated with the hospitalization of patients. Other costs borne by the Ministry of Health (MOH), such as, radiation therapy, outpatient follow-up, and chemotherapy are not considered. As only in-hospital costs are considered, the analysis is perhaps best used

by decision makers not at the level of the MOH but at by decision makers working at specific hospitals. The stated perspective of this economic analysis was that of the MOH.

Hospital-based decision makers may be a more appropriate audience for this project. Decisions about surgical equipment are usually made at the institutional level rather than by the MOH. Furthermore, clinical practices will vary among institutions. Consequently, the model used in this study will be more applicable to some institutions than others. For illustrative purposes the standard clinical pathway for patients undergoing FTT at the Ottawa Hospital is displayed graphically in Appendix F.

#### *6.2.2 Use of Local Cost Data*

The cost data that was used in this study was drawn from the Ottawa Hospital. This is tertiary teaching hospital. The costs associated with FTT procedures at other institutions may be substantially different. Some centres routinely admit patients to intensive care units (ICU) following FTT surgery. At the Ottawa Hospital, patients are admitted to a step-down unit (SDU) after surgery. The SDU provides a closer level of care than a regular hospital ward bed but it does not provide the same level of care as an ICU setting. Stay in an SDU is also less costly than stay in an ICU. Centres that do not admit patients to an SDU post-operatively may have different costs than the Ottawa Hospital.

Free tissue transfer surgery is complicated and includes multiple steps. How surgeons bill the MOH varies not only between institutions and between provinces, but in some cases, between surgeons (at the Ottawa Hospital the billing practices of surgeons were the same). Surgical fees outside Canada may vary even more. Although the cost of the initial surgery was excluded in this analysis, subsequent surgeries were included.

Decision makers using this analysis should consider local surgical costs and how those local costs may affect the applicability of this project.

### *6.2.3 The Use of Evidence from Case-Series and Expert Opinion*

The data used to create the significant parts of this model were derived from case series and expert opinion. Traditional epidemiologic rankings rate such sources of evidence very poorly in terms of quality.

Coyle and Lee have created a hierarchy of evidence used in economic analyses (Coyle and Lee, 2002). This hierarchy ranks evidence depending on the type of parameter. For example, the ranking for data used for clinical effect sizes is different from the ranking used for baseline clinical data. These differences recognize that the literature available for certain parameters is more robust than that available for other parameters.

In our model, data from pooled case series was used to determine the clinical effectiveness of IDT. This type of data is ranked as the second lowest level of evidence under the Coyle – Lee scale. Baseline clinical data was derived from both primary data gathering from a case series of FTT procedures at the Ottawa Hospital and from expert opinion. The Coyle – Lee scale ranks primary data derived from the jurisdiction of interest as the highest possible level of evidence. However, expert opinion is ranked as the lowest level of evidence.

The use of this low quality evidence obviously limits the findings of this project. Specifically, the use of a systematic review of published data may subject this project to a significant publication bias. Unfortunately, it is unlikely that more robust evidence will be provided in the future to guide decision makers. The primary outcome of FTT failure

is rare and conducting a randomized control trial (RCT) would require multicentre participation and over multiple years. While feasible, it is unlikely any such study will be performed. The IDT has been available for nearly three decades, and no such study has been performed. Waiting for data from an RCT will further delay determining whether IDT should be adopted.

The conclusions of this project should not be ignored because expert opinion and case series data was used to create the decision analytic model. As stated earlier in this thesis, to account for the use of expert opinion we incorporated large confidence intervals into the sensitivity analysis. Furthermore, other parameters of the decision analytic model were created using the highest levels of evidence on the Coyle – Lee scale. The utility parameters were determined primarily by a direct utility assessment. Resource use and costs were determined by prospective cost calculation and data collection from reliable and detailed administrative sources. Given the available evidence, this project has attempted to draw the most robust conclusions possible.

### 6.3 Implications for Decision Makers

This study suggests that IDT use in monitoring FTT procedures performed to reconstruct defects due to the treatment of head and neck cancers is a cost-effective intervention. Decision makers should discuss the findings with local clinicians and healthcare managers to determine how applicable this study's model was to the local conditions. Decision makers should also remember that the model chosen was as broad as possible and in some cases, especially the use of a shorten time-horizon, probably underestimated the benefit of the IDT.

### 6.4 Future Research

#### *6.4.1 Design of a Randomized Control Trial to Assess Implantable Doppler Technology*

Ultimately the conclusions of this thesis are limited by the effectiveness data currently available. The majority of the uncertainty identified by the value of information analysis was found in the take back rates and the salvage rates associated with the different types of monitoring.

An RCT designed to determine whether the use of IDT in patients undergoing FTT for reconstruction of surgical defects from treatment of HNC improves FTT salvage rate and overall FTT survival rate would significantly strengthen the ability of health economic analysis to determine the benefit of IDT.

Were an RCT to be designed, overall FTT survival should be the primary outcome. Improved FTT salvage rate and take back rates should be secondary outcomes.

An RCT of IDT use would involve randomization to one of two arms. In one arm participants would be monitored by traditional clinical methods alone. In the other arm, participants would be monitored by both IDT and traditional clinical methods. Obviously, it would not be possible to blind either observers or participants to the intervention being employed.

While the clinical methods used to monitor FTT procedures vary between institutions, these methods will require standardization for the purposes of the trial. This will mean standardizing the method used, the frequency of its use, and the personnel performing these checks. The standardized method should be determined by a group of health professionals caring for the participants in the study.

The use of IDT should also be standardized. Based on the recommendation of Swartz and of the manufacturer, a single IDT probe should be used and placed on the

anastomotic vein. Health care personnel should be trained to listen to the audible signal for signs of FTT failure by professionals experienced with IDT technology.

The response to signs of FTT failure should also be standardized. All flaps showing either clinical signs of failure or signs of failure as per IDT should be taken to the operating room immediately for exploration and possible repair of the microvascular anastomosis.

Data should be collected on take back rates, salvage rates, and overall survival rates. Data on risk factors for FTT failure such as tobacco use and use of vein grafts should also be collected. Demographic data such as participant age, type of FTT procedures, participant gender should also be collected. In parallel to the clinical data being gathered, cost data and resource data should be accrued, ideally using a microcosting technique. Furthermore, overall patient survival should be measured as another secondary outcome. If delays in receiving adjuvant therapy are associated with FTT failure, it is reasonable to expect that improving FTT success rates may improve overall patient survival.

The number of participants required for this trial can be calculated as follows:

$P_1$ : the overall survival of FTT procedures monitored by traditional methods alone.

Based on the systematic review in this project this is 0.957.

$P_2$ : the overall survival of FTT procedures monitored by IDT as well as traditional methods. Based on the systematic review in this project this is 0.984.

ES: effect size of the intervention =  $P_2 - P_1$ . In this case, ES = 0.027.

$\alpha$ : accepting an  $\alpha$  error of 0.05 in a two tailed test.

$\beta$ : accepting a power of 0.80,  $\beta = 0.20$ .

Using these inputs, the trial would require 615 participants per arm or 1230 participants in total. Given that the average tertiary care centre in Ontario performs 60 FTT procedures a year, a single centre would take 20 years to recruit enough patients to perform this study (assuming a 100% recruitment rate). If seven centres were involved in the study (there are currently five, high-volume centres in Ontario performing FTT procedures for Head and Neck cancer), the necessary recruitment of participants could be achieved within three years, again assuming a 100% recruitment rate.

#### *6.4.2 Utility Values of Head and Neck Health States*

This study highlighted the dearth of research on utility states in head and neck cancer (HNC). While disease-specific quality of life scales such as the University of Washington Head and Neck scale and the EORTC head and neck scale have been developed and used extensively, no author has determined utility states in patients undergoing surgical treatment of HNC (Hassan and Weymuller, 1993; Bjordal et al, 1999). Studies examining cost-effectiveness of interventions in head and neck cancer have tended to extract utility data from disease-specific scales instead of true utility scales (Greenhalgh et al, 2009; Parthan et al, 2009).

The utility survey which formed part of this project is the first known attempt to quantify utility in this group of patients. This project's utility survey was limited to only four relevant health states; however, the utility of multiple other HNC health states remains undetermined. Future research should be directed towards better determination of HNC health utility states.

Utility data could be collected either by primary methods such as standard gamble and time trade-off as described earlier or by multi-attribute questionnaires.

Multi-attribute questionnaires are likely the best way to capture large datasets on various health states as they could easily be integrated into the standard patient encounter at any head and neck unit. This would effectively capture a large amount of utility data on patients in nearly all important HNC health states.

As mentioned before the EQ-5D, the SF-6D, the HUI2, and the HUI3 are commonly used multi-attribute scales (Essink-bot, Stouthard, and Bonsel, 1993; Brazier, Roberts, and Deverill, 2002; Furlong et al., 2001). Drummond points out that these scales are not interchangeable (Drummond 2005). Yet, each scale has a substantial body of evidence validating their use and each scale has been used widely with success. Consequently, it is quite likely that any of the EQ-5D, the SF-6D, or the HUI2/3 would be an adequate choice of scale.

The benefits of utility research in HNC are obvious. If utility data of various HNC states are defined, economic evaluation will be easier and more accurate. Improved economic evaluations in HNC will complement the existing efficacy and effectiveness data.

## **7. Conclusion**

This study performed an economic evaluation of the use of Cook-Swartz implantable Doppler monitoring of free tissue transfer procedures performed to reconstruct defects arising from the treatment of head and neck cancer. A Markov model based on the clinical pathways used by surgeons at the University of Ottawa was created. The model used primarily gathered cost and utility data as inputs. It relied on a systematic review of case series to determine the effectiveness of implantable Doppler technology. Implantable Doppler technology was found to be a cost effective intervention.

## **References.**

1. Adronis L., Barton P., Bryan S. (2009) Sensitivity analysis in economic evaluation: an audit of NICE current practice and a review of its use and value in decision-making. *Health Technology Assessments*. 13(29), 1-61.
2. Barnett PG. (2009). An improved set of standards for finding the cost for cost-effectiveness analysis. *Medical Care*, 47(7 suppl 1), S82-88.
3. Bernier J., Cooper JS. (2005) Chemoradiation after surgery for high-risk head and neck cancer patients: how strong is the evidence? *Oncologist* 10(3), 215-224.
4. Bjordal, K., Hammerlid, E., Ahlner Elmquist, M., de Graeff, A., Boysen, M., Evensen, J. F., . . . Kaasa, S. (1999). Quality of life in head and neck cancer patients: Validation of the european organization for research and treatment of cancer quality of life questionnaire-H N35. *Journal of Clinical Oncology*, 17(3), 1008-1019.
5. Black, M. J., Chait, L., O'Brien, B. M., Sykes, P. J., & Sharzer, L. A. (1978). How soon may the axial vessels of a surviving free flap be safely ligated: A study in pigs. *British Journal of Plastic Surgery*, 31(4), 295-299.
6. Blackwell, K. E. (1999). Unsurpassed reliability of free flaps for head and neck reconstruction. *Archives of Otolaryngology--Head Neck Surgery*, 125(3), 295-299.

7. Blackwell, K. E., Brown, M. T., & Gonzalez, D. (1997). Overcoming the learning curve in microvascular head and neck reconstruction. *Archives of Otolaryngology--Head Neck Surgery*, 123(12), 1332-1335.
8. Brazier, J., Roberts, J., & Deverill, M. (2002). The estimation of a preference-based measure of health from the SF-36. *Journal of Health Economics*, 21(2), 271-292.
9. Briggs, A. H. (2000). Handling uncertainty in cost-effectiveness models. *Pharmacoeconomics*, 17(5), 479-500.
10. Briggs, A., Sculpher, M., & Buxton, M. (1994). Uncertainty in the economic evaluation of health care technologies: The role of sensitivity analysis. *Health Economics*, 3(2), 95-104.
11. Briggs, A. (2005). Probabilistic analysis of cost-effectiveness models: Statistical representation of parameter uncertainty. *Value in Health*, 8(1), 1-2.
12. Briggs A., Claxton K, Sculpher M. (2006) *Decision Modelling for Health Economic Evaluation*. Oxford: Oxford University Press.
13. Buxton, M. J., Drummond, M. F., Van Hout, B. A., Prince, R. L., Sheldon, T. A., Szucs, T., & Vray, M. (1997). Modelling in economic evaluation: An unavoidable fact of life. *Health Economics*, 6(3), 217-227.
14. The Canadian Agency for Drugs and Technology in Health. Guidelines for Economic Evaluation of Health Technologies: Canada. 3<sup>rd</sup> Edition. Ottawa: Canadian Agency

for Drugs and Technology in Health (CADTH) 2006. Available:

[http://www.cadth.ca/media/pdf/186\\_EconomicGuidelines\\_e.pdf](http://www.cadth.ca/media/pdf/186_EconomicGuidelines_e.pdf).

15. Canadian Medical Association. Health Care Transformation in Canada: Change that works. Care that lasts. Ottawa, 2010.  
[http://www.cma.ca/multimedia/cma/content\\_images/inside\\_cma/advocacy/HCT/HC-T-2010report\\_en.pdf](http://www.cma.ca/multimedia/cma/content_images/inside_cma/advocacy/HCT/HC-T-2010report_en.pdf).
16. Chang, WT, Collins ED, Kerrigan CL (2001). An internet-based utility assessment of breast hypertrophy. *Plastic and Reconstructive Surgery*, 108(2), 370-377.
17. Chepeha, D., Nussenbaum, B., Bradford, C., & Teknos, T. (2002). Leech therapy for patients with surgically unsalvageable venous obstruction after revascularized free tissue transfer. *Archives of Otolaryngology--Head Neck Surgery*, 128(8), 960-965.
18. Chilcott, J., Brennan, A., Booth, A., Karnon, J., & Tappenden, P. (2003). The role of modelling in prioritising and planning clinical trials. *Health Technology Assessment*, 7(23), iii, 1-iii125.
19. Claxton, K., Sculpher, M., McCabe, C., Briggs, A., Akehurst, R., Buxton, M., . . . O'Hagan, T. (2005). Probabilistic sensitivity analysis for NICE technology assessment: Not an optional extra. *Health Economics*, 14(4), 339-347.
20. Cook, D. J., Guyatt, G. H., Laupacis, A., Sackett, D. L., & Goldberg, R. J. (1995). Clinical recommendations using levels of evidence for antithrombotic agents. *Chest*, 108(4 Suppl), 227S-230S.

21. Cooper, N., Coyle, D., Abrams, K., Mugford, M., & Sutton, A. (2005). Use of evidence in decision models: An appraisal of health technology assessments in the UK since 1997. *Journal of Health Services Research Policy*, 10(4), 245-250.
22. Corv, R. (2007). Evidence-based radiation oncology in head and neck squamous cell carcinoma. *Radiotherapy and Oncology*, 85(1), 156-170.
23. Coyle D, Lee KM. Evidence-based economic evaluation: how the use of different data sources can impact results In: Donaldson C, Mugford M, Vale L, eds. (2002) *Evidence-Based Health Economics: From Effectiveness to Efficiency in Systematic Review*. London: BMJ Publishing Group, 55–66.
24. Coyle, D., & Oakley, J. (2008). Estimating the expected value of partial perfect information: A review of methods. *The European Journal of Health Economics*, 9(3), 251-259.
25. Coyle, D., Tolley, K. (1992). Discounting of health benefits in pharmacoeconomic analysis of drug therapies: an issue for debate? *Pharmacoeconomics*, 2(2), 153-162.
26. Daniels, M., & Zhao, Y. (2003). Modelling the random effects covariance matrix in longitudinal data. *Statistics in Medicine*, 22(10), 1631-1647.
27. Dalziel, K., Round, A., Stein, K., Garside, R., Castelnuovo, E., & Payne, L. (2005). Do the findings of case series studies vary significantly according to methodological characteristics? *HealthTechnology Assessment*, 9, iii-iv, 1.

28. Dassonville, O., Poissonnet, G., Chamorey, E., Vallicioni, J., Demard, F., Santini, J., . . . Bozec, A. (2008). Head and neck reconstruction with free flaps: A report on 213 cases. *European Archives of Oto-Rhino-Laryngology*, 265(1), 85-95.
29. de Bree, R., Reith, R., Quak, JJ., . . . Leemans, CR. (2007) Free radial forearm flap versus pectoralis muscle myocutaneous flap reconstruction of oral and oropharyngeal defects: a cost analysis. *Clinical Otolaryngology*, 32(4), 275-82.
30. Deeks JJ, Higgins JPT, Altman DG (editors). Chapter 9: Analysing data and undertaking meta-analyses. In: Higgins JPT, Green S (editors). *Cochrane Handbook for Systematic Reviews of Interventions*. Version 5.0.1 [updated September 2008]. The Cochrane Collaboration, 2008. Available from [www.cochrane-handbook.org](http://www.cochrane-handbook.org).
31. DerSimonian, R., & Laird, N. (1986). Meta-analysis in clinical trials. *Controlled Clinical Trials*, 7(3), 177-88.
32. Dickersin, K. (1990). The existence of publication bias and risk factors for its occurrence. *JAMA (Chicago, Ill.)*, 263(10), 1385-1389.
33. Drummond M.F., Sculpher M.J., Torrance G.W., O'Brien B.J., Stoddart G.L. (2005) *Methods for the Economic Evaluation of Health Care Programmes*. Toronto: Oxford University Press.
34. Eckardt, A., Meyer, A., Laas, U., & Hausamen, J. (2007). Reconstruction of defects in the head and neck with free flaps: 20 years experience. *The British Journal of Oral Maxillofacial Surgery*, 45(1), 11-15.

35. Essink Bot, M. L., Stouthard, M. E., & Bonsel, G. J. (1993). Generalizability of valuations on health states collected with the EuroQolc-questionnaire. *Health Economics*, 2(3), 237-246.
36. Felli, J. C., & Hazen, G. B. (1998). Sensitivity analysis and the expected value of perfect information. *Medical Decision Making*, 18(1), 95-109.
37. Ferguson, RH., Yu, P. (2009) Techniques in monitoring buried free flaps. *Plastic and Reconstructive Surgery*, 123(2), 526-532
38. Fitzpatrick-Lewis, D., Ciliska D., & Thomas, H. (2009). The Methods for the Synthesis of Studies without Control Groups. Hamilton, ON: National Collaborating Centre for Methods and Tools. [[http://www.nccmt.ca/pubs/non-RCT2\\_FN.pdf](http://www.nccmt.ca/pubs/non-RCT2_FN.pdf)].
39. Frick KD. (2009) Microcosting quantity data collection methods. *Medical Care*, 47(7 suppl 1), S76-81.
40. Furlong, W. J., Feeny, D. H., Torrance, G. W., & Barr, R. D. (2001). The health utilities index (HUI) system for assessing health-related quality of life in clinical studies. *Annals of Medicine*, 33(5), 375.
41. Furnas, H., & Rosen, J. M. (1991). Monitoring in microvascular surgery. *Annals of Plastic Surgery*, 26(3), 265-272.
42. Furness, S., Glenly, A., Worthington, H., Pavitt, S., Oliver, R., Clarkson, J., . . . Conway, D. (2011). Interventions for the treatment of oral cavity and oropharyngeal

cancer: Chemotherapy. *Cochrane Database of Systematic Reviews*, 4, CD006386-CD006386.

43. Furr, M., Cannady, S., & Wax, M. (2011). Interposition vein grafts in microvascular head and neck reconstruction. *The Laryngoscope*, 121(4), 707-711.
44. Genden, EM., Rinaldo, A., Suarez, C., Wei, WI., Bradley, PJ., Ferlito, A. Complications of free flap transfers for head and neck reconstruction following cancer resection. (2004) *Oral Oncology*, 40(10), 979-84.
45. Gold, M. R., Franks, P., McCoy, K. I., & Fryback, D. G. (1998). Toward consistency in cost-utility analyses: Using national measures to create condition-specific values. *Medical Care*, 36(6), 778.
46. Green, C., Brazier, J., & Deverill, M. (2000). Valuing health-related quality of life. A review of health state valuation techniques. *Pharmacoeconomics*, 17(2), 151-165.
47. Greenhalgh, J., Bagust, A., Boland, A., Fleeman, N., McLeod, C., Dundar, Y., . . . Shaw, R. (2009). Cetuximab for the treatment of recurrent and/or metastatic squamous cell carcinoma of the head and neck. *Health Technology Assessment*, 13 Suppl 3, 49-54.
48. Guillemaud, J., Seikaly, H., Cote, D., Allen, H., & Harris, J. (2008). The implantable cook-swartz doppler probe for postoperative monitoring in head and neck free flap reconstruction. *Archives of Otolaryngology--Head Neck Surgery*, 134(7), 729-734.

49. Hassan, S. J., & Weymuller, E. A. (1993). Assessment of quality of life in head and neck cancer patients. *Head Neck, 15*(6), 485-496.
50. Haughey, B. H., Wilson, E., Kluwe, L., Piccirillo, J., Fredrickson, J., Sessions, D., & Spector, G. (2001). Free flap reconstruction of the head and neck: Analysis of 241 cases. *Otolaryngology and Head and Neck Surgery, 125*(1), 10-17.
51. Hedges, L.V., Pigott, TD. (2001) The power of statistical tests in meta-analysis. *Psychological Methods, 6*, 203-217.
52. Hoffmann, J., Ehrenfeld, M., Hwang, S., & Schwenzer, N. (1998). Complications after microsurgical tissue transfer in the head and neck region. *Journal of Cranio-Maxillo-Facial Surgery, 26*(4), 255-259.
53. Hyodo, I., Nakayama, B., Kato, H., Hasegawa, Y., Ogawa, T., Terada, A., & Torii, S. (2007). Analysis of salvage operation in head and neck microsurgical reconstruction. *The Laryngoscope, 117*(2), 357-360.
54. Jones, NF., Jarrahy, R., Song, JI., Kaufman, MR., Markowitz, B. (2007) Postoperative medical complications – not microsurgical complications – negatively influence the morbidity, mortality, and true costs after microsurgical reconstruction for head and neck cancer. *Plastic and Reconstructive Surgery, 119*(7), 2053-2060.
55. Kassirer, J. P., & Angell, M. (1994). The journal's policy on cost-effectiveness analyses. *The New England Journal of Medicine, 331*(10), 669-670.

56. Kerrigan, C. L., Zelt, R. G., & Daniel, R. K. (1984). Secondary critical ischemia time of experimental skin flaps. *Plastic and Reconstructive Surgery*, 74(4), 522-526.
57. Krabbe, P. F., Essink-Bot, M. L., & Bonsel, G. J. (1997). The comparability and reliability of five health-state valuation methods. *Social Science Medicine*, 45(11), 1641-1652.
58. Kroll, SS., Evans, GR., Goldberg, D., . . . Schusterman, MA. (1997) A comparison of resource costs for head and neck reconstruction with free and pectoralis major flaps. *Plastic and Reconstructive Surgery*, 99(5), 1282-1286.
59. Kubo, T., Yano, K., & Hosokawa, K. (2002). Management of flaps with compromised venous outflow in head and neck microsurgical reconstruction. *Microsurgery*, 22(8), 391-395.
60. Luu, Q., & Farwell, D. G. (2009). Advances in free flap monitoring: Have we gone too far? *Current Opinion in Otolaryngology Head and Neck Surgery*, 17(4), 267-269.
61. Manning WG, Fryback DG, Weinstein MC. Reflecting uncertainty in cost-effectiveness analysis. In: Gold MR, Siegel JE, Russel LB, Weinstein MC, eds (2002). *Cost-effectiveness in Health and Medicine*. New York: Oxford University Press.

62. May, J. W., Chait, L. A., O'Brien, B. M., & Hurley, J. V. (1978). The no-reflow phenomenon in experimental free flaps. *Plastic and Reconstructive Surgery*, 61(2), 256-267.
63. Miller, MJ., Swartz, WM., Miller, RH., Harvey, JM. (1991). Cost analysis of microsurgical reconstruction in the head and neck. *Journal of Surgical Oncology*, 46(4), 230-4.
64. Ministry of Health and Long-term Care. Ontario Health Insurance Plan (OHIP). Schedule of Fees and Services. Physician Benefits.  
[http://www.health.gov.on.ca/english/providers/program/ohip/sob/phsserv/physserv\\_mn.html](http://www.health.gov.on.ca/english/providers/program/ohip/sob/phsserv/physserv_mn.html). (accessed June 2010)
65. Nakamizo, M., Yokoshima, K., & Yagi, T. (2004). Use of free flaps for reconstruction in head and neck surgery: A retrospective study of 182 cases. *Auris Nasus Larynx*, 31(3), 269-273.
66. Nakatsuka, T., Harii, K., Asato, H., Takushima, A., Ebihara, S., Kimata, Y., . . . Ichioka, S. (2003). Analytic review of 2372 free flap transfers for head and neck reconstruction following cancer resection. *Journal of Reconstructive Microsurgery*, 19(6), 363-8.
67. Novakovic, D., Patel, R., Goldstein, D., & Gullane, P. (2009). Salvage of failed free flaps used in head and neck reconstruction. *Head Neck Oncology*, 1, 33-33.

68. O'Brien, C. J., Lee, K. K., Stern, H. S., Traynor, S. J., Bron, L., Tew, P. J., & Haghghi, K. S. (1998). Evaluation of 250 free-flap reconstructions after resection of tumours of the head and neck. *Australian and New Zealand Journal of Surgery*, 68(10), 698-701.
69. Olivas, T. P., Saylor, T. F., Wong, H. P., Stephenson, L. L., & Zamboni, W. A. (2001). Timing of microcirculatory injury from ischemia reperfusion. *Plastic and Reconstructive Surgery*, 107(3), 785-788.
70. Parthan, A., Posner, M., Brammer, C., Beltran, P., & Jansen, J. (2009). Cost utility of docetaxel as induction chemotherapy followed by chemoradiation in locally advanced squamous cell carcinoma of the head and neck. *Head Neck*, 31(10), 1255.
71. Parthan, A., Posner, M., Brammer, C., Beltran, P., & Jansen, J. (2009). Cost utility of docetaxel as induction chemotherapy followed by chemoradiation in locally advanced squamous cell carcinoma of the head and neck. *Head Neck*, 31(10), 1255-1262.
72. Paydar, K., Hansen, S., Chang, D., Hoffman, W., & Leon, P. (2010). Implantable venous doppler monitoring in head and neck free flap reconstruction increases the salvage rate. *Plastic and Reconstructive Surgery*, 125(4), 1129-1134.
73. Philips, Z., Bojke, L., Sculpher, M., Claxton, K., & Golder, S. (2006). Good practice guidelines for decision-analytic modelling in health technology assessment: A review and consolidation of quality assessment. *Pharmacoeconomics*, 24(4), 355-371.

74. Podrecca, S., Salvatori, P., Squadrelli Saraceno, M., Fallahdar, D., Calabrese, L., Cant, G., & Molinari, R. (2006). Review of 346 patients with free-flap reconstruction following head and neck surgery for neoplasm. *Journal of Plastic, Reconstructive Aesthetic Surgery*, 59(2), 122-129.
75. Pohlenz, P., Blessmann, M., Blake, F., Li, L., Schmelzle, R., & Heiland, M. (2007). Outcome and complications of 540 microvascular free flaps: The hamburg experience. *Clinical Oral Investigations*, 11(1), 89-92.
76. Preston, C., Ashby, D., & Smyth, R. (2004). Adjusting for publication bias: Modelling the selection process. *Journal of Evaluation in Clinical Practice*, 10(2), 313-322.
77. Pratt JW, Raiffa H, Schlaifer R. (1995) *Introduction to Statistical Decision Theory*. Cambridge: MIT Press.
78. Pryor, S., Moore, E., & Kasperbauer, J. (2006). Implantable doppler flow system: Experience with 24 microvascular free-flap operations. *Otolaryngology and Head and Neck Surgery*, 135(5), 714-718.
79. Reeves BC, Deeks JJ, Higgins JPT, Wells GA. Chapter 13: Including non-randomized studies. In: Higgins JPT, Green S (editors), *Cochrane Handbook for Systematic Reviews of Interventions*. Version 5.0.1 [updated September 2008]. The Cochrane Collaboration, 2008. Available from [www.cochrane-handbook.org](http://www.cochrane-handbook.org).

80. Rocchi, A., Menon, D., Verma, S., Miller, E. (2008). The role of economic evidence in Canadian reimbursement decision-making: to lambda and beyond. *Value Health, 11(4), 771-783.*
81. Ross, D., Ariyan, S., Restifo, R., & Sasaki, C. (2003). Use of the operating microscope and loupes for head and neck free microvascular tissue transfer: A retrospective comparison. *Archives of Otolaryngology--Head Neck Surgery, 129(2), 189-193.*
82. Ross, G., Ang, E. S. W., Lannon, D., Addison, P., Golger, A., Novak, C., . . . Neligan, P. (2008). Ten-year experience of free flaps in head and neck surgery. how necessary is a second venous anastomosis? *Head Neck, 30(8), 1086-1089.*
83. Rozen, W., Chubb, D., Whitaker, I., & Acosta, R. (2010). The efficacy of postoperative monitoring: A single surgeon comparison of clinical monitoring and the implantable doppler probe in 547 consecutive free flaps. *Microsurgery, 30(2), 105-110.*
84. Rozen, W., Enajat, M., Whitaker, I., Lindkvist, U., Audolfsson, T., & Acosta, R. (2010). Postoperative monitoring of lower limb free flaps with the cook-swartz implantable doppler probe: A clinical trial. *Microsurgery, 30(5), 354-360.*
85. Schmulder, A., Gur, E., & Zaretski, A. (2010). Eight-year experience of the cook-swartz doppler in free-flap operations: Microsurgical and reexploration results with regard to a wide spectrum of surgeries. *Microsurgery,*

86. Simpson, K. H., Murphy, P. G., Hopkins, P. M., & Batchelor, A. G. (1996). Prediction of outcomes in 150 patients having microvascular free tissue transfers to the head and neck. *British Journal of Plastic Surgery*, *49*(5), 267-273.
87. Singh, B., Cordeiro, P. G., Santamaria, E., Shaha, A. R., Pfister, D. G., & Shah, J. P. (1999). Factors associated with complications in microvascular reconstruction of head and neck defects. *Plastic and Reconstructive Surgery*, *103*(2), 403-411.
88. Sinno HH, Thibaudeau S, Duggal A, Lessard L (2010). Utility scores for facial disfigurement requiring facial transplantation [outcomes article]. *Plastic and Reconstructive Surgery*, *126*(2), 443-449.
89. Smit, J., Zeebregts, C., Acosta, R., & Werker, P. M. N. (2010). Advancements in free flap monitoring in the last decade: A critical review. *Plastic and Reconstructive Surgery*, *125*(1), 177-185.
90. Smith, RB., Sniezek, JC., Weed, DT., Wax, MK. (2007) Utilitization of free tissue transfer in head and neck surgery. *Otolaryngology-Head and Neck Surgery*, *137*(2) 182-91.
91. Spiegel, J., & Polat, J. (2007). Microvascular flap reconstruction by otolaryngologists: Prevalence, postoperative care, and monitoring techniques. *The Laryngoscope*, *117*(3), 485-490.

92. Stang, A. (2010). Critical evaluation of the newcastle-ottawa scale for the assessment of the quality of nonrandomized studies in meta-analyses. *European Journal of Epidemiology*, 25(9), 603-605.
93. Stollenwerk, B., Stock, S., Siebert, U., Lauterbach, K., & Holle, R. (2010). Uncertainty assessment of input parameters for economic evaluation: Gauss's error propagation, an alternative to established methods. *Medical Decision Making*, 30(3), 304-313.
94. Steuten, L. M., Vrijhoef, H. J., van-Merode, G. G., Severens, J. L., & Spreeuwenberg, C. (2004). The health technology assessment – Disease management instrument reliably measured methodologic quality of health technology assessments of disease management. *Journal of Clinical Epidemiology*, 57, 881-888.
95. Stuart A, Ord JK. (1994) *Kendall's Advanced Theory of Statistics* (6th edition). London: Edward Arnold.
96. Suh, J., Sercarz, J., Abemayor, E., Calcaterra, T., Rawnsley, J., Alam, D., & Blackwell, K. (2004). Analysis of outcome and complications in 400 cases of microvascular head and neck reconstruction. *Archives of Otolaryngology--Head Neck Surgery*, 130(8), 962-966.
97. Swartz, W. M., Izquierdo, R., & Miller, M. J. (1994). Implantable venous doppler microvascular monitoring: Laboratory investigation and clinical results. *Plastic and Reconstructive Surgery*, 93(1), 152-163.

98. Swartz, W. M., Jones, N. F., Cherup, L., & Klein, A. (1988). Direct monitoring of microvascular anastomoses with the 20-MHz ultrasonic doppler probe: An experimental and clinical study. *Plastic and Reconstructive Surgery*, *81*(2), 149.
99. Syin, D., Woreta, T., Chang, D., Cameron, J., Pronovost, P., & Makary, M. (2007). Publication bias in surgery: Implications for informed consent. *The Journal of Surgical Research*, *143*(1), 88-93.
100. Tabah, R. J., Flynn, M. B., Acland, R. D., & Banis, J. C. (1984). Microvascular free tissue transfer in head and neck and esophageal surgery. *The American Journal of Surgery*, *148*(4), 498-504.
101. Tamai, S. (2009). History of microsurgery. *Plastic and Reconstructive Surgery*, *124*(6 Suppl), e282.
102. Teerawattananon, Y., Mugford, M., & Tangcharoensathien, V. (2007). Economic evaluation of palliative management versus peritoneal dialysis and hemodialysis for end-stage renal disease: Evidence for coverage decisions in thailand. *Value in Health*, *10*(1), 61-72.
103. Thoma, A., Jansen, L., & Sprague, S. (2009). Outcomes in microsurgery. *Plastic and Reconstructive Surgery*, *124*(6 Suppl), e303-e312.
104. Torrance, G. W., Feeny, D., & Furlong, W. (2001). Visual analog scales: Do they have a role in the measurement of preferences for health states? *Medical Decision Making*, *21*(4), 329-334.

105. Torrance, G., Furlong, W., & Feeny, D. (2002). Health utility estimation. *Expert Review of Pharmacoeconomics Outcomes Research*, 2(2), 99-108.
106. Tsue, T., Desyatnikova, SS., Deleyiannis, FWB., ... Glenn, MG. (1997). Comparison of cost and function in reconstruction of the posterior oral cavity and oropharynx: free vs. pedicled soft tissue transfer. *Archives of Otolaryngology – Head and Neck Surgery*, 123(7), 731-7.
107. Urken, M. L., Weinberg, H., Buchbinder, D., Moscoso, J. F., Lawson, W., Catalano, P. J., & Biller, H. F. (1994). Microvascular free flaps in head and neck reconstruction. report of 200 cases and review of complications. *Archives of Otolaryngology--Head Neck Surgery*, 120(6), 633-640.
108. Van Hout, B. A., Al, M. J., Gordon, G. S., & Rutten, F. F. (1994). Costs, effects and C/E-ratios alongside a clinical trial. *Health Economics*, 3(5), 309-319.
109. Von Neumann J. and Morganstern O. (1944) *Theory of Games and Economic Behavior*. Princeton: Princeton University Press.
110. Weinstein, M. C., Siegel, J. E., Gold, M. R., Kamlet, M. S., & Russell, L. B. (1996). Recommendations of the panel on cost-effectiveness in health and medicine. *JAMA (Chicago, Ill.)*, 276(15), 1253-1258.
111. Yoshimoto, S., Kawabata, K., & Mitani, H. (2010). Analysis of 59 cases with free flap thrombosis after reconstructive surgery for head and neck cancer. *Auris Nasus Larynx*, 37(2), 205-211.

112. Yoshimoto, Y. (2003). Publication bias in neurosurgery: Lessons from series of unruptured aneurysms. *Acta Neurochirurgica*, *145*(1), 45-48.
113. Yu, P., Chang, D., Miller, M., Reece, G., & Robb, G. (2009). Analysis of 49 cases of flap compromise in 1310 free flaps for head and neck reconstruction. *Head Neck*, *31*(1), 45-51.

## **Appendix A: Excluded Full Papers**

<b>Article</b>	<b>Reason for Exclusion</b>
Ferguson, RH., Yu, P. (2009) Techniques in monitoring buried free flaps. <i>Plastic and Reconstructive Surgery</i> , 123(2), 526-532	Specialized free flap population
Schlenker, JD., Robson, MC., Parsons, RW. (1981) Methods and results of reconstruction with free flaps following resection of squamous cell carcinoma of the head and neck. <i>Annals of Plastic Surgery</i> , 6(5), 362-373.	Take back rate, salvage rate not described
MacLeod, AM., O'Brien, BMC., Morrison, WA. (1979) Microvascular techniques in reconstruction following major resections for cancer of the head and neck. <i>Australian and New Zealand Journal of Surgery</i> , 49(6), 648-653.	Take back rate, salvage rate not described
Morais-Besteiro, J., Cernea, CR., Medina dos Santos, LR., . . . Ferraz, AR. (1990) Microvascular flaps in head and neck reconstruction. <i>Head and Neck</i> , 12(1), 21-30.	Take back rate, salvage rate not described
Waris, TH., Kaarela, OI., Teerikangas, HE., Raatikainen, TK. (1994) Fifty consecutive free flaps in a new microsurgical unit. <i>Microsurgery</i> , 15(2), 150-52.	Not all head and neck cancer patients, unable to discern rates for head and neck cancer patients alone.
Myers, LL., Sumer, BD., Defatta, RJ., Minhajuddin, A. (2008) Free tissue transfer reconstruction of the head and neck at a veterans affairs hospital. <i>Head and Neck</i> , 30(8), 1007-11.	Take back rates and salvage rates not available
Ryan, MW., Hochman, M., (2000) Length of stay after free flap reconstruction of the head and neck. <i>The Laryngoscope</i> , 110(2), 210-16.	Take back rates and salvage rates not available.
Jones, NF., Jarrahy, R., Song, JI., Kaufman, MR., Markowitz, B. Postoperative medical complications-not microsurgical complications-megatively influence the morbidity. <i>Plastic and Reconstructive Surgery</i> , 119(6), 2053-2060.	Monitoring method not described.
Blackwell, KE. (1999) Unsurpassed	Included in larger series.

reliability of free flaps for head and neck reconstruction. <i>Archives of Otolaryngology and Head and Neck Surgery</i> , 125(3), 295-299.	
Khariwala, SS., Vivek, PP., Lorenz, RR., Escalamado, RM., Wood, B., Strome, M., Alam, DS. (2000) Swallowing outcomes after microvascular head and neck reconstruction: A prospective review of 191 cases. <i>The Laryngoscope</i> , 117(8), 1359-63.	Take back rates and salvage rates not available.
Pohlenz, P., Blessmann, M., Heiland, M., Blake, F., Schmelzle, R., Lei, L. (2007) Postoperative complications in 202 cases of microvascular head and neck reconstructions. <i>Journal of Cranio-Maxillofacial Surgery</i> , 35(6-7) 311-15.	Included in larger series
Haughey, BH., Wilson, E., Kluwe, L, Piccirillo, J., Fredrickson, J., Sessions, D., Spector, G. (2001) Free flap reconstruction of the head and neck: analysis of 241 cases. <i>Otolaryngology-Head and Neck Surgery</i> , 125(1), 10-17.	Take back rates and salvage rates not available.
Jones, NF, Johnson JT., Shestak, KC., Myers, EN., Swartz, WM. (1996) Microsurgical reconstruction of the head and neck: Interdisciplinary collaboration between head and neck surgeons and plastic surgeons in 305 cases. <i>Annals of Plastic Surgery</i> , 36(1), 37-43.	Implantable Doppler population not isolated.
Schusterman, MA., Miller, MJ., Reece, GP., Kroll, SS., Marchi, M., Goepfort, H. (1994) A single centres experience with 308 free flaps for repair of head and neck cancer defects. <i>Plastic and Reconstructive Surgery</i> , 93(3), 472-78.	Monitoring technique not described, take back rates not described.
Eckardt, A., Fokas, K., (2003) Microsurgical reconstruction in the head and neck region: an 18-year experience with 500 consecutive cases. <i>Journal of Cranio-Maxillofacial Surgery</i> , 31(4), 197-201.	Included in larger series
Okazaki, M., Asato, H., Takushima, A., Sarukawa, S., Nakatsuka, T., Yamada, A., Harii, K. (2007) Analysis of salvage treatments following the failure of free flap	Take back rates and monitoring technique not included.

transfer caused by vascular thrombosis in reconstruction for head and neck cancer. <i>Plastic and Reconstructive Surgery</i> , 119(4), 1223-32.	
Yu, P., Chang, D., Miller, M., Reece, G., Robb, G. (2009) Analysis of 49 cases of flap compromise in 1310 free flaps for head and neck reconstruction. <i>Head and Neck</i> , 31(1) 45-51.	Implantable Doppler population not isolated.
Rosenberg, JJ., Fornage, BD., Chevray, PM. (2006) Monitoring buried free flaps: limitations of implantable Doppler and use of color duplex sonography as a confirmatory test. <i>Plastic and Reconstructive Surgery</i> , 118(7), 109-113.	Specialized flap population.
Kroll, SS., Schusterman, MA., Reece, GP., Miller, MJ., Evans, GR., Robb, GL., Baldwin, BJ. (1996) Choice of flap and incidence of free flap success. <i>Plastic and Reconstructive Surgery</i> , 98(3), 459-63.	Did not isolate head and neck population.
Miller, MJ., Swartz, WM., Miller, RH., Harvey, JM. (1991) Cost analysis of microsurgical reconstruction in the head and neck. <i>Journal of Surgical Oncology</i> , 46(4), 230-4.	Take back rates and salvage rates not included.
Disa, J., Cordeiro, PG., Hidalgo, DA. (1999) Efficacy of Conventional monitoring techniques: an 11-year experience in 750 consecutive cases. <i>Plastic and Reconstructive Surgery</i> , 104(1), 97-101.	Did not isolate head and neck population.
Bozec, A., Poissonnet, G., Chamorey, E., . . . Dassonville, O. (2008) Free-flap head and neck reconstruction and quality of life: a 2 year prospective study. <i>Laryngoscope</i> , 118(5), 874-80.	Included in larger series.
Brown, JS., Magennis, P., Rogers, SN., Cawood, JI., Howell, R., Vaughan, ED. (2006) Trends in head and neck microvascular reconstructive surgery in Liverpool. <i>British Journal of Oral and Maxillofacial Surgery</i> , 44(5), 364-70.	Take back rates and salvage rates not included.
Heinz, TR., Cowper, PA., Levin, SL. (1999) Microsurgery costs and outcome. <i>Plastic and Reconstructive Surgery</i> , 104(1), 89-96.	Head and neck population not isolated.

<p>Schusterman, MA., Horndeski, G. (1991) Analysis of the morbidity associated with immediate microvascular reconstruction in head and neck cancer patients. <i>Head and Neck</i>, 13(1), 51-5.</p>	<p>Take back rates and salvage rates not included.</p>
<p>Brown, MR., McCulloch, TM., Funk, GF., Graham, SM., Hoffman, HT. (1997) Resource utilization and patient morbidity in head and neck reconstruction <i>The Laryngoscope</i>, 107(8), 1028-1031.</p>	<p>Take back rates and salvage rates not included.</p>
<p>Swartz, WM., Izquierdo, R., Miller, MJ. (1994) Implantable venous Doppler microvascular monitoring: laboratory investigation and clinical results. <i>Plastic and Reconstructive Surgery</i>, 93(1), 152-163.</p>	<p>Head and Neck population not isolated.</p>
<p>Clark, JR., McCluskey, SA., Hall, F., . . . Gilbert, R. (2007) Predictors of morbidity following free flap reconstruction for cancer of the head and neck. <i>Head and Neck</i>, 29(12) 1090-1101.</p>	<p>Included in larger study</p>
<p>Oliver, DW., Whitaker, IS., Giele, H., Critchley, P., Cassell, O. (2005) The Cook-Swartz venous Doppler probe for the post-operative monitoring of free-tissue transfers in the United Kingdom: a preliminary report. <i>British Journal of Plastic Surgery</i> 58(3), 366-370.</p>	<p>Head and Neck population not isolated.</p>

## Appendix B: Screen Captures from Online Utility Survey

Grab File Edit Capture Window Help

[SURVEY PREVIEW MODE] Utility Values in Head and Neck Surgery Survey

### Utility Values in Head and Neck Surgery

#### 1. Introduction

1 / 7 14%

Introduction

My name is Michael Gupta. I am a second-year student in the Master's in Epidemiology program and a Head and Neck Surgery Fellow in the Department of Otolaryngology at the University of Michigan. The data gained from this survey will be used to analyze whether the benefits of implantable Doppler technology (a special method used to measure blood flow) are worth the cost of my thesis project.

It will take 10 minutes to complete. The survey consists of brief instructions followed by four questions.

Thank you very much for your help. If you have any questions or comments about the survey please do not hesitate to email Michael Gupta @ [michaelkgupta@hotmail.com](mailto:michaelkgupta@hotmail.com)

Next

Powered by **SurveyMonkey**  
Create your own [free online survey](#) now!

## Utility Values in Head and Neck Surgery

### 2. General Information

2 / 7

#### General Information

- The following questionnaire is designed to determine how "health states" are valued.
  - Health states are descriptions of health status, i.e. mildly ill with a viral infection.
  - You will be asked to consider what length of life time in one health state you consider equally preferable to 10 months in another state.
  - All of the questions have the same structure:
    1. Two health states are described.
    2. This first health state involves a specific disease or disability, for example a leg amputation. This health state is always assigned a length of 10 months. In this scenario, you do not
    3. Next, a second health state is described. The second health state is always perfect health. There is no period of time assigned to the second health state.
    4. Finally, you are asked how much time in the second health state you consider equally preferable to 10 months spent in the first state.
- \*\* Answers may contain decimal points and can be any length of time between 0 and 10 months.

A sample question is next . . .

Prev Next

Powered by **SurveyMonkey**  
Create your own [free online survey](#) now!

### Utility Values in Head and Neck Surgery

#### 3. Sample Question

3 / 7

43%

#### Sample Question

You have a choice between the two health states described below.

##### Health State A:

- You have both of your legs amputated above the knee.
- You are confined to a wheelchair.
- This health state lasts 10 months and is followed by death.

##### Health State B:

- You are in perfect health for your age.
- You are able to do all activities that you wish to do.

How many months of life in health state B would you consider equally preferable to 10 months of life in Health State A?

Sample Answers: As answers will vary, I have provided three samples of answers and the logic behind these answers.

##### Sample Answer 1. 7 months.

In this case the respondent requires 3 extra months of life in Health State A before considering it equally preferable to the increased quality of life Health State B. The result is the respondent

##### Sample answer 2. 8 months.

This respondent values Health State A (with bilateral amputations) more highly than the previous respondent. As a result, they require only an extra 2 months of life in state A to be equivalent. The result is the respondent chooses 8 months as their answer

##### Sample Answer 3. 9 months.

In this case, the respondent values the health state with bilateral amputations even more than in Sample Answer 2. Here the respondent would chose to live only one month less before consi equivalent.

Starting on the next page, we will begin the survey.

Prev

Next

Powered by **SurveyMonkey**  
Create your own [free online survey](#) now!

### Utility Values in Head and Neck Surgery

#### 4. Question 1

4 / 7

57%

**\*1. Question 1.**

**Health State A:**

- You are in hospital.
- You have a tracheotomy tube (a breathing tube in your neck) and are not capable of speech.
- You are being fed through a feeding tube in your nose and are not able to take either liquids or solids by mouth.
- Your activity is very limited, you are capable of only very short walks and sitting in a chair.
- You are otherwise well
- This health state lasts 10 months and is followed by death.

**Health State B:**

- You are perfectly healthy for your age.
- You are able to do all things you wish to do.

How many months of life in health state B would you consider equally preferable to 10 months of life in Health State A?

Answer in box below (answers should be between 0 and 10 months):

Prev

Next

Powered by **SurveyMonkey**  
Create your own [free online survey](#) now!

Utility Values in Head and Neck Surgery

5. Question 2

5 / 7

71%

\*1. Question 2.

Health State A:

- You have had surgery for cancer of the head and neck.
- You are admitted to hospital.
- You have a tracheotomy tube (a breathing tube in your neck) and are not capable of speech.
- You are being fed through a feeding tube that enters your stomach through an incision in your abdomen.
- You are not able to take either liquids or solids by mouth.
- You require twice daily dressing changes for an open wound on your neck.
- You are capable of most non-strenuous normal activities.
- You are otherwise well
- This health state lasts 10 months and is followed by death.

Health State B:

- You are perfectly healthy for your age.
- You are able to do all things you wish to do.

How many months of life in health state B would you consider equally preferable to 10 months of life in Health State A?

Answer in box below (answers should be between 0 and 10 months):

Prev

Next

## Utility Values in Head and Neck Surgery

### 6. Question 3

6 / 7

86%

#### 1. Question 3.

##### Health State A:

- You have had surgery for cancer of the head and neck.
- You are at home.
- You are capable of slightly altered, but intelligible speech that can be understood on the telephone.
- You are able to eat most food.
- You are capable of most non-strenuous normal activities.
- You are otherwise well
- This health state lasts 10 months and is followed by death.

##### Health State B:

- You are perfectly healthy for your age.
- You are able to do all things you wish to do.

How many months of life in health state B would you consider equally preferable to 10 months of life in Health State A?  
Answer in box below (answers should be between 0 and 10 months):

Prev

Next

Powered by **SurveyMonkey**  
Create your own [free online survey](#) now!

### **Appendix C: Derivation of Revision Proportion in Decision Analysis Model**

1. The proportion of free flaps allowed to granulate in was derived from expert opinion that approximately 25% of free flap failures were left to heal with dressing changes and wound care.
2. The proportion of revised flaps treated with pedicled flaps was derived from expert opinion that held that two-thirds of flaps were revised with pedicled flaps.
3. The proportion of revised flaps treated with free tissue transfer was derived from expert opinion that held that one-third of flaps were revised with free tissue transfer.

## **Appendix D: Sample Calculations of Transition Probabilities**

### **TP1:**

TP1 was calculated as a daily transition probability. This is the calculation for the daily probability of a flap not being taken back for the first 3 days in the conventional group.

This value is denoted as TP13daysconv.

1. First, the probability of a flap not being taken back over the entire length of the first 3 days (totalnontbr3days) was calculated.

a) Portion of take backs in first three days (expert opinion) = prop3days = 0.95

b) Pooled take back proportion in conventional group (literature review) = tbr = 0.065

$$\text{totalnontbr3days} = 1 - \text{prop3days} * \text{tbr} = 1 - 0.95 * 0.065 = 0.938$$

2. Next daily probability of a flap not being taken back over the first 3 days (dailynontbr3days) was calculated.

a) Total probability of take back in the first three days = totaltbr3days = 0.938

$$\text{dailynontbr3days} = \text{totaltbr3days}^{(1/3)} = 0.979$$

3. Then the daily chance of successful salvage in the first three days (dailyprosalv3days) is calculated

a) Daily chance of take back in first three days = 1 - dailynontbr3days = 1 - 0.979 = 0.021 = dailytbr3days

b) Proportion of successful salvage (literature review) = 0.51 = sr.

$$\text{dailyprosalv3days} = \text{dailytbr3days} * \text{sr} = 0.021 * 0.51 = 0.011.$$

4. Finally, the TP1 for the first three days was calculated

a) Daily probability of not being taken back in first three days = 0.979 =  
dailyontbr3days.

b) Daily chance of salvage in the first three days = 0.011 = dailyprosalv3days

TP13daysconv = dailyontbr3days + dailyprosalv3days = 0.979 + 0.011  
= 0.99.

## **TP2:**

TP2 was also calculated as a daily probability. It was the daily probability of successful revision. It is denoted as dailyrevsuccess.

1. First the likelihood of successful revision (revsuccess) was calculated.

a) Chance of having a FTT revision (expert opinion) = 0.25 = chanceFTT.

b) Probability this FTT revision is successful (expert opinion) = 0.8 =  
successFTT.

c) Chance of having a pedicled flap revision (expert opinion) = 0.50 =  
chancepedicle

Revsuccess = (chanceFTT\*successFTT) + (chance pedicle)

= (0.25\*0.8) + (0.5)

= 0.7

2. Second the likelihood of successful revision (revsuccess) is multiplied by the chance of having surgery on a particular day to give the daily chance of successful revision

a) Chance of having surgery = chancesurg = 1/(average wait for surgery = 3 days  
(expert opinion)) = 0.333

b) Chance of successful revision from above = revsuccess = 0.7

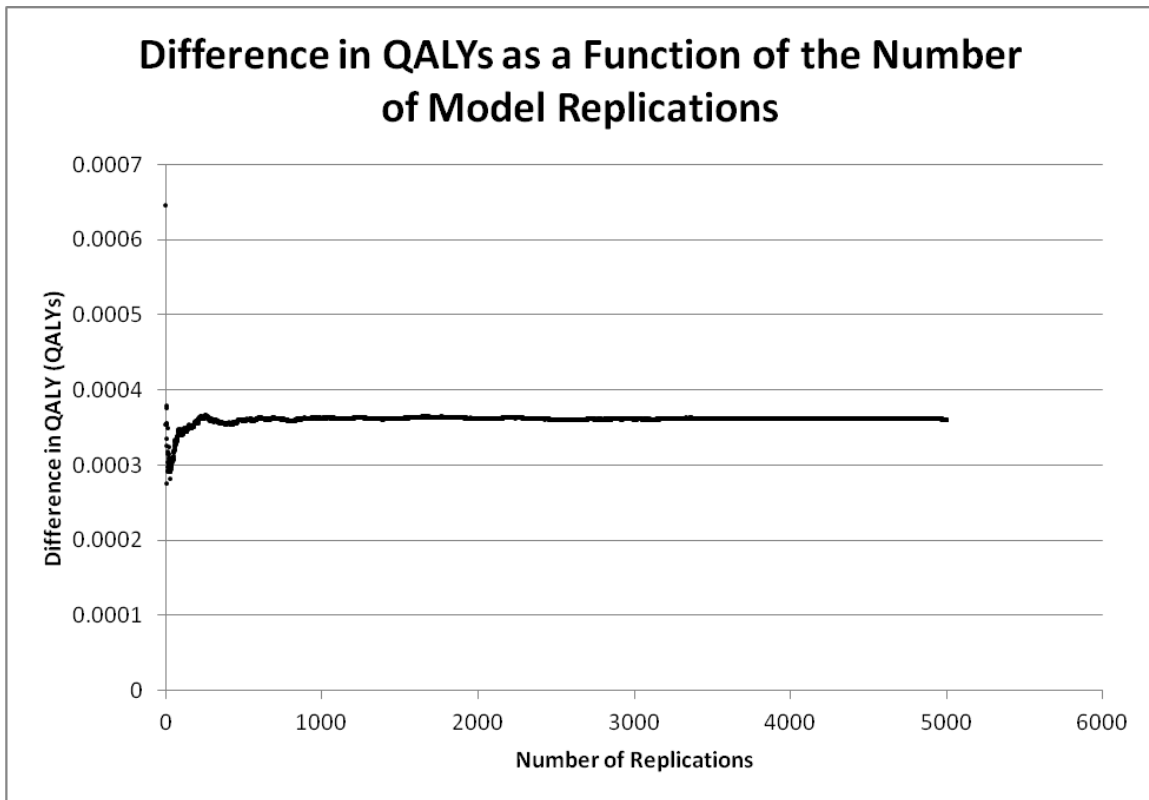
$$\begin{aligned} \text{Dailyrevsuccess} &= \text{chancesurg} * \text{revsucces} = 0.7 * 0.33 \\ &= 0.231. \end{aligned}$$

**TP3:** This transition probability refers to the daily probability of discharging a patient who has a failed FTT and is waiting revision or granulation. The probability is calculated based on the average length of stay of patients who have failed. In the deterministic case it is calculated as zero everyday until the day at which the average length of stay of failures is reached at which case it is calculated as one.

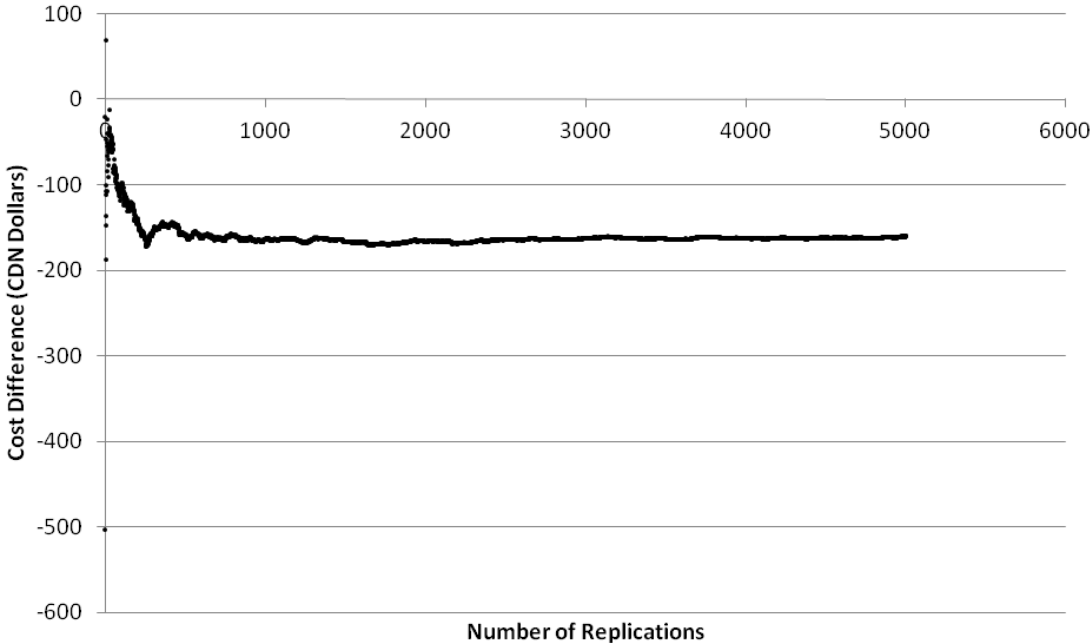
**TP5:** This is calculated in the same manner as with the exception that the average length of stay of successful FTT is used.

## **Appendix E: Results of Decision Analysis as a Function of the Number of Model Replications**

This Appendix graphically displays the results of the economic evaluation as a function of the number of replications of the model. The simplest case with a single Doppler probe is used as an example.



### Difference in Cost as a Function of the Number of Model Replications



**Appendix F: Clinical Pathway of Free Tissue Transfer in the Head and Neck at the**

**Ottawa Hospital**

**Clinical Pathway of FTT at TOH**

