Thyroid Fine-Needle Aspiration Biopsy: The Winnipeg Experience

by

Christina R. Lang

A practicum submitted to the Faculty of Graduate Studies of

The University of Manitoba

in partial fulfilment of the requirements of the degree of

MASTER OF SCIENCE

Faculty of Medicine
Department of Pathology
University of Manitoba
Winnipeg, Manitoba

THE UNIVERSITY OF MANITOBA

FACULTY OF GRADUATE STUDIES

*****

COPYRIGHT PERMISSION

Thyroid Fine-Needle Aspiration Biopsy: The Winnipeg Experience

By

Christina R. Lang

A Thesis/Practicum submitted to the Faculty of Graduate Studies of The University of Manitoba in partial fulfillment of the requirement of the degree

Of

Master of Science

ACKNOWLEDGEMENTS

I would like to thank Dr. H. R. Wightman for his continued guidance and patience during the evolution of this practicum project. I would also like to thank Dr. E. Ravinsky, Mr. L. Fuczek, and Dr. D. Inglis for their interest and guidance in the undertaking of this project.

I would like to thank my parents, Ellen and Robert Lang, for their continuous encouragement and support and for their steadfast belief in my ability to succeed academically, regardless of the educational and financial obstacles encountered along my journey. To my grandfather H.F. Lang, your love and unwavering belief in me has given me the courage and determination to face every academic challenge. I would like to thank my sister, Victoria, for her continuous support. To my brother, Michael, who has always stood by me, and whose phone calls and supportive text messages have given me the strength and encouragement I needed to complete this latest academic journey, I extend my greatest thanks.

I would like to thank Lynda and Ed Taylor for opening their home and their hearts to me. Their continued friendship, support, and encouragement has been an unexpected and greatly appreciated blessing.

Abstract

Objectives: To investigate the accuracy of thyroid fine needle aspiration (FNA) biopsies in Winnipeg and the relative risk of malignancy associated with each FNA category. To provide evidence in support of the separation of the "indeterminate" cytology category into high and low risk malignancy categories.

Materials and methods: A retrospective search for thyroid FNA reports and corresponding histology reports was initiated. Cytological-histological correlations evaluated the accuracy of each FNA diagnosis and the associated risk of malignancy.

Results: Eight hundred and ninety-three thyroid FNA cytology reports were obtained with 31% being, 22% follicular lesion (FL), 6% follicular neoplasm (FN), 7% suspicious for malignancy, 4% malignant, 29% unsatisfactory. Histologically-confirmed malignancy rates were benign 5%, FL 17-19%, FN 46%, suspicious for malignancy 80%, malignant 93%, unsatisfactory 12.5%.

Conclusions: Thyroid FNA biopsies are performed accurately in Winnipeg and malignancy rates are consistent with published literature values. The differences in the malignancy rates of FL (17-19%) and FN (46%) categories supports the separation of indeterminate category.

Table of Contents

I. Table of Figures

II. Table of Tables

III. List of Abbreviations

ACL..... atypical cellular lesion

HC.....Hurthle cell carcinoma

HN.....hyperplastic nodule

HSC.....Health Sciences Centre

FC.....follicular carcinoma

FL.....follicular lesion

FMTC.....familial medullary thyroid carcinoma

FN.....follicular neoplasm

FNA.....fine needle aspiration

FVPC.....follicular variant papillary carcinoma

MEN-II.....multiple endocrine neoplasia type II

MNG.....multi-nodular goiter

MTC.....medullary thyroid carcinoma

NCI.....National Cancer Institute

PHIN.....Personal Health and Information Number

PPV.....positive predictive value

PTC.....papillary thyroid cancer

RET.....rearranged during transfection[59]

SBGH.....St. Boniface General Hospital

TMG.....toxic multinodular goiter

TSH.....thyroid stimulating hormone

US.....ultrasound

WHO..... World Health Organization

IV. Introduction

The thyroid gland is a peripheral endocrine gland which lies in the anterior neck over the trachea just below the larynx and is palpable during swallowing.[1] Weighing only 15 to 25 grams, the thyroid gland alone is responsible for producing and circulating enough thyroid hormone to control the metabolic rate of every cell in the human body, with the exception of the brain, spleen, testes, and uterus.[1, 2] Like any other organ of the body, abnormalities of the thyroid, particularly those producing palpable nodules are of concern to both patients and physicians. Worldwide, 4-7% of adults have a palpable thyroid nodule and with the use of advanced diagnostic imaging, particularly ultrasound, this number increases significantly.[3, 4, 21, 26] In the past, thyroid nodules were most often identified by palpation. With the introduction of diagnostic imaging as many as 30-40% of the general public have been found to have incidental non-palpable nodules.[4] On average approximately 5% of thyroid nodules are malignant.[25, 26] Various studies have documented malignancy rates from 5% to 38%.[5, 6] With potentially hundreds of millions of individuals worldwide having abnormal thyroid glands and an average of 5% of abnormal thyroids harboring carcinomas (the remaining 95% being benign) physicians have turned to multiple laboratory and diagnostic tests to help differentiate between malignant and benign conditions.[7-10]

Thyroid cancer accounts for only 1% of visceral cancers and even fewer deaths.[3, 12] When all cancers were considered together the World Health Organization (WHO) estimated that 7 million individuals worldwide died of cancer in 2000 and less than 0.5% of these deaths were due to thyroid cancer.[3, 12, 26] Interestingly, the total number of cancer deaths in 2000 is only about one-half to one-third the number of Americans with thyroid

nodules.[3, 12, 26] It is estimated that 10 to 20 million Americans have a clinically detectable thyroid nodule,[11] but the lifetime probability of developing a thyroid malignancy is 0.22% and 0.6% in males and females respectively.[4, 11, 15] The lifetime probability of dying from thyroid cancer is even lower, 0.03% in males and 0.06% in females.[4] Thyroid cancer accounts for only 1% of visceral cancers and even fewer deaths.[3, 12] It is estimated that 4300 Canadians will be diagnosed with thyroid cancer in 2008 and of these 180 will die from the disease.[13] Since 1997, Canada has seen a 5.5% and 10.1% increase in the number of reported thyroid cancers in men and women respectively, while the mortality rate has remained unchanged. Early detection of malignant nodules, through ultrasound and FNA biopsies, has increased the number of thyroid cancers diagnosed.[13] However, a steady mortality rate has been achieved through the effective surgical excision of these early cancers.[13] Between 1960 and 1991 American statistics reported an increase in the five year survival rate of thyroid cancer from 83% to 95%.[4,20] The most recent edition of Canadian Cancer Statistics cites the five year survival rate at 98%, giving thyroid cancer the highest five year survival rate of any malignancy.[13]

A majority of thyroid cancers arise in the follicular epithelial cells and include follicular, papillary, and anaplastic carcinomas.[14] Medullary thyroid carcinomas (MTC) arise from the parafollicular cells.[14] Papillary thyroid carcinoma (PTC) is the most common form of thyroid cancer and accounts for 75-85% of thyroid cancers.[14] The features of PTC are varied; these tumors can be solitary or multifocal, encapsulated or invasive. Follicular carcinomas (FC) are less common and account for only 5-20% of thyroid cancers.[14] Most often occurring as a single nodule, encapsulated FC are distinguished from their benign counterparts by their invasion through the capsular wall

or into adjacent vessels.[14] Diffusely invasive FC are rare and more aggressive than their encapsulated counterpart.[14] Malignancies of the parafollicular cells (medullary carcinoma) account for approximately 5% of all thyroid cancers.[14] Both familial and sporadic forms of MTC have been identified.[14] Familial medullary thyroid carcinoma (FMTC) is the result of a mutation in the rearranged during transfection[59] (RET) oncogene, passed from one generation to the next through an autosomal dominant pattern of inheritance, and it accounts for 25% of medullary cancers.[14, 16] The remaining 75-80% of medullary cancers are sporadically arising parafollicular malignancies.[14] Papillary, follicular, and medullary carcinomas represent differentiated thyroid cancers, in which the malignant cells still bear some resemblance to their cells of origin.[14] Differentiated thyroid cancers which have an excellent prognosis include papillary and encapsulated follicular carcinomas, while medullary and diffusely invasive follicular carcinomas carry a good prognosis.[14] Conversely, anaplastic thyroid carcinoma represents an undifferentitiated follicular epithelial tumor.[14] These rapidly enlarging aggressive tumors account for less than 5% of thyroid cancers and patients rarely survive, most dying within one year of diagnosis.[14]

A thorough patient history and physical examination are the first steps in assessing a patient with a thyroid nodule and determining which ancillary tests may provide useful information regarding the risk of malignancy for the nodule in question.[16] A rapidly enlarging nodule or sudden changes in the size of a longstanding nodule are indicators of a potential malignancy. Airway compression, hoarseness (vocal cord paralysis), and dysphagia may indicate invasion into adjacent tissue.[16, 26] Radiation exposure is a known environmental factor which contributes to the development of

palpable nodules and cancer.[16, 25] Radiation exposure damages cell DNA and therefore malignancy rates are higher in these nodules, ranging from 20-50%.[12, 24] During the 1950's and as late as the 1970's, therapeutic radiation was a common treatment for enlarged tonsils, adenoids, or thymus as well as skin infections of the neck.[16] When therapeutic radiation was a common treatment, the rate of thyroid nodules in children was 35-40%, significantly higher than the 0.2-1.4% of affected children today.[3,17] It is important to identify immigrants who lived in Russia, Ukraine, and Belarus at the time of the Chernobyl accident of April 26, 1986 and inquire if they were exposed to radiation, for they have a higher risk of malignancy as well.[16] The increased incidence of thyroid cancer in these patients is the result of exposure to radioactive iodine (a product of the nuclear fission process)[60] concentrated in the milk of cows grazing on vegetation exposed to the Chernobyl nuclear fallout.[18, 19] Because the cells of the thyroid take up and concentrate the radioactive iodine in the ingested milk the thyroid gland receives a radiation dose 1000 times greater than the rest of the body. [18, 19] Children were the most affected not only because their growing thyroid glands take up more iodine than adults, but also because they consumed more contaminated milk, exposing their growing thyroid glands to more radioactive iodine.[18, 19]

With increasing genetic associations being identified, a family history of benign or malignant thyroid pathology should be sought. This is especially important when a history of multiple endocrine neoplasia type II (MEN-II) is suspected.[16] MEN-II is a genetically inherited syndrome in which endocrine organs are prone to developing lesions which are at a higher risk of being malignant.[14, 16] FMTC is a subtype of MEN-II in which the thyroid gland is the only endocrine gland prone to developing neoplastic

lesions.[14] In these patients the risk of a thyroid nodule being malignant is significantly higher than in the general population. A family history of MTC, parathyroid hyperplasia or pheochromocytoma (present in 50% of MEN-IIa and MEN-IIb patients) is an indication for genetic testing. MEN-II and FMTC are linked to a mutation in the RET oncogene, and RET proto-oncogenic testing identifies more than 98% of patients with this gene mutation.[12, 16] Given an autosomal dominant pattern of inheritance, genetic testing of all first degree family members is recommended following identification of any MEN-II or FMTC patients with RET mutations.[16] Early identification of family members with the RET mutation, especially children, allows for early intervention in the form of prophylactic thyroidectomy. A prophylactic thyroidectomy greatly reduces the risk of cancer development and spread to regional lymph nodes, and the earlier the surgery is performed the greater the survival rate of the patient, and the need for a central neck dissection is less likely.[16] Research has shown that the position and type of RET mutation can be used to predict the disease phenotype and a time frame in which the malignancy is likely to develop. This enables clinicians to determine when a prophylactic thyroidectomy should be scheduled. However, the disease does not always follow the predicted course and a malignant thyroid nodule and metastasis to regional lymph nodes may occur sooner than expected. The yearly testing of calcitonin levels has been suggested as a method to identify cases in which earlier resection may be necessary.[16] Optimal results are achieved when the surgery is initiated before the age of five, and may be performed in children as young as six months.[16] In general, RET testing is performed on all patients diagnosed with MTC even in the absence of a family history.

In the absence of a genetic mutation MTC is considered to be sporadic in nature, with approximately 75% of MTC being sporatic.[12, 14, 16]

Patients with a thyroid nodule and a history of multiple benign or malignant tumors throughout the body also have an elevated risk of thyroid cancer.[16] In these cases both clinical information and patient history may be important in determining the risk of malignancy. Multiple tumors are common in Cowden and Gardner syndrome and patients with these autosomal dominant syndromes are at an increased risk for thyroid cancer.[16, 22] Multiple benign and neoplastic lesions are commonly identified in patients with Cowden syndrome and two-thirds of these patients have abnormalities of the thyroid gland.[22] Thyroid cancer is the second most common malignancy in these patients and their life-time risk of developing thyroid cancer is 7 to 10%, substantially higher than that of the general population (approximately 5 %).[22] Gardner syndrome is an uncommon variant of familial polyposis which affects 1 in 1million Americans.[23] With potentially malignant tumors arising in every body system, individuals who inherit this rare autosomal dominant condition are rigorously screened for an extensive list of potential malignancies, including thyroid cancer. Because of the increased risk of multicentric PTC, annual clinical and US evaluation of the thyroid is mandatory.[23]

Men and women differ not only in the frequency with which thyroid nodules occur, but also in the lifetime risk of developing and dying from thyroid cancer.[4, 16] More women are affected by thyroid cancer than men, as is evidenced by Canadian cancer statistics reporting 1 in 81.3 women will be diagnosed with thyroid cancer in their lifetime and 1 in 1,160 will die of their disease.[13] Thyroid cancer is not even one of the top 16 most commonly diagnosed cancers in Canadian men.[13] In 2004 3,237 new cases

of thyroid cancer were reported in Canada from 673 men and 2564 women.[13] While more women are diagnosed with thyroid cancer the morality rate is higher in men. Of the 673 men diagnosed with thyroid cancer, 1 in 10 died while only 1 in 25 of the women died.[13] Women are four times more likely to develop a thyroid nodule and two times more likely to develop a malignancy.[3] However, a man with a thyroid nodule is at greater risk of thyroid cancer because there are proportionally fewer men with thyroid nodules, with any single nodule demonstrating a higher risk of malignancy.[3]

Thyroid cancer can occur at any age but there are significant trends in the incidence of malignancy within certain age groups which allow a patient's age to be statistically useful when assessing a thyroid nodule for malignancy.[3, 16] Thyroid nodules are uncommon in children (0.2-1.4%), but are more likely to be malignant (15-20%) than nodules in young to middle-aged adults.[3, 17] Most thyroid cancers occurring in children are slow-growing low-grade PTC, and if detected early, these children have a better prognosis than their adult counterparts who are diagnosed after the age of 40.[3, 17] Thyroid nodules are more common in adults, especially between the ages of 30 and 50, and the rate of thyroid cancer also increases in this group. With an increase in both the incidence of thyroid nodules and thyroid cancer in adults the risk of malignancy is approximately 5%, significantly lower than the 15-20% risk of malignancy in childhood thyroid nodules. After the age of 60 the incidence of mortality increases, with an additional increase in the percentage of thyroid cancer.[3, 4] A thorough investigation is essential in children, because they are at an increased risk for developing cancer and early detection improves their survival rate, but older individuals must be aggressively investigated as well, because their prognosis and mortality rates are significantly elevated. [3, 4] In summary, patient

history, family history, gender, and age are all significant factors in assessing thyroid nodules.

In the presence of a palpable thyroid nodule, physical examination of the thyroid and neck can provide the clinician with information about the number of nodules, size, consistency, and invasion into adjacent structures.[25] Invasion of the recurrent laryngeal nerves often presents as vocal cord paralysis and assessment of vocal cord mobility may be helpful in identifying invasive nodules.[16] When palpating the neck of patients with palpable and non-palpable thyroid nodules, special attention should be given to the assessment of the cervical lymph nodes.[16, 24] Enlarged, firm, or fixed lymph nodes may be an indicator of metastatic disease, and abnormal cervical adenopathy is common in patients with PTC, where lymphatic invasion is the most common form of metastatic spread.[14, 16, 24, 25]

Patient history, gender, age, and family history are important in assessing the potential for cancer in a thyroid nodule.[11] Although aggressive or high grade thyroid malignancies may be obvious on clinical examination, most thyroid cancers present as solitary nodules without signs of malignancy. Surgical excision and histological examination of any thyroid nodule is the only way to absolutely rule out a malignant lesion. However, there are millions of adults with thyroid nodules[3, 4, 21] and surgically removing millions of thyroids in order to ensure all malignant nodules (5%)[24,25] are excised is unrealistic and surgery is not without significant risks.[33,36] With this in mind, additional ancillary tests have evolved to further assess the risk of malignancy. These include ultrasonographic imaging, radioisotope scanning, and FNA biopsies.[11, 16]

Ultrasonographic imaging is a relatively recent addition to the armamentarium of diagnostic tools available for the evaluation of thyroid nodules. It provides detailed information regarding the size, location, composition, and appearance of thyroid nodules.[25] Due to the high number of non-palpable clinically-insignificant thyroid nodules within the general population which would be evident on ultrasound (US), this test is not recommended for universal screening of thyroids.[25] It is, however, recommended for patients with a palpable nodule, history of head and neck radiation, a family history of MTC, MEN or PTC, or unexplained cervical adenopathy.[25] Under these circumstances US may be used to accurately measure nodules, determine consistency (solid, cystic, or mixed), assess nodule borders, identify solitary nodules within a multinodular goiter, and identify any additional non-palpable nodules.[16, 25] While US cannot be used to distinguish benign from malignant nodules, microcalcifications, irregular boundaries, hypoechogenicity, and chaotic intranodular vasculature are features which may raise the index of suspicion of the radiologist that the nodule may have an increased risk of malignancy, especially when more than one feature is present.[16, 25, 26] On their own, no one of these features is consistently predictive of malignancy, but when two or more features are present the risk of malignancy increases.[16, 25] Extracapsular extension is of particular concern especially when irregular borders appear to extend beyond the thyroid into adjacent structures.[25] These findings may warrant cytological evaluation of the nodule or may be used to overrule a benign FNA diagnosis.[25]

One of ultrasound's most important roles is its use in ultrasound-guided FNA biopsies. Here US provides visual confirmation that the clinically significant nodule is the nodule being sampled.[25] When biopsying a cyst or complex nodule, US can also be

used to distinguished solid from cystic components, thereby enabling clinicians to target solid areas where they are more likely to obtain a sufficient number of follicular cells to allow a diagnosis, compared to cystic fluid which contains few cells and is often non-diagnostic.[24, 25]

Prior to the introduction of FNA biopsies, thyroid scintigraphy (a nuclear medicine radioisotope scan) played a much more significant role in discriminating between benign and malignant nodules.[50, 51] A "hot" nodule was considered to have a low malignant potential (essentially benign) while a "cold" nodule was at greater risk for malignancy and surgical intervention was often initiated.[16] A radioisotope scan of the thyroid is used to determine if the nodule is functioning (hot-producing thyroid hormone) or non-functioning (cold). A thyroid scintigraph is based on the premise that thyroid follicular cells are the only cells in the body which take up iodine. Iodine is then combined with the amino acid tyrosine to produce the thyroid hormones thyroxine (T4) and triiodthyronine (T3).[27] By injecting radioactive iodine isotopes and tracking the uptake of iodine by the thyroid, a radiologist is able to identify normal, hypofunctioning, and hyperfunctioning areas. A hyperfunctioning "hot" nodule is manufacturing more thyroid hormone than the surrounding normal tissue and takes up more iodine, appearing as a concentrated dark spot on the thyroid scan. Conversely, a "cold" nodule does not take up radioactive iodine and therefore appears as a void on the thyroid scan.[27, 28] Hot nodules account for approximately 5% of nodules while cold nodules account for about 80-90%.[25, 26] A normal functioning thyroid is "warm" on scintigraphy and about 5-15% of thyroid nodules are warm with 9% of warm nodules being malignant.[24, 27, 28] Hot nodules are rarely malignant (1-4%)[16, 25, 30] while 5-15% of cold nodules are malignant.[16,

[25] If every cold nodule was removed for suspicion of malignancy, many patients would be subjected to needless surgery.[24-26] Thus, scintigraphy only spares approximately 5% of patients (those with hot nodules) but the remaining pool of patients with cold nodules (80-90%) have similar statistical odds of having a malignancy to patients who have not undergone scintigraphy. Radioisotope scanning cannot be used to distinguish between malignant and benign nodules. [16] In Europe, scintigraphy continues to contribute valuable information to the overall clinical evaluation of the malignant potential of a nodule. However, in North America, its role is largely limited to assessing the degree of nodular function, especially in patients with suppressed serum thyroid stimulating hormone (TSH) levels. [26, 29, 30] Graves disease and a toxic multinodular goiter (TMG) both produce excess thyroid hormones, suppressing serum TSH levels, and are associated with hyperthyroidism.[30] A solitary concentration of iodine uptake in an already diffusely "hot" thyroid gland (Graves disease) or multiple "hot" nodules in a normal or cold thyroid gland of a TMG (autonomous hot nodules can suppress iodine uptake in the remaining thyroid tissue) on scintigraphy confirm benign conditions and patients are treated medically rather than surgically.[30] Scintigraphy may also be beneficial in identifying ectopic thyroid tissue or substernal extension of a large multinodular goiter.[30] Due to the high percentage of "cold" nodules on radioisotope scan the test is of limited value in identifying thyroid malignancies. In Europe, an area of the world with a higher incidence of "hot" nodules, thyroid scanning continues to play a role in differentiating those patients who are treated clinically (hot nodules) from those who require further investigation.[29, 30]

In practice, the assessment of a thyroid nodule involves a patient history, physical examination, and FNA biopsy, the later being performed with or without ultrasound guidance. The introduction and use of FNA biopsies in the assessment of thyroid nodules has been credited with not only reducing the number of thyroid nodules requiring surgical intervention but also with increasing the number of thyroid glands removed for malignant tumors.[5, 11, 39] FNA biopsies of the thyroid were first introduced by Soderstrom in 1952 but did not gain popularity in North America until the late 1970s.[11] Prior to the introduction of FNA biopsies, triaging thyroid nodules was primarily based on a collection of clinical symptoms and diagnostic test results from which the clinician favored a benign or malignant diagnosis.[50, 51] Because of the imprecise nature of the clinical diagnosis, only 10-25% of the thyroids surgically removed for a suspected malignancy actually harbored a malignancy.[3, 50, 51] With the introduction of FNA biopsies, the incidence of malignancy following surgical resection has doubled, with some investigators reporting as many as 50% of surgical resections demonstrating cancer. However, most report malignancy values between 15-50%.[31] The primary purpose of the various thyroid diagnostic tests is to screen the millions of thyroid nodules and identify those most likely to harbor a malignancy.[6, 15] Because an FNA biopsy directly samples nodular tissue it can, in many situations, provide a diagnosis and is often the first test recommended for those with a thyroid nodule.[11, 32] While surgery is the only absolutely accurate way to assess nodules, it is not only impossible to surgically remove every nodule, the surgery itself is not without risk.[32] In its infancy, thyroid surgery mortality rates were largely attributed to damage to and loss of parathyroid glands.[36] This was due to a poor understanding of both anatomical relationships between the thyroid and

adjacent neck structures and also of parathyroid physiology. Hypoparathyroidism and vocal cord paralysis were two of the most common complications of surgery, because of the anatomical proximity of the thyroid gland to the parathyroid glands and the recurrent laryngeal nerve.[36]

Lying just behind the posterior thyroid capsule, the parathyroid glands can be damaged, accidentally removed, or their blood supply inadvertently compromised during surgery.[2, 3, 36, 37] The parathyroid gland is responsible for regulating and maintaining calcium levels within the body, and calcium ions are essential for the conduction of nerve impulses, and are therefore responsible for insuring the body's neurons and muscles are functioning properly.[34] In the absence of parathyroid hormone, hypocalcemia results, causing sensory neuron excitability (perioral numbness, tingling, abdominal pain, and nervousness) and motor neuron excitability (paraesthesias of extremities and muscle cramping).[33, 34] Severe hypocalcemia can be fatal, especially when the muscles controlling the respiratory system go into tetanic spasm.[34] Most patients experience some degree of hypoparathyroidism following thyroidectomy surgery. The number of parathyroid glands involved and the extent of damage they sustain determines whether the condition is temporary or permanent.[2, 33, 35, 36] When calcium ion levels fall, parathyroid hormones act on the kidneys, bones, and the intestines to return calcium levels to normal, making the medical management of calcium levels extremely difficult, and motivating careful surgical dissection around the thyroid.[34] Autotransplantation of at least one parathyroid gland to the sternoclidomastoid muscle is one technique surgeons are using to reduce the likelihood of permanent hypocalcemia in patients undergoing a total thyroidectomy.[33, 35]

The recurrent laryngeal nerves run laterally along the edges of the thyroid gland and are responsible for controlling the laryngeal muscles which produce vocal cord vibration (voice production).[2, 37] Unilateral vocal cord paralysis results in a poor voice or hoarseness of the voice, while bilateral vocal cord paralysis results in almost no voice production.[37] Thyroidectomy is the third leading cause of iatrogenic bilateral vocal cord paralysis.[33] In the last several decades, the understanding of anatomical relationships and parathyroid physiology has dramatically reduced the morbidity rates and almost eliminated the risk of death associated with thyroid surgery. Although the morbidity and mortality rates associated with thyroidectomy surgery are relatively low, injury to the recurrent laryngeal nerve and hypoparathyroidism remain potential risks in every surgery.[33, 36] FNA biopsies have made great progress in sparing those patients with a benign diagnosis the risk of surgery.[5, 11, 39] While ultrasound provides feedback about the composition of the lesion, changes in its size, or the presence of calcifications,[25, 26] and radioactive scans can indicate if the nodule is functional or not,[26] FNA biopsies give specific information about the cytologic nature of the lesion, providing a definitive or differential diagnosis of "benign" or "malignant". FNA biopsies provide a definitive malignant diagnosis for PTC, MTC, and anaplastic cancer, and a definitive benign diagnosis for routine colloid nodules.[3] A differential diagnosis is most often associated with hypercellular follicular lesions and Hurthle cell proliferations.[32] In these cases the distinction between a benign hyperplastic nodule (HN), adenoma, and a carcinoma cannot be made on cytology alone. The nodule is most often called a "follicular/Hurthle cell lesion" and the differential diagnosis typically includes HN versus a FN.[32] The term neoplasm is used with both follicular and Hurthle cell lesions because the distinction

between the benign adenoma and malignant carcinoma is based on capsular or vascular invasion, a feature which can only be assessed on histology.[32] Without recourse to surgically removing every thyroid nodule, FNA biopsies are currently the most accurate and sensitive test available, and provide diagnostically significant information regarding the malignant potential of the nodule. [3]

An FNA biopsy can be performed by a clinician (or a radiologist if an ultrasound-guided biopsy is required) and can be completed in as little as twenty minutes.[38] The patient lies in a supine position exposing the neck, and the nodule to be biopsied is palpated. [38] A radiologist may use US to locate a non-palpable nodule and to insure the intended nodule is indeed sampled.[38] Once the nodule has been identified, the area is sterilized and a 23-25 gauge needle is inserted into the nodule, typically without the need for a local anesthetic.[38] In order to obtain a satisfactory sample of follicular epithelial cells, the needle is gently rocked within the nodule before being withdrawn.[38] The vascular nature of the thyroid gland often results in the aspiration of too much blood. By using a "non-aspiration" technique in which negative pressure is not applied to the syringe the aspiration of excessive blood is reduced.[39] The aspirated material is applied to microscope slides using the two-slide contact and smear technique. One of the paired slides is air-dried while the other is fixed in 95% ethanol. The air-dried slide is stained using Diff-Quik while the alcohol-fixed slide is stained using the Papanicolaou technique. The Diff-Quik slides can be stained immediately and the specimen can be assessed by the cytopathologist at hand for an adequate population of follicular cells before the patient has left. Assessing specimen adequacy while the patient is still present can be beneficial to both the patient and the clinicians handling the patient's care.[29] Immediate rebiopsy

can often provide an adequate specimen, sparing the patient the stress of a non-diagnostic/unsatisfactory result and a second trip to have the nodule rebiopsied. Clinicians are also better able to provide appropriate care to the patient when an satisfactory specimen is received and a diagnosis is made.

Assessing whether the specimen received is sufficient to provide a diagnosis is the first step in the diagnostic process.[12] There are various criteria (all of which are based on the number of follicular cells present in the sample) for determining those specimens which are satisfactory. Following examination by a pathologist, the sample may be placed in a number of diagnostic categories depending on the presence or absence of specific cytological features. Because the main role of an FNA biopsy is to aid clinicians and surgeons in evaluating and selecting the most effective treatment for each patient, the categories used by pathologists to classify FNA biopsies play a critical role in the type of treatment selected, particularly the need for surgery (total or partial thyroidectomy).[7, 44] Currently there are no established guidelines for the classification of FNA biopsies, with individual hospitals and research centers each appearing to have their own systems. Classification systems can vary from 2 categories (benign and malignant) up to 6 or more categories (which rate the amount of cellular material present and grade it as to the degree of atypical features[40,41]).[49] While pathologists vary in the descriptive terminology they use to classify FNA biopsies, clinicians vary in the treatment plans which they implement for each classification.[44] In the literature there is consensus with regard to the surgical excision of malignant FNA biopsies and clinical follow-up for benign nodules.[42, 45,46] When it come to suspicious or indeterminate categories, some clinicians advocate the

removal of all suspicious or indeterminate nodules[25, 40, 57] while others favor observation, follow up, and repeat FNA.[42, 45, 46, 49]

Both false positive and false negative results are of concern in thyroid cytology. A false negative biopsy can be problematic, delaying a diagnosis and resulting in significant patient morbidity and mortality.[25, 50] A false positive result is felt to be of less concern by many clinicians, particularity given the lower rates of morbidity associated with thyroid surgery and the recent and relative ease of thyroid hormone replacement.[50] This is particularity true when the surgical intervention is a partial thyroidectomy. Many feel that a false positive result is of less concern than a false negative.[50] A false positive biopsy may result in an unnecessary surgery but prior to the introduction of FNA biopsies these patients would have undergone surgery anyway. A false negative result is of more concern to both clinicians and patients because it means a cancer was missed.[50] Thus, a false positive has the potential for morbidity, but a false negative has the potential for mortality.

The easiest nodules to treat are those diagnosed as malignant on FNA. With a clear cut diagnosis, all clinicians agree surgery is the appropriate treatment course.[24, 25,46] The treatment of choice for a carcinoma diagnosis on FNA is surgical intervention, in the form of a total or subtotal thyroidectomy.[24, 42] The most common exception is an FNA diagnosis of anaplastic thyroid carcinoma.[24] The aggressive nature of anaplastic lesions, (rapid growth, invasion into adjacent neck structures, and metastasis to distant organs) means the life expectancy of most patients is measured in months rather than years.[14] For this reason surgical excision is rarely advised, unless the nodule is compressing the patient's airway.[14, 16, 24] In place of surgery, the most common alternative therapies

include radiotherapy and chemotherapy.[24] A subtotal thyroidectomy, rather than a lobectomy, is the preferred treatment for a diagnosis of PTC on FNA, because the multicentric/multifocal nature of the malignancy often means the other lobe harbors malignant cells.[24, 43] A lobectomy may remove the primary lesion, but a subtotal thyroidectomy removes the potential for recurrence if additional malignant foci are present in the other lobe. Research has shown that local recurrence rates and nodal metastasis are significantly lower (2% and 6% respectively) in patients who underwent a total or subtotal thyroidetomy compared with those who had a unilateral lobectomy (14% and 19% respectively).[43] A subtotal thyroidectomy is preferred over a total thyroidectomy because of the complex mechanisms regulating and maintaining calcium and phosphate homeostatsis controlled by the parathyroid glands.[43] These mechanisms are essential to life and are difficult to maintain exogenously. Leaving as little as 200mg of thyroid tissue can insure the integrity of at least one parathyroid gland.[43] Following a histological diagnosis of PTC some researchers recommend ipsilateral central compartment (level VI) lymph node dissection, in addition to completion thyroidectomy.[24] A lymph node dissection is recommended because lymphatic invasion is the most common form of metastatic PTC spread and these patients are at a higher risk of regional lymph node metastasis.[14, 24, 61] An ipsilateral or bilateral modified neck dissection is recommended for patients with positive level VI lymph nodes and can reduce the risk of recurrent metastatic spread and mortality.[61] As with PTC, a subtotal thyroidectomy, as well as a central compartment (level VI) lymph node dissection, is recommended for those with an FNA diagnosis of MTC.[24]

Surgery is avoided, in the absence of airway compression or cosmetic concerns, following a benign FNA diagnosis, and the current treatment protocol is continued clinical observation.[24, 25] Benign nodules on FNA continue to be monitored because, unlike malignant nodules which are surgically removed (thereby minimizing the risk of recurrent or metastatic disease) benign nodules remain in situ and are associated with a low but relevant false negative rate of approximately 5%.[6, 24, 50] Continued observation may include measuring serum thyrotropin levels, neck palpation, ultrasonography or repeat FNA biopsy.[24,26] There are circumstances in which clinical findings (a family history of MTC or MEN-II, a rapid increase in nodular size, a firm or hard nodule, fixation of the nodule to adjacent structures, vocal cord paralysis, regional lymphadenopathy, males, patients younger than 20 or older than 70 years, or distant metastasis) may supersede a benign FNA diagnosis.[24, 25, 26] However, there is no consensus as to what constitutes a malignant growth pattern, and a repeat FNA biopsy is often performed when an increase in nodular size is identified.[24]

The indeterminate category is challenging for both pathologists and clinicians.[10, 44] This category includes a broad spectrum of diagnostic possibilities due to the fact that there is significant overlap in the cytologic features defining each diagnosis, thus preventing pathologists from precisely identifying an individual pathological process as benign or malignant.[44, 49] The major entities included in the differential diagnosis of an indeterminate nodule include hyperplastic/adenomatoid nodule, FN (which encompasses follicular adenoma and FC), Hurthle cell neoplasm (Hurthle cell adenoma and HC), and FVPC.[32, 42] For the pathologist, overlapping cytologic features of benign and malignant conditions complicate diagnostic certainty. A major limitation for clinicians lies in the

wide range of terminology used by the pathologist.[44] With no standardized meaning, terms such as "atypical, indeterminate, and suspicious" provide the clinician with suboptimal information about the malignant potential of a nodule, and often the same term can initiate a wide variety of treatment plans.[44] "Favor benign colloid nodule," FL, FN, "cannot rule out FN" and "favor FN" are more descriptive phrases used to describe indeterminate lesions, and differences in interpretation by the clinician may have major implications for patient treatment.[32] Certain phrases are less ambiguous and provide clinicians with more information about the risk of malignancy. A diagnosis of FN defines a differential diagnosis between follicular adenoma and FC, resulting in a lobectomy, while the phrase "cellular nodule cannot rule out FN" often leads to continued observation and repeat FNA biopsy.[32] It is widely accepted that the only way to distinguish follicular adenomas from FC is on histology, where evidence of capsular or vascular invasion is the key feature defining FC.[32] Following an indeterminate diagnosis on FNA, current literature supports surgical excision in the form of a thyroid lobectomy as the treatment of choice. As the risk of malignancy associated with an indeterminate FNA diagnosis ranges between 5% and 42% over half of the nodules with a indeterminate diagnosis will be benign. A lobectomy not only allows for histological examination of the nodule but also enables those patients who had a benign nodule to retain a functional thyroid lobe, thus avoiding dependence on thyroid hormone replacement.[45] If, following a lobectomy, a histological diagnosis of malignancy (FC, HC or FVPC) is made, the recommended course of treatment is completion thyroidectomy.[24, 25] With a widely invasive FC a completion or sub-total thyroidectomy is performed to assist with radioiodine ablation.[24] Radioiodine ablation is administered to

destroy any residual thyroid tissue and its benefits are two-fold. First, the potential for microscopic cancer foci in the thyroid tissue bed or micro-metastasis are eliminated.[30] Second, thyroglobulin and radioiodine imaging tests, used for long-term follow-up, more accurately detect tumor recurrence and metastasis when thyroglobulin and iodine uptake are not the result of normal thyroid tissue but instead are the result of thyroid cancer.[30] In the case of minimally invasive FC, where lymphovascular invasion is absent, a lobectomy is curative.[24] A lymph node dissection is not recommended for a histological diagnosis of widely invasive FC because metastatic spread tends to be hematogenous.[24] There are several groups who argue for the surgical removal of all indeterminate FNA nodules because there is a potential risk of malignancy and cytology alone cannot rule out a FC.[15, 24, 25,] The terms benign and malignant are easily interpreted by clinicians and patients readily receive appropriate treatment. However, confusion and lack of consensus surround the categorization and descriptive terminology of FNA biopsies which fall into the indeterminate category, and this makes treatment decisions more difficult.

In the past the term indeterminate was used to encompass a wide variety of cellular follicular specimens for which a benign or malignant diagnosis could not be made based on the cytologic specimen.[45] In literature, malignancy rates for these specimens range between 12-41%.[45] In recent years several multi-tiered classification systems have been proposed in an attempt to establish standardized terminology for indeterminate biopsies as well as consistent treatment protocols.[49] Many of these classification schemes have divided the indeterminate category into high and low risk subgroups based on the degree of architectural or cellular atypia.[42, 45,46, 49] The category "FL/atypia of undetermined significance" has been proposed for those specimens where

cellular or architectural findings are neither benign nor "suggestive of FN." The category FN/suspicious for FN includes both follicular and Hurthle cells lesions for which the distinction between benign and malignant can only be determined on histology. This category includes cellular adenomatoid nodules, follicular/Hurthle cell adenoma, follicular/ Hurthle cell carcinoma, and FVPC.[46]

An FNA biopsy is termed unsatisfactory (non-diagnostic) when the specimen fails to contain enough follicular epithelial cells to allow the pathologist to make a diagnosis.[45, 46] While each institution often has its own criteria for identifying unsatisfactory specimens, there continue to be discrepancies between institutions and even individual pathologists as to what constitutes an unsatisfactory FNA biopsy.[44] An acceptable number of unsatisfactory biopsies is approximately 20%.[47] Following an unsatisfactory diagnosis most nodules are re-biopsied and of these up to 50% will have a second unsatisfactory specimen. If US-guided FNA biopsies had not been previously performed, this may be useful when rebiopsying.[39] Inflammatory cells and blood often accumulate at the initial biopsy site and waiting at least 4 weeks before re-biopsying the nodule allows time for the reabsorbtion of blood and for inflammatory cells to leave the site, thus increasing the likelihood of obtaining a satisfactory specimen on repeat FNA.[24] In 10.9-12% of unsatisfactory specimens, a malignant diagnosis was identified on histology.[42, 45] For this reason, patients with two consecutive unsatisfactory FNA samples are considered potential candidates for surgery.[24, 26, 32] Clinical risk factors may also be used to categorize patients as high risk (young, male, rapidly growing cold nodule) or low risk (older, female, stable nodule).[39] Surgical excision is recommended

for "high risk" nodules while observation and annual clinical follow-up is advised for patients with "low risk" nodules.[39]

FNA is currently considered to be the most accurate and sensitive diagnostic tool available for the assessment of thyroid nodules. It is the primary screening test allowing clinicians to triage malignant nodules for surgery and benign nodules for observation and clinical follow up.[24, 32, 41] Clinical findings, not cytologic evaluations, allowed at best, only an inference of a benign or malignant diagnosis.[50, 51] This lead to the removal of many benign nodules.[11] Unlike US and radioisotope scans, which provide a broad picture of the structure and function of the nodule,[16, 25,26, 30] an FNA biopsy samples cells of the nodule and enables the pathologist to make a diagnosis based on the cells present within the nodule rather than inferring a benign or malignant diagnosis from often unreliable features of US or thyroid scans.[11] In distinguishing between benign and malignant conditions, FNA biopsies have spared thousands of patients with benign nodules from undergoing unnecessary surgery, while at the same time ensuring that those patients with malignant or potentially malignant nodules receive prompt surgical intervention.[5, 11, 39] The FNA biopsy procedure itself is quick, painless, and inexpensive when compared to the cost of surgery, has relatively few complications (most commonly hematoma and local discomfort) and delivers a diagnosis in 85% of satisfactory biopsies.[38, 48] One of the most critical components of FNA reporting is clear communication between pathologists and clinicians, to ensure that the terminology used by pathologists clearly conveys the type of cells present in the nodule and whether the condition poses a threat of malignancy.[44] With this information the clinician is able to provide the best possible treatment to the patient.[7, 44] As with any diagnostic procedure, experience at all levels,

from the clinician performing the aspiration, the pathologist assessing the sample, to the clinician providing treatment, makes a great difference in ensuring satisfactory samples, correct diagnosis, and proper treatment.[44]

As there are many categories of thyroid nodules, there is also a wide array of treatment options for patients, and the treatment choice is often based largely on FNA biopsy results.[7] The usefulness of a thyroid FNA biopsy is linked to its ability to accurately and effectively triage those with thyroid nodules for surgical or non-surgical (observation and follow up) treatment.[42, 49] One of the major drawbacks of thyroid FNA biopsies is the inconsistency with which sample results are reported and in turn how these reports are interpreted by clinicians treating the patient.[44] Currently, reporting schemes and terminology differ between institutions as well as between the various reporting pathologists.[49] Since the inception and acceptance of FNA biopsies in the 1970's many classification systems and mathematical equations have been developed to convey information between the disciplines involved in patient care.[49] Most major academic teaching centers across North America and Europe have conducted a series of retrospective studies examining and correlating cytological and histological data in order to ascertain the usefulness of FNA biopsies at their site. The positive predictive value (PPV) is a statistical value often used to evaluate the usefulness of a diagnostic test and is a measure of the proportion of individuals with a positive result that were correctly diagnosed using the diagnostic tool under evaluation.[58] The greater the PPV of a given test the more likely a person with a positive test result has the condition.[58] In the case of thyroid FNA biopsies PPVs have been used to identify those who are candidates for surgery and those who are best followed clinically by identifying which thyroid FNA

categories, in a given scheme, carry the highest risk of malignancy. Surgeons routinely use a PPV of 20% as the point at which surgical intervention is recommended.[46, 49] Those categories with a PPV of less than 20% are unlikely to have a malignant nodule and observation, repeat FNA, and clinical follow-up are the most appropriate form of treatment.[24, 46, 49]

Sensitivity and specificity values have historically been used to establish the accuracy of a wide variety of diagnostic medical procedures,[52] including thyroid FNA biopsies. Sensitivity is a statistical measure of the proportion of patients with a positive diagnostic test who were subsequently proven to have the disease/condition the test was being used to predict.[52] In the case of thyroid FNA biopsies sensitivity is the proportion of biopsies cytologically classified as positive/malignant which were proven malignant on histological examination. Specificity is a measure of the proportion of patients with a negative diagnostic test who did not have the disease/condition being tested.[52] Specificity represents the percentage of thyroid FNA biopsies in the benign category which were proven to be benign following surgery and histological examination.[49] In the diagnosis of thyroid nodules, histopathological examination is considered the "gold standard" and thyroid sensitivity and specificity values are calculated based on the final histological diagnosis. Sensitivity and specificity values vary greatly between institutions and can range from as low as 7% to as high as 100%.[44] A great deal of this variation can be attributed to the type of classification system used to categorize thyroid FNA biopsies and specifically how the "indeterminate" category treated.[42] Whether these "indeterminate" samples are considered a positive test result or a negative test result has a substantial impact on calculated sensitivity and specificity values. If the indeterminate biopsies are

considered a positive test result the overall sensitivity of the thyroid FNA diagnostic test increases while the specificity decreases.[50] The opposite is true if the indeterminate category is considered to represent a negative test result.[50]

While sensitivity and specificity values are important in determining if thyroid FNA biopsies are useful diagnostic procedures, these values may have little relevance in the assessment and treatment of individual patients. In these situations a more useful measurement is the concept of PPV. While sensitivity and specificity values can account for the overall diagnostic performance of thyroid FNA biopsies, [52] PPVs reflect the probability that an individual with a positive test has been correctly diagnosed and does indeed have the underlying condition being tested.[58]

PPVs are especially useful with respect to the treatment of indeterminate nodules. These nodules can be subdivided into various categories based on the degree of atypia present and the PPV is calculated for each subdivision.[49] Instead of surgically removing every indeterminate nodule (accounting for 5%-42% of thyroid FNA biopsies),[45] treatment plans can be tailored to each category and can range from continued observation, to repeat FNA, to surgery.

Diagnostic categories with high PPVs are at greater risk of malignancy and therefore require more aggressive treatment, often in the form of a hemithyroidectomy or total thyroidectomy, while categories with a PPV below 20% can be followed clinically.[49]

The purpose of this study is to examine the correlation between an FNA biopsy diagnosis and the histologic diagnosis following excision of the nodule in two Winnipeg regional hospitals, and how these results compare to published literature. We will also

establish malignancy rates for the diagnostic categories identified in our study and compare the Winnipeg hospitals' values to those presented in current published literature. This study will determine the percentage of unsatisfactory specimens in this population and how this value compares to the published literature, as well as determine the percentage of FNA biopsies termed "intermediate follicular lesion" and examine how this diagnosis was interpreted and treated by clinicians. It will also examine the role of terminology and its influence on treatment.

V. Literature Review

Gharib H and Goellner JR. 1993. Thyroid fine-needle aspiration biopsies: progress, practice, and pitfalls. *Endocrine practice* 9(2): 128-136

In their review of 7 published thyroid FNA studies between 1982 and 1991 Gharib and Goellner investigated the usefulness of FNA biopsies in treating patients with thyroid nodules.[50] Thyroid aspirations were divided into 4 diagnostic categories: non-diagnostic (unsatisfactory), benign, suspicious (indeterminate), and malignant. The suspicious category includes those biopsies suggesting a FN or a Hurthle cell neoplasm, as well as those biopsies suggesting a malignant process but lacking sufficient atypical cells for definitive diagnosis. The combined results from these 7 studies showed an average of 69% of cytology findings were benign, 10% were suspicious (indeterminate), 17% were unsatisfactory, and 3.5% were malignant. Gharib and Goellner also investigated the accuracy of a clinical examination, thyroid scan, and US in the identification of malignant thyroid nodules.[50] They found only 10-15% of nodules clinically identified as malignant contained a malignancy on histology. Alternatively, 17-51% of surgically excised nodules in the 7 reviewed FNA studies contained histologically malignant lesions. Hamberger *et al.* also investigated the percentage of malignant lesions identified prior to and after the introduction of FNA biopsies.[51] Of the patients identified as having a malignant nodule based on patient history, physical exam, and thyroid scanning who subsequently underwent surgery, only 14% had a histologically confirmed malignancy. Of the patients with an FNA diagnosis who underwent surgery, 29% had a histologically confirmed malignancy, a malignancy rate twice that of resected nodules diagnosed prior to the introduction of FNA biopsies.[51]

With the classification of thyroid aspirations as either benign, malignant, suspicious, or unsatisfactory the rate of false-negative cases ranged between 1.3-11.5% while the range of false-positive cases was 0-7.7%. In the case of false-negative rates, Gharib and Goellner point out that absolute false-negative rate cannot be determined unless every patient with a benign FNA diagnosis undergoes surgery.[50] However, it is generally accepted that, were all patients with a benign FNA diagnosis to undergo surgery the resulting false-negative rate would be less than 5%.[6, 50] Because false-negative cases may result in delayed surgical intervention for patients with a malignant nodule, Gharib and Goellner emphasized the importance of identifying and reducing potential factors which may contribute to false-negative cases.[50] Both Gharib and Goellner and the authors of the 7 studies they reviewed used sensitivity and specificity values as a means of assessing how accurately malignant lesions were identified on FNA biopsies.[50] The diagnostic accuracy of FNA biopsies was 95% while the overall sensitivity and specificity values were 83% and 92%. However, Gharib and Goellner point out that calculated sensitivity and specificity values are dependent on the treatment of suspicious (indeterminate) biopsies.[50] If a suspicious biopsy is considered positive, sensitivity increases while specificity decreases. On the other hand, if a suspicious biopsy is considered to be negative for malignancy, sensitivity decreases while specificity increases. As there are no standardized guidelines for the treatment of indeterminate biopsies, published sensitivity and specificity values range between 65-98% and 72-100% respectively.[50] The inability of sensitivity and specificity values to provide useful clinical information about the risk of malignancy associated with an indeterminate FNA diagnosis is one of the reasons current thyroid reporting schemes have started to adopt 5-7 tiered

classification schemes using cytologic/histologic correlations to determine the risk of malignancy associated with each diagnostic FNA category. In this way the indeterminate category can be broken down into lesions with a high (FN) or low (FL) risk of malignancy.

The inability of the indeterminate category to distinguish between benign or malignant nodules is 1 of 2 limitations of FNA biopsies cited by Gharib and Goellner.[50] Unsatisfactory aspirations are the second limitation of FNA biopsies. Cystic and vascular lesions are particularly problematic because these lesions contain a limited population of follicular cells and a definitive diagnosis cannot be made. Gharib and Goellner listed the skill level of the clinician aspirating the nodule as a factor in unsatisfactory rates, and suggested that developing and maintaining effective aspiration skills is important in increasing the likelihood of obtaining a satisfactory specimen from vascular or cystic nodules.[50] Based on their findings, Gharib and Goellner recommend clinical observation and a repeat FNA biopsy in one year for patients with a benign FNA diagnosis.[50] A repeat FNA is recommended for those patients with an unsatisfactory diagnosis. Surgical resection is recommended for recurrent cysts and large cysts (greater than 4 cm). Since the risk of malignancy in suspicious nodules is approximately 30%, Gharib and Goellner recommend surgical excision of these nodules.[50]

Wang HH. 2005. Reporting thyroid fine-needle aspirations: literature review and a proposal. *Diagnostic cytopathology* 34(1): 67-76

Accurate documentation and clear communication are essential components of interdisciplinary medicine and are essential for correlating disease progression, patient

prognosis, and treatment.[49] In the case of thyroid FNA biopsies, where a particular lesion is placed in a classification system has important clinical implications and for this reason clear communication between pathologists and clinicians is critical.[7, 44, 49] Unfortunately, in its' nearly three decades as a widely accepted diagnostic tool, the use of confusing, often cryptic terminology and classification schemes continues to hinder the clinical interpretation of thyroid nodules and their usefulness in directing patient care. This problem has only been compounded by an expansive spectrum of published thyroid classification schemes. Using the PaperChase database at the Beth Israel Deaconess Medical Centre, Wang searched for thyroid publications in which a formal classification scheme, with or without criteria, was published.[49] From the 87 compiled publications, the number of categories and diagnostic titles were recorded, along with cytohistological correlations, sensitivity, specificity and positive and negative predictive values, when available. Based on the number of categories utilized in each publication, reporting schemes were classified as two, three, four, five or six or more category schemes and further subdivided based on the use of similar classification terminology.[49] Non-diagnostic/insufficient categories, representing FNA biopsies where a diagnosis could not be made based on an insufficient amount of cellular material, were not included in the reporting schemes. Of the 87 publications, 8 were dedicated entirely to classifying either FL or PTC.[49]

At its' simplest, the two-category scheme directly addresses the essential question, surgery or no surgery, and FNA biopsies were either benign (no surgery) or malignant (surgery).[49] With only two choices, those biopsies showing any degree of atypia were classified as malignant and surgically removed, resulting in widely variable

specificity (47–87%) and PPVs (34-92%) on cytologic-histologic correlation. Sensitivity (83-86%) and negative predictive values (88-97%) showed significantly less variability, most likely due to the classification of any atypical biopsies as malignant.

Almost half of the published reporting schemes added a third category, ranging from non-tumor conditions to FN, and this category could be further subdivided into two distinct schemes.[49] One scheme utilized an abnormal category for those biopsies which were neither definitely malignant nor definitely benign, while another included a FN group, used specifically for follicular or Hurthle cell neoplasm. In both schemes the additional category represented those biopsies for which the degree of atypia or inability to assess the capsular invasion, on cytology alone, made an exact diagnosis difficult. The addition of an "atypical" group decreased the sensitivity of the malignant category, as a significant proportion of malignant lesions were now being included in the atypical group. However, the PPV of a malignant FNA diagnosis increased, as suspicious or slightly atypical lesions were no longer being placed in the malignant category.[49]

Wang, found that when a fourth category was added to the classification scheme it most often resulted in 2 "less than definitive" indeterminate categories, or 1 "less than definitive" category and a separate category for follicular/Hurthle cell neoplasms, or 1 "less than definitive" category and a separate category for inflammatory nodules.[49] The use of 2 indeterminate categories allowed a distinction between nodules with a high or low degree of suspicion, and subsequent risk of malignancy. As with the three category schemes, the use of two indeterminate categories reduced the overall sensitivity of the malignant category from 67-100% down to 7-84%, while the PPV of the malignant category ranged from 50-100% (with the majority ranging from 87-100%).[49] With two

"less than definitive" categories the range of calculated sensitivity values also increased substantially (7-84%).[49]

The addition of a the fifth category resulted in either a range of indeterminate categories different from one another in the degree of atypia and associated suspicion of malignancy, or the creation of a separate category for follicular/Hurthle cell lesions or inflammatory nodules.[49] The combination of all the indeterminate categories, in each of the schemes, resulted in a range of sensitivity values from 63-99%.[49] Wang also noted a trend between the increasing PPV of malignant and suspicious categories and an increase in the number of indeterminate categories employed in a particular scheme.[49]

The separation of PTC aspirations from follicular type aspirations was one of the notable differences in those schemes which employed 6 or more categories.[49] Various schemes also employed one or more separate categories for Hurthle cell lesions while others included a category for non-diagnostic/borderline biopsies. In one particular scheme, biopsies were first classified as either non-neoplastic or neoplastic before being placed in more specific diagnostic categories. Wang noted there was no increase in sensitivity or PPV when moving from a 5 category scheme to a 6 or more category scheme.[49]

Of the publications dealing only with the reporting of papillary carcinomas on FNA, Wang identified two different classification schemes.[49] In the first, FNA biopsies were either positive or suspicious for papillary carcinoma, while the second further stratified the classification of these FNA biopsies by adding a third, atypical/indeterminate category. The two category scheme had a combined (positive and

suspicious biopsies) sensitivity of 50% while the addition of the atypical/indeterminate category increased the overall sensitivity for the three category system to 81%. Both schemes showed significant separation in PPV values, 93-100% and 54-57% for positive and suspicious biopsies in the two category system and 97-100%, 64-84%, and 20-43% for positive, suspicious, and atypical/indeterminate biopsies in the three category system.[49] However, Wang questions the benefit of further stratification, particularly when the current standard of care recommends surgical excision of all lesions with a minimum 20% risk of malignancy and even the lower risk atypical/indeterminate category has at least a 20% risk of malignancy.[49] On the other hand, stratification may be advantageous if different PPVs are associated with different treatment recommendations. A lobectomy is more appropriate for suspicious nodules in the two category system with a PPV of 54%, while a total thyroidectomy may be appropriate for suspicious biopsies in the three category system with a PPV of 84%.[49]

In addressing the classification of FL, Wang included four publications which discussed only FL and which divided these lesions into 2-3 categories as well as the 6 or more category schemes which had 2 or more follicular categories.[49] A comparison of sensitivity and specificity was not possible and PPV values ranged from 2-91%.[49]

Following her extensive sorting and categorizing of available thyroid reporting schemes Wang compiled the important aspects of the various schemes and constructed her own reporting system.[49] Wang's categories related to the range of ideal PPVs with the intention that these categories be clinically significant for individual patients.[49] Positive biopsies consisted of those showing features consistent with PCT, MCT or other specific malignancies, and had an ideal PPV of 100%. Those specimens which were

suspicious but not diagnostic (typically due to a lack of cellularity) for PCT, MCT, or other specific malignancies were placed in the suspicious for malignancy category. This category had an ideal PPV greater than 65%.[49] Wang highlights the role of follicular lesions and the so-called indeterminate category as the major problem for cytopathologists.[49] Wang's "indeterminate for malignancy" category consists of indeterminate for PTC, microfollicular neoplasms (most consistent with FN), and Hurthle cell neoplasms.[49] In her "indeterminate for malignancy" group Wang included those specimens which were indeterminate for PTC, as well as the entire spectrum of FN/Hurthle cell neoplasms, ie. indeterminate for follicular/Hurthle cell neoplasm, suspicious for follicular/Hurthle cell neoplasm, and outright follicular/Hurthle cell neoplasms.[49] One of the major drawbacks of Wang's proposed scheme is that follicular/ Hurthle cell lesions are not divided into high and low risk categories.[49]

NCI Thyroid Fine Needle Aspiration State of the Science Conference. 2008. Review and Conclusions [internet]. Bethesda, Maryland.

Available from: http://thyroidfna.cancer.gov/pages/conclusions

In 2005 Wang stressed that the lack of standardized classification schemes and terminology, particularly with respect to indeterminate biopsies[49], is a major limitation of thyroid FNA cytology for both cytopathologists and treating clinicians.[46] This topic was a major focal point of the National Cancer Institute (NCI) Thyroid Fine-Needle Aspiration State of the Science Conference.[46] Interdisciplinary committees, comprised of cytopathologists, surgical pathologists, endocrinologists, radiologists, and surgeons collectively assembled and analyzed a wide range of information related to the diagnosis

of thyroid disease and the usefulness of FNA biopsies. Topics ranged from the identification of the thyroid nodules by palpation and the role of imaging of these nodules to detailed discussions of the cytological features, reporting terminology, and patient treatment following an FNA diagnosis.[46] Prior to the conference, an online forum which focused on the terminology and the criteria used to diagnose thyroid nodules on FNA biopsy provided those health-care professionals involved in the assessment and treatment of thyroid nodules with an opportunity to participate in a wide range of topics. This was an ideal format and opportunity to address the lack of standardized thyroid diagnostic categories and to set forth standardized diagnostic categories as well as more standardized treatment procedures for each category. The main diagnostic objective of a thyroid FNA biopsy is to provide a rational approach to triage patients, not only for clinical follow-up or surgery, but also for the most appropriate type of surgery.[46] A major advantage of the interdisciplinary approach was that it enabled the various health-care disciplines to come to a consensus and establish a more standardized approach to the reporting of FNA biopsies. Contributions to various on-line discussions and surveys were presented at the conference and used to propose a classification scheme which correlated diagnostic terminology and cytomorphologic criteria to their relative risk of malignancy and appropriate treatment options.[46] The resulting classification scheme consisted of 6 categories, benign, FL/atypia of uncertain significance, FN/suspicious for FN, suspicious for malignancy, malignant (in increasing risk of malignancy), and non-diagnostic.[46]

A nodule classified as benign carried a malignancy risk of less than 1% and included nodular goiters, chronic lymphocytic thyroiditis, and hyperplastic/adenomatoid

nodules.[46] The treatment approach suggested by the NCI for those patients with a benign FNA diagnosis included clinical follow-up and periodic radiologic examinations.[46]

FNA biopsies which were neither benign nor malignant fell into one of 3 categories of increasing malignant potential, FL/atypia of undetermined significance, FN/suspicious for FN, and suspicious for malignancy.[46] While suspicious lesions are not outright malignant they carry a higher risk of malignancy than FL and FN. The suspicious diagnosis is often the result of inadequate cellularity, the malignant features being present but not to the extent which is required for an outright malignant diagnosis.[46]

The classification and treatment of follicular/Hurthle cell patterned FNA biopsies which are neither benign nor malignant is controversial, yet many of those in attendance at the NIC conference agreed with the creation of two categories, FL/atypia of undetermined significance and FN/suspicious for FN.[46] The FL category includes those biopsies in which the degree of cellular architectural atypia increases the likelihood of malignancy but is not enough to warrant a diagnosis of FN. Those specimens which contain low cellularity, poor fixation, or excessive blood may also be placed in this category. The malignancy risk associated with these lesions ranges from 5-10% and the recommended follow-up includes repeat FNA as well as correlating FNA findings with available clinical and radiologic findings.[46] The FN category includes non-papillary follicular pattern neoplasms and Hurthle cell neoplasms and carries a 20-30% risk of malignancy.[46] Cellular features such as monotonous cell populations, microfollicles, nuclear overlapping/crowding and less watery colloid may be helpful in identifying FN. Because this category applies to follicular and Hurthle cell lesions for which the

distinction between a benign adenoma or malignant carcinoma can only be made on histology, these patients are potential candidates for surgery, most commonly a lobectomy.[46] By differentiating between FN and FL lesions, those with a FN diagnosis are treated surgically while those with a FL diagnosis are spared an invasive surgical procedure but a repeat FNA is performed and the nodule is followed for any clinical changes, such as an increase in size.[46]

The suspicious for malignancy category encompasses those indeterminate nodules with a higher risk of malignancy, 50-75%.[46] This category most often refers to those nodules which contain some of the nuclear features of PTC but lack the cellularity necessary to render a malignant diagnosis. However, this category can be further subdivided into suspicious for PTC, suspicious for MTC, and suspicious for other malignancies. These lesions are most often referred to a surgeon, especially if they are suspicious for PTC. Controversy exists as to whether a lobectomy or a total thyroidectomy is optimal.[46]

A malignant FNA diagnosis is rendered when the degree of cytological atypia and nuclear features are specific for a PTC, MTC, anaplastic carcinoma, lymphoma, or metastasis and carries an ideal malignancy risk of 100%.[46] As with suspicious for malignancy, there is controversy surrounding the extent of surgical intervention. A variety of factors must be taken into account and include the size and type of lesion as well as the clinical status of the patient. A total thyroidectomy is the most common surgical procedure, though a lobectomy may be sufficient in the case of PTC less than 1.5 cm in size. Depending on the extent of the disease a central neck dissection may also be

recommended, most commonly in the case of large bulky lesions or involvement of the recurrent laryngeal nerve. [46]

The final category, non-diagnostic, applies to those FNA biopsies of limited cellularity, an absence of follicular cells, or poorly preserved samples. [46] A risk of malignancy has not been applied to this category and a repeat FNA biopsy is recommended, at least 3 months after the initial procedure. Surgical resection may be recommended if the repeat FNA is also non-diagnostic. However, if the lesion is 1cm or less, surgery should not be considered unless there is an increase in size. Cystic lesions can also fall into this category and the low but significant risk of a cystic PTC has led to treatment controversy - either clinical follow up including a repeat FNA, or surgical resection if the second FNA is non-diagnostic. [46]

Yang J, Schnadig V, Logrono R Vasserman PG. 2007. Fine-needle aspiration of thyroid nodules: a study of 4703 patients with histological and clinical correlations. *Cancer (cancer cytopathology)* 111(5):306-315

In 1996, the Papanicolaou Society of Cytopathology developed a series of recommendations for the diagnosis and reporting of thyroid FNA biopsies in an attempt to address a long-standing limitation of FNA biopsies - the lack of standardized terminology and standardized reporting. [42, 56] As Wang (2005) outlined, this limitation is not only problematic for clinicians but it also greatly impedes and often prevents the comparison of data, particularly the effectiveness of certain reporting schemes or treatment plans between different institutions. [49] Using the terminology and reporting scheme outlined by the Papanicolaou Society of Cytopathology, Yang *et al.* compiled and

categorized the 12 and 13 years of FNA aspirations (and corresponding surgical specimens when available) from the Long Island Jewish Medical Center in New York and the University of Texas Medical Branch in Galveston. [42] Of interest is the categories utilized by Yang and colleagues are consistent with those outlined in the classification scheme proposed by the NCI, and even though this study was published prior to the NCI conference, it can be considered an NCI "field test."

FNA biopsies from the two hospitals were placed into one of six categories. Biopsies lacking fewer than six groups of cells on each of at least 2 slides, and/or contained predominately blood or lacked colloid and follicular cells were considered unsatisfactory.[42] Benign/Negative for malignancy biopsies consisted of abundant colloid, flat sheets or macrofollicles of unremarkable follicular cells, fluid containing macrophages, colloid, and small uniform follicular cells with coarse chromatin, lymphocytes, Hurthle cells or other inflammatory cells.[42]

Like the NCI classification system, Yang and colleagues have used the degree of architectural and nuclear atypia to subdivide the indeterminate category.[42] The atypical cellular lesion (ACL) category, is equivalent to the FL of uncertain significance category used by the NCI, and consists of abundant flat sheets of follicular cells with microfollicles and macrofollicles, or focal areas of enlarged nuclei with irregular nuclear membranes and pale chromatin.[42] A PTC cannot be ruled out due to the presence of architectural or nuclear atypia and this category would be expected to contain a portion of follicular/Hurthle cell adenomas, carcinomas and FVPC. Interestingly, Wang places mixed micro- and macrofollicular and macrofollicular aspirations in the "most probably benign" category.[42] However, Wang found that when PTC biopsies were separated into

malignant, suspicious, and atypical/indeterminate categories the atypical biopsies had a significant PPV, 20-40% and she therefore proposed the placement of these atypical PTC aspirations in an indeterminate for malignancy category.[49] The FN/Indeterminate category used by both the NCI and Yang *et al.* encompasses follicular/Hurthle cell adenomas and carcinomas, and FVPC.[42, 46] These biopsies typically consisted of scant or absent colloid and follicular cells or Hurthle cells with a syncytial or thick 3-dimentional pattern.[42] These follicular patterned lesions have enlarged, round, and overlapping nuclei and the typical features associated with PTC are absent.[42]

Aspirations which were not quantitatively or qualitatively sufficient to diagnose PCT but which contained atypical nuclei with nuclear grooves, intra-nuclear pseudoinclusions, and fine granular chromatin were placed in the suspicious for malignancy category.[42] In the early 1990's FNA biopsies at one of the hospitals placed follicular and Hurthle cell neoplasms with marked atypia in the suspicious for malignancy category. The placement of these aspirations in this category did not continue for the duration of the reporting period. However, their inclusion in this category impacted the overall malignancy rate.[42]

All specimens containing cytological features of PCT, MTC, anaplastic carcinoma, lymphoma, or metastatic cancers were placed in the positive for malignancy category. [42] When a discrepancy between the cytological diagnosis and the histological diagnosis (considered the gold standard) was encountered, these cases were reviewed in order to understand and correct the source of error. [42]

Of the 4703 identified thyroid FNA biopsies 64.6% were benign, 3.2% were ACL, 11.6% were FN, 2.6% were suspicious and 7.6% were malignant.[42] The remaining 10.4% were unsatisfactory biopsies. Of the 247 (9.8%) patients with a benign FNA diagnosis who underwent surgery 7.3% had a histologically confirmed cancer. The majority of patients with a suspicious FNA underwent surgery (86.1%) and 64.8% where found to be malignant on histology. While only 79.3% of those with a malignant FNA diagnosis underwent surgery 98.5% of the excised lesions were malignant on histology. ACL FNA biopsies had a surgical excision rate of 40.6% with 19.2% of those nodules harboring a malignancy. The surgical excision rate and malignancy rate were both higher in the FN/indeterminate category, 63.1% and 32.2% respectively.[42]

The ultimate objective of Yang and colleagues' retrospective study was to take the information gathered, particularly malignancy rates and effective treatment options, and develop a thyroid reporting system in which the risk of malignancy is effectively communicated as well as the establishment of a consistent approach to treatment.[42]

Wu HH, Jones JN, Osman J. 2006. Fine-needle aspiration cytology of the thyroid: ten years experience in a community teaching hospital. *Diagnostic cytopathology* 34**(2)**:93-96

Similar to Wang and Yang *et al.*, Wu *et al.* felt the ambiguous terminology used to report indeterminate thyroid FNA biopsies which are neither benign nor malignant significantly limited the usefulness of these biopsies in triaging patients with thyroid nodules.[42, 44, 45] This is particularly important when 5-42% of thyroid nodules are indeterminate and the malignancy rate of these nodules ranges from 12-41%.[45] A broad

spectrum of terminology has been used to describe these lesions, including atypical, FL, FN, or suspicious, and each term can have a different cytological description or risk of malignancy depending on the reporting scheme being used. Therefore, little useful information is gained from these diagnostic categories.[45] However, as surgical intervention is currently indicated when the malignancy rate is 20%[46, 49] one of the most useful parameters which might be gained from these indeterminate FNA biopsies is their likelihood of being malignant. This would help clinicians and surgeons to better identify patients in need of surgical intervention. In an attempt to pass this valuable information on to the treating physicians, Wu *et al.* set about identifying malignancy rates and the relative cancer risk associated with indeterminate thyroid FNA biopsies.[45]

A ten year retrospective computer search was performed at the Ball Memorial Hospital in Muncie, Indiana, identifying 1621 thyroid FNA biopsies, of which 401 had an accompanying thyroidectomy report.[45] In order for a diagnosis to be rendered, FNA biopsies had to contain 8 to 10 clusters of follicular cells, with each cluster having at least 10 follicular cells. All adequate biopsies were placed in one of six diagnostic categories. Nodular goiter, colloid/adenomatoid nodules, and lymphocytic/Hashimoto thyroiditis were classified as benign non-neoplastic lesions. Those specimens with an inadequate number of benign follicular cells and abundant colloid were diagnosed as benign colloid nodules and included a notation in the report identifying limited cellularity and the need for clinical correlation.[45] A malignant diagnosis was given to those biopsies with cellular features of a specific thyroid cancer, lymphoma or sarcoma. If nuclear features and cellular patterning suggested a specific type of malignancy but specimen inadequacy precluded a specific diagnosis the biopsy was classified as suspicious for malignancy.

The presence of nuclear features consistent with PTC (nuclear enlargement, nuclear grooves, nuclear pseudo-inclusions, and prominent nucleoli) lead to an atypical diagnosis, and given the diagnostic criteria used, PTC cannot be ruled out.[45] Biopsies consisting of syncytial sheets or microfollicles with enlarged nuclei, coarse chromatin, and prominent nucleoli fell into the FN category. A diagnosis of FL was given when both benign and FN cellular features were present.[45]

Of the 159 benign FNA biopsies which were surgically resected, 11 (7%) had a thyroid malignancy identified on histology.[45] Of these 11 false negative cases, 6 of the identified malignancies were not found at the site of the FNA biopsy and were deemed incidental carcinomas. FVPC accounted for 9 of the 11 false-negative cases, including the 5 cancers which were not the result of undersampling. A retrospective review of these 5 cases identified focal areas of atypia, which included nuclear enlargement, nuclear grooves, and syncytial sheets in a background of abundant colloid. While the presence of abundant colloid or cystic components resulted in a benign FNA diagnosis, Wu *et al.* concluded the presence of syncytial sheets, nuclear grooves, nuclear enlargement, and fine chromatin can serve as indicators of FVPC in FNA biopsies.[45]

When grouped together, the suspicious, atypical, FN, and FL FNA biopsies constituted the indeterminate category and accounted for 24.3%.[45] Of the 19 surgically resected suspicious lesions a malignant tumor was confirmed on histology in 13 (69%). Sixteen (48%) of the 33 surgically resected lesions with an atypical FNA diagnosis had a corresponding malignant tumor identified on histology. Of the 88 patients with a FN FNA diagnosis who were treated surgically 29 (33%) had a corresponding malignant lesion while only 7 (14%) of the 51 surgically excised FLs were malignant on histology.

When considered together, the indeterminate group accounted for 24.3% of reported FNA biopsies and had a malignancy rate of 34%,[45] and by current surgical standards all patients with an indeterminate FNA diagnosis would be candidates for surgery.[49] However, Wu *et al.* further subdivided the indeterminate category into FL, FN, atypical, and suspicious for malignancy and determined the malignancy rates for each subdivision, respectively 14%, 33%, 48%, and 68%.[45] Not only do the subdivisions differ in terms of the severity of the disease and likelihood of a malignancy on histology but also in treatment options. With malignancy rates of 48% and 68%, patients with atypical or suspicious FNA biopsies are referred to a surgeon. On the other hand surgical excision of FL or FN is less straight-forward and Wu *et al.* emphasized the importance of correlating these FNA findings with other clinical data and patient history before deciding for or against surgery.[45]

Overall, thyroid FNA biopsies at the Ball Memorial Hospital had a sensitivity of 87%, specificity of 100%, and a false negative rate of 3%.[45]

The interchanging of descriptive terminology used to identify "indeterminate" FNA biopsies is evident when comparing Wu and colleagues' FL categories with Yang and colleagues' ACL category. In their ACL category, Yang and colleagues include those aspirations with cellular atypia which prevents inclusion in the benign category but is not sufficient for inclusion in the FN category.[42] In their FL category Wu and colleagues describe aspirations with cytologic features of both a benign nodule and possible FN.[45] Both authors are describing similar nodules but are using very different terminology. Adding further support to the notion that both authors are describing similar nodules are their PPVs, FL 14%[45] and ACL 13.5%.[42] To add to the confusion Wu

et al. has a seventh category, "atypical" which consists of those specimens with nuclear atypia such as enlarged nuclei, nuclear grooves, nuclear pseudoinclusions and prominent nucleoli.[45] Unlike Yang and colleagues' atypical cellular lesion category, which would be scheduled for continued observation, Wu and colleagues' atypical category carries a PPV of 48% and should be treated surgically with ideally a hemi-thyroidectomy. Inadvertently confusing these similar terms could have serious repercussions for the patient.[45]

VI. Materials and Methods

The Delphic Laboratory Information Services (LIS) and Anatomical Pathology (AP) programs used by Diagnostic Services of Manitoba (DSM) and Winnipeg Pathology Departments were used to generate a computerized retrospective search for all thyroid cytology reports submitted to the Departments of Pathology at the Health Sciences Center (HSC) and St. Boniface General Hospital (SBGH). A list of thyroid cytology specimens was generated beginning March 19, 2007 (the date HSC and SBGH were both using the LIS/AP programs) and ending December 31, 2007 (the month prior to the initiation of the data collection component of this practicum project). A list of cytology specimens was generated by entering the hospital site, specimen type, and the start and finish dates for specimen collection. For both hospitals the specimen type was entered as cytology non-gynecology. At HSC the code THY was used to identify all thyroid cytology specimens whereas the code FTHY was entered in order to identify fine needle thyroid cytology specimens at SBGH. Thyroid FNA cytology reports from both the HSC and SBGH were printed in the Cytology department at HSC. To identify patients with a corresponding surgical specimen an "independent" patient history search was initiated. The LIS/AP inquiry function and the patient's personal health and information number (PHIN) were used to locate possible surgical thyroid specimens or additional thyroid FNA biopsy reports. Once each cytology report had been searched all collected data was arranged in alphabetical order, by patient last name, in order to eliminate any duplicate reports.

Clinical information collected on each patient included date of birth, age at the time the FNA biopsy was taken, gender, cytology number, hospital site where the biopsy was signed out, cytology category, and a brief description of the sample. If a surgical

excision or biopsy was performed the surgical pathology number and histological diagnosis were recorded.

The initial search identified less than 100 thyroid cytology reports with a corresponding surgical report. To increase the number of cytologic/histologic correlations and the likelihood of generating statistically significant results the time frame used to search for cytology reports was widened to include HSC and SBGH cytology and histology reports between January 1, 2008 and June 30, 2008. The Dephic LIS/AP program began September 23, 2006 at the SBGH, seven months before the program began at HSC, and our second search also included these SBGH cytology and histology reports. Thyroid cytology reports and cytology/histology pairs were sorted into multiple diagnostic categories based on the description or diagnosis provided on the cytology report. For patients who underwent repeat FNA biopsies all cytology reports, within the designated time frame, were included. When a patient underwent multiple FNA biopsies and had a histological report the cytologic diagnosis which evoked the surgery was used for a cytologic/histologic correlation and statistical calculations. A total of 775 patients were included in the study, 893 FNA biopsies were reported and resulted in 197 surgical excisions or biopsies.

The diagnostic classification scheme recommended by the NCI was used to categorize Winnipeg thyroid cytology reports as benign, unsatisfactory, FL, FN, suspicious for malignancy, and positive for malignancy.[46]

The benign category included the following: colloid nodule, multinodular goiter, multinodular goiter undergoing cystic degeneration, benign adenomatous nodule, cysts,

and cysts in which the underlying nature of the cyst cannot be determined.[46] Cytology reports describing benign follicular epithelial cells or the absence of malignant cells present were placed in the benign category. A differential diagnosis which favored a colloid nodule or a hyperplastic colloid nodule over a FN was included in the benign category. However, if a benign nodule was favored over a FN but the reporting pathologist recommended further investigation or re-aspiration of the nodule the cytology report was classified as FL. Additionally, if the reporting pathologist favored a benign process but could not rule out a neoplasm the report was included in the FL category.

The "unsatisfactory" category included FNA biopsies in which limited or no follicular epithelial cells were present as well as biopsies in which the follicular cells were obscured by the presence of excess blood. However, FNA biopsies which lacked follicular epithelial cells but were consistent with cyst content (excess watery fluid and macrophages[53]) were placed in the benign group. The term "non-diagnostic" was used to describe both FNA biopsies with little or no follicular epithelial cells and biopsies in which follicular atypia was present but a distinction between a benign or neoplastic process could not be determined. "Non-diagnostic" cytology reports containing few, if any, follicular epithelial cells were placed in the unsatisfactory category while cytology reports describing atypical cells or reports suspicious of a neoplastic process were placed in the FL and FN categories respectively.

Using the NCI classification recommendations, the "indeterminate" category was subdivided into FL and FN categories.[46] The FL category included FNA biopsies in which the atypical architectural or morphologic features of follicular epithelial cell populations prevent a benign diagnosis, but were not sufficient for a neoplastic diagnosis.

As per the NCI classification scheme, the FL category consists of follicular patterned lesions with and without cytological atypia.[46] For the purpose of this study, the follicular patterned lesion without atypia group was further divided into truly indeterminate and assorted indeterminate lesions. The follicular patterned lesions with atypia group was divided into follicular-type atypia, and atypia cannot rule out PTC. Cytology reports describing a "microfollicular and macrofollicular cell pattern" or "syncytial and honeycomb follicular cell pattern" without indicating whether the microfollicular or syncytial pattern was dominant (indicative of a follicular neoplasm) were included in the truly indeterminate category. A FL diagnosis without a benign or malignant qualifier was included in the truly indeterminate category. Cytology reports which mentioned Hurthle cells, or a dominant microfollicular or syncytial pattern were placed in the assorted indeterminate category. Thyroid reports with a differential diagnosis which included PTC, or could not rule out PTC were included in the atypia cannot rule out PTC category. This category also included FNA reports which described a limited number of cells with features of PTC (including nuclear grooves, nuclear inclusions, or fine chromatin). Cytology reports describing atypical cellular features of uncertain significance were placed in the follicular-type atypia category. Aspirations describing follicular atypia in the absence of any features of PTC were also included in this category.

The suspicious for malignancy category consisted of thyroid aspirations displaying architectural or morphological features suggestive of a specific malignant process, but inadequate cellularity prevented a definitive malignancy diagnosis.[46] This group of aspirations was further divided into two subcategories - suspicious for PTC and suspicious for a malignancy other than PTC. Cytology reports with a suspicious for PTC

diagnosis or reports describing nuclear features consistent with PTC in a borderline aspiration were classified as "suspicious for PTC." Thyroid FNA biopsies describing features suspicious for a malignancy other than PTC or FVPC were placed in the suspicious for "other" malignancy category. This included aspirations suspicious for anaplastic, or MTC, as well as reports questioning lymphoma. Reports describing suspicious cell of uncertain significance were also placed in this subcategory.

The positive for malignancy category included those FNA biopsies with an outright malignant diagnosis or cytology reports describing an adequate population of malignant cells associated with a specific malignant process. Positive thyroid FNA reports were divided into "PTC" and "other malignancy" categories. Thyroid reports with an outright PTC diagnosis and reports describing a sufficient population of follicular cells displaying PTC nuclear features were included in the PTC category. Thyroid FNA biopsies describing or reporting a malignancy other than PTC, FC, or FVPC were placed in the "other malignancy" category. This category included biopsies reported as anaplastic or MTC, as well as lymphoma.

Based on the indolent nature of thyroid malignancies incidental nodules measuring less than 1.0 cm in greatest dimension, usually identified during imaging of the neck, are monitored by US rather than undergoing an FNA biopsy.[16] A review of 9 autopsy studies found that an average of 36% of autopsied thyroids contained occult/incidental microcarcinomas[4] and this provides additional support for monitoring incidental nodules clinically. Because incidental nodules are not biopsied and the purpose of this study is the assessment of Winnipeg FNA biopsies, using cytologic/histologic correlations, surgical reports identifying a microcarcinoma (less than

1.0 cm) were considered to be benign on histology. All indeterminate cytology reports (FL, and FN) were reviewed by Dr. H. R. Wightman, practicum supervisor and Laboratory Director at the Grace General Hospital. In addition, all cytologic-histologic correlates, excluding benign-benign and positive for malignancy-malignant, were reviewed by Dr. H. R. Wightman. This retrospective study was not a formal review, and therefore cytological and histological slides were not reviewed.

Cytologic-histologic correlations and malignancy rates were calculated for each thyroid cytology category. The on-line bio-statistical graphpad software was used to calculate the statistical differences between FNA categories using the Fisher's exact test - two tailed.

VII. Results

Eight hundred and ninety-three thyroid FNA reports from 755 patients were identified through the LIS data base, including 749 (84%) cytology reports from 642 females between the ages of 13 and 93, and 142 (16%) cytology reports from 112 males between the ages of 21 and 87. Two (0.2%) thyroid reports from a 66-year-old individual of undetermined gender were also identified. The 893 thyroid reports were reviewed and grouped into one of six categories: 280 (31%) benign, 40 (4.5%) malignant, 63 (7%) suspicious for malignancy, 51 (6%) FN, 196 (22%) FL, and 263 (29%) unsatisfactory. One hundred and twenty-seven (17%) patients underwent at least one repeat FNA while 197 (26%) were referred for surgery (tissue biopsy, lobectomy, or total thyroidectomy). Of the 197 patients with histological follow-up there was a single histology report in which it was difficult to determine whether or not the malignant nodule identified on histology had been sampled or if it was an incidental carcinoma. Therefore, the total number of patients with a benign thyroid nodule is 109-110 (57-58%) while 87-88 (42-43%) patients have a malignant thyroid nodule. (Table I)

Sixty-nine patients with an initial unsatisfactory FNA biopsy underwent at least 1 repeat FNA biopsy. Thirty-seven (54%) patients had a significant diagnosis which included 18 benign, 12 FL, 3 FN, 3 suspicious, and 1 malignant. Of the 32 (46%) patients with a 2nd unsatisfactory diagnosis 5 (16%) underwent a 3rd FNA biopsy, of which 1 (20%) patient had a malignant diagnosis and 4 had an unsatisfactory diagnosis. One patient had a 4th biopsy which was unsatisfactory. Of the 225 patients with an unsatisfactory diagnosis, 24 (11%) patients underwent a thyroidectomy and 3 (12.5%)

malignant and 21 (87.5%) benign nodules were identified on histology. The PPV was 12.5%.

Of the 271 patients with a benign FNA biopsy, 21 (8%) patients underwent a surgical resection. In 17 of the resected nodules the benign FNA diagnosis was confirmed on histology, while PTC was diagnosed on histology in the remaining 4 nodules. Of the 4 false-negative cases 3 of the malignant lesions arose within a cystic nodule, the nature of which had previously been identified on US. Cystic lesions carry a higher risk of malignancy than solid nodules and these nodules are generally monitored more closely than other cytologically benign lesions. For this reason the three cystic papillary carcinomas were not considered false negatives and were not included in the calculated PPV for benign FNA biopsies. The benign category yielded a PPV/false-negative rate of 5.5%.

The 40 positive for malignancy FNA biopsies were subdivided into 31 (77.5%) PTC and 9 (22.5%) other malignancy. Twenty-four (77%) patients with an FNA diagnosis of PTC underwent surgery and histology confirmed 22 (92%) malignancies (21 PTC and 1 anaplastic carcinoma) and identified 2 benign lesions. (Table II) Five (56%) of the 9 patients with non-PTC malignancy FNA diagnosis underwent surgery. All 5 patients had malignant lesions confirmed on histology: 2 anaplastic thyroid carcinomas, 1 MTC, and 2 non-Hodgkin lymphomas. Overall the malignant category had a PPV of 93% and a false-positive rate of 7%.

Of the 63 patients with a suspicious for malignancy diagnosis 59 (94%) were suspicious for PTC while the remaining 4 (6%) were suspicious for other malignancies.

Thirty-eight (64%) patients with a suspicious for PTC diagnosis were treated surgically. Histology confirmed 30 (79%) malignancies (29 PTC, and 1 FC) and 8 benign nodules and yielded a PPV of 79%. (Table III) Two (50%) patients with an aspiration diagnosed as suspicious for a malignancy were treated surgically and 2 (100%) malignant lesions (a PTC and a sarcoma) were confirmed on histology. The suspicious for malignancy category had an overall PPV of 80%.

Forty-nine patients had an FNA biopsy diagnosis of FN. Twenty-four thyroid nodules were surgically resected from these patients and 11(46%) were malignant (6PTC, 4FC, and 1HC) while 13 were benign on histology. The PPV was 46%.

One hundred and eighty-five patients had an FNA biopsy diagnosis of FL. Of the 59 patients diagnosed with a FL who underwent surgery 10 (17%) had a malignancy confirmed on histology, 7 PTC, 1 FVPC, and 2 FC. In addition, there was one histology report in which it was difficult to determine whether or not the malignant nodule identified on histology had been sampled or if it was an incidental carcinoma. For this reason the overall risk of malignancy associated with the FL category was determined to be 17-19%. The 196 FL cytology reports (Figure I) were subdivided into FL without atypia and atypical cytology(FL with atypia), where the FL without atypia accounted for 73% of FL cytology and FL with atypia accounted for 27%. The 143 FL without atypia were subdivided into true indeterminate cytology and assorted intermediate cytology. Of the 63 truly indeterminate cytology reports 15 (24%) patients were treated surgically. (Table IV) One malignant and one possible malignant lesion were identified on histology and the PPV was 7-13%. Of the 80 assorted indeterminate cytology reports 30 (38%) patients underwent surgery and 4 (13%) malignancies were identified on histology. The

PPV was 13%. The 53 cytology reports classified as FL with atypia were further subdivided into follicular-type atypia or atypia cannot rule out PTC. Of the 27 patients with a diagnosis of follicular-type atypia 7 (26%) underwent surgery. Histology identified 1PTC and 6 benign nodules. The PPV was 14%. Twenty-two patients had an atypia cannot rule out PTC diagnosis and 7 (32%) underwent surgery. Four (57%) PTC and 3 (43%) benign nodules were identified on histology and the PPV was 57%.

The malignant FNA category was not statistically different from the suspicious for malignancy category with a p value of 0.17. The FN category was statistically different from the suspicious for malignancy category and the FL category, with p values of 0.007 and 0.01, respectively. The FL category was not statistically different from the benign category with a p value of 0.27. The adjusted FL category, which did not include atypia cannot rule out PTC biopsies, was statistically different from the atypia cannot rule out PTC category with a p value of 0.012.

VIII. Discussion

Thyroid FNA biopsies have long been used as a primary diagnostic tool for triaging patients with thyroid nodules. The results of this study confirm that Winnipeg FNA biopsies are performed and diagnosed accurately and also support their continued use in the assessment and treatment of thyroid nodules. Analysis of 893 thyroid FNA biopsies, utilizing the classification scheme recommended by the NCI, revealed that 31% of the FNA biopsies were benign, 29% were unsatisfactory, 22% were FL, 6% were FN, 7% were suspicious for malignancy, and 4% were positive for a malignancy.

In the past, sensitivity and specificity were the parameters most utilized to assess the accuracy of thyroid FNA biopsies. However, the interpretation of indeterminate specimens (FL, and FN) as positive or negative greatly affects sensitivity and specificity values, and has contributed to the wide range of published values.[42, 50] The calculation of sensitivity and specificity values therefore gives little information with regard to the malignant potential of the indeterminate category, and in some studies this category has been excluded from the calculations altogether.[42] In light of the limited information sensitivity and specificity values provide, the Papanicolaou Society of Cytopathology and the American Association of Clinical Endocrinologists have recommended surgical removal of all thyroid nodules with an indeterminate diagnosis.[42, 56]

In recent years there has been a shift in FNA reporting. Specifically, an effort has been made to establish the absolute and relative malignancy rates in standardized and uniform diagnostic categories. In our study NCI diagnostic categories were used and PPVs were established. A benign diagnosis was associated with a 5.5% risk of

malignancy, and the "unsatisfactory" category with a malignancy risk of 12.5%. The FL and FN categories were associated with malignancy rates of 17-19% and 46% respectively. The "suspicious for malignancy" category was associated with an 80% risk of malignancy and the malignant category had a malignancy risk of 93%. The PPV in these categories provides clinically useful information to the clinician since according to the NCI recommendations, these categories can be associated with treatment recommendations.[46] Clinical follow-up is reasonable for benign lesions.[46] Excision is recommended for FN, suspicious for malignancy, and malignant aspirations.[46] Follow-up or repeat aspiration is recommended for FL.[46] At the present time, specific clinical recommendations are not included in thyroid cytology reports in Winnipeg.

In Winnipeg, benign thyroid FNA biopsies are performed and diagnosed accurately at the cytologic level, as evidenced by the excellent cytological/histological correlation and a PPV of 5%. A benign diagnosis was the most common diagnosis made on FNA (31%), but was closely followed by an unsatisfactory FNA diagnosis (29%). Of the 271 patients with a benign FNA diagnosis 250 (92%) were followed clinically, while the remaining 21 (8%) patients were treated surgically. Of the 21 resected nodules, 4 malignancies were identified on histology, 3 of which arose in nodules which had previously been identified on imaging as cystic lesions. Cystic fluid contains few, if any of the follicular cells necessary for a diagnosis and generally carries a greater risk of malignancy, approximately 10%, when compared to solid nodules whose risk of malignancy is approximately 5%.[4, 16] The increased risk of malignancy in cystic lesions leads many clinicians to advise surgical excision, particularly if the cyst recurs.[16] For this reason, the cystic PTC lesions in this study were not included in the overall malignancy

rate for benign lesions. Wang and the NCI cite the risk of malignancy for benign lesions as 0%[49] and less than 1%[46] respectively. However, as noted, these are ideal values. Wang restricts her benign category to non-neoplastic conditions.[49] The NCI malignancy rate appears unreasonably low. As the rate is based on cytologic-histologic correlations, clinically atypical nodules will be over-represented, raising the overall malignancy rate of the benign category. Supporting this interpretation are the malignancy rates for benign nodules in the literature. Both Yang *et al.* and Wu *et al.* calculated the risk of a histological malignancy in a cytologically benign nodule to be 7%.[42, 45]

The high accuracy of thyroid FNA in the diagnosis of malignant thyroid nodules, especially PTC, has been well-documented in the literature, and the Winnipeg data is no exception. A "positive for malignancy" FNA diagnosis accounted for 4% of Winnipeg thyroid cytology reports, and of the 29 patients who subsequently underwent surgery, 27 (93%) had a corresponding malignant diagnosis on histology (3 anaplastic carcinomas, 1 medullary carcinoma, 2 non-Hodgkin lymphoma, 3FVPC, and 18 PTC). Two benign colloid nodules were also identified on histology, resulting in a false positive rate of 7%. When considered separately, a PTC diagnosis accounted for 78% of malignant cytology specimens, and had a PPV of 93%. A malignant diagnosis, other than papillary carcinoma, accounted for 22% of the malignant cytology reports with a PPV of 100%.

As cytologic and histologic microscopic slides were not reviewed in this study, the source of error responsible for the false-negative and false-positive cases cannot be determined. In the literature, sampling error, in which the malignant nodule identified on histology was not the nodule aspirated, accounts for the majority of false-negative cases. Interpretive error is a less common source of error.[42] In their review of 1621 thyroid

FNAs, Wu *et al.* report under-diagnosis of FVPC as the most common cause of interpretive error.[45] Diagnostic difficulties were attributed to the presence of colloid and monolayer sheets of follicular cells with a limited number of PTC nuclear features. Yang *et al.* determined that the diagnosis of a benign nodule on inadequate or borderline aspirations was responsible for a majority of their false-negative cases, and cystic PTC was a common contributing factor to their false negative rate of 7.3%.[42] In our study, cystic papillary carcinomas were specifically excluded from the benign category malignancy rate calculation.

In our study, 2 patients with a malignant FNA diagnosis had histologically benign nodules (false-positive). The presence of pseudo-nuclear inclusions, nuclear grooves, fine chromatin, and psammoma bodies in hyalinizing trabecular adenomas can result in a malignant FNA diagnosis and is a common contributing factor of FNA false positives.[7] Ylagan *et al.* reported the overlapping cytologic features of FN, adenomatous nodules, and lymphocytic thyroiditis as additional factors responsible for FNA false positive cases.[55] Interestingly, the two false-positive cytology samples in this study were sent to an outside consultant with a special interest in thyroid lesions.

A malignancy was confirmed in 80% of those cases diagnosed as suspicious for malignancy on FNA, a PPV well above the NCI malignancy range (50-75%).[46] Of the 63 patients who had an FNA biopsy which was suspicious for malignancy, 59 (94%) were suspicious for papillary carcinoma, and 4(6%) were suspicious for another malignancy. Thirty-eight (65%) patients with a thyroid nodule suspicious for PTC underwent surgery. Of the 38 patients, histology confirmed 30 (79%) malignant nodules, including 1 HC, 7 FVPC, and 21 PTC, and identified 8 benign nodules. Of the four patients diagnosed with

a nodule suspicious for another malignancy, 2 (50%) underwent surgery. With the histological diagnoses of a PTC and sarcoma, the "suspicious for other carcinoma" category had a PPV of 100%. The high PPV in our data raises the possibility that a subtotal thyroidectomy may be appropriate for a "suspicious for malignancy" diagnosis, particularly, if clinical suspicion is high. This is an issue best left to clinical judgment. In our study population subtotal thyroidectomies were only rarely performed after a "suspicious" diagnosis, despite some literature recommendations supporting that approach.[49]

The "classically" indeterminate category (FL and FN) accounted for 28% of Winnipeg thyroid cytology reports, and yielded a PPV of 24-26%. In accordance with current surgical practice, all of these indeterminate lesions would require surgical intervention. However, one of the main aims of our study was the separation of the indeterminate category into FL and FN, and to determine whether these categories demonstrate clinically significant differences in their PPVs. In our study, there is excellent separation between the PPV for FL (17-19%) and FN (46%). The difference in FL and FN PPVs is statistically significant ($p=0.011$). While our values are higher than the ideal malignancy rates cited by the NCI for FL and FN cytology (5-10% and 20-30% respectively),[46] they are consistent with those in the literature. Yang *et al.* reported FL and FN PPV of 13.5% and 32.2%.[42] Wu *et al.* reported FL and FN PPV of 14% and 33% respectively.[45] Perhaps of more importance is the fact that the FL PPV is less than 20%, below the level generally considered by head and neck surgeons to require surgery, while the FN PPV (46%) supports surgical resection. Since FL biopsies account for 22% of the

total case load, sparing these patients invasive surgery would not only have a significant clinical impact but would also significantly reduce patient morbidity.

In this study the FL category, as per NCI definition, consists of follicular patterned lesions with and without cytological atypia.[46] This category therefore includes thyroid aspirations which demonstrate borderline features of papillary carcinoma. When the category of "atypia cannot rule out papillary carcinoma" is excluded from the FL category the PPV of the FL category dropped to 12-13%. The PPV for "atypia cannot rule out papillary carcinoma" is 57%, significantly different from the adjusted FL category ($p=0.012$). However, the number of "atypia cannot rule out papillary carcinoma" cytology reports with a histologic correlation is small ($n=7$). The PPV of this category is essentially the same as the PPV of the FN category and these lesions should probably be excised.

A further objective of this study was to determine how various FNA categories influenced treatment. In our study, 92% of patients with a final benign FNA diagnosis were followed clinically, while 72% of patients with a malignant diagnosis and 71% of patients with a suspicious for malignancy diagnosis were treated surgically. As June 30th 2008 was the cut-off date for both cytology and histology reports, many patients with cytology reports in April, May or June who were to be treated surgically may not have undergone surgery within the time scope of our study. Additional patients may have been treated surgically outside our catchment area. Patients may also have had additional medical factors which prevented them from being surgical candidates. Of those with an unsatisfactory diagnosis, 54% had at least 1 repeat aspiration and 13% of patients with a final unsatisfactory FNA diagnosis underwent surgery. The treatment of patients with the

FN diagnosis was also inconsistent, with 53% undergoing surgery while 47% were followed clinically. A FL diagnosis favored clinical follow-up (66%) over surgical excision (34%). The exclusion of the "atypia cannot rule out PTC" category had little effect on patient treatment: 35% of patients underwent surgery and 65% of patients were followed clinically. The inconsistent treatment of patients with a FN diagnosis and surgical resection of one-third of all FL diagnosed lesions suggests that information currently included in thyroid cytology reports does not effectively/clearly communicate the risk of malignancy to the treating clinician. The clinically significant difference between our adjusted FL and FN malignancy rates clearly supports the appropriateness of different treatment approaches for these categories. A PPV of 46% warrants the surgical resection of FN nodules and is consistent with the NCI recommendation that patients with an FN diagnosis undergo a lobectomy.[46] This may be followed by a completion thyroidectomy if a malignancy is identified on histology. The low risk of malignancy associated with the adjusted FL category suggests that immediate surgical intervention is unnecessary. Close clinical follow-up and a repeat FNA is appropriate and is consistent with the more conservative approach taken by the NCI.[46] They recommend a repeat FNA performed in 3-6 months time, and surgical resection can be considered if the second FNA biopsy is diagnosed as a FL or a lesion with a higher risk of malignancy.[46] Wu *et al.* emphasized the importance of correlating these diagnoses with available clinical data before deciding on a definitive treatment.[45] Yang *et al.* found that 73% of patients with an initial FL diagnosis who underwent a repeat FNA biopsy had a benign diagnosis.[42] They recommended a repeat FNA biopsy for these patients.[42] Conversely, nodules

diagnosed as a FN carried a risk of malignancy of 32% and Yang and colleagues recommended these nodules be surgically excised.[42]

The NCI recommends a lobectomy for patients with nodules suspicious for a malignancy.[46] There is some controversy as to whether patients with a malignant FNA biopsy should undergo a total thyroidectomy or a lobectomy. In these cases the NCI recommends correlating the type of malignancy diagnosed on FNA with available clinical data, particularly size and distribution, which may be useful in determining which surgical procedure is most appropriate.[46] Wang suggests that nodules with a substantial risk of malignancy should undergo a total thyroidectomy while a lobectomy may be warranted for nodules with a lower, but significant, risk of malignancy.[49] This is particularly true for nodules demonstrating malignant, suspicious, or atypical nuclear features of papillary carcinoma. A repeat FNA biopsy is the most commonly recommended clinical follow-up for patients with an unsatisfactory diagnosis. Surgical excision may be warranted for patients with two unsatisfactory aspirations, particularly in the case of recurrent thyroid cysts.[16, 46] Interestingly, in our study 4 patients with 3 or 4 unsatisfactory aspirations have yet to undergo surgery. When treating patients with a benign thyroid nodule there is a general consensus within the published literature.[16,42,45,49] These nodules are followed clinically and a repeat FNA biopsy or surgical excision is recommended when clinical features are consistent with a malignant process[16,25, 42,45,49]

The second objective of this study was to compare the Winnipeg unsatisfactory cytology rate to that of the published literature. In our study this category accounted for 29% of the Winnipeg thyroid fine-needle cytology, a value well above the acceptable range (10-20%) of unsatisfactory biopsies.[47] There are several factors which may have

contributed to our high unsatisfactory rate. These include the experience of the clinician aspirating the nodule and the experience of the pathologist examining the aspirate, as well as the nature of the nodule itself.[47] Cystic nodules are a common contributing factor to this category, particularly because these nodules contain at least one milliliter of fluid and very few follicular cells.[47] Excessive air-drying, hemorrhagic material, or poor fixation may also contribute to an unsatisfactory diagnosis.[47] In their retrospective review of thyroid cytology reports from 37, 895 patients, Ravetto *et al.* reported an unsatisfactory rate of 1.64%.[29] Ravetto *et al.* emphasized the importance of the pathologist's experience level, time devoted to assessing thyroid cytology, as well as number of thyroid cases assessed each week in achieving an "unsatisfactory" rate of 1.64%.[29] Of particular significance, in their study the treating clinician localized the thyroid nodule and the pathologist performed the aspiration. The cytology specimen was immediately evaluated microscopically and this provided an opportunity for any concerning features to be discussed between the pathologist and clinician.[29] In addition, if the initial aspirate was unsatisfactory a second aspiration could be obtained immediately, as the patient remained present during initial evaluation of the specimen.[29]

Currently there are no standardized requirements for the number of follicular cells needed for a satisfactory FNA diagnosis.[44] Consequently, institutional guidelines vary in the number of follicular cells required for a satisfactory diagnosis, and literature values for unsatisfactory rates range from 1.6-43%.[29, 40] In Winnipeg, the terms "unsatisfactory" and "non-diagnostic" were used interchangeably to describe aspirations with an insufficient number of follicular cells. The term non-diagnostic was also used to describe indeterminate biopsies in which the underlying pathology could not be determined. In

our study, the term "non-diagnostic " is not used and these cases were assigned to appropriate NCI categories based on reported microscopic descriptions.

One reason for the presence of an indeterminate category is that FNA biopsies are unable to reliably distinguish between a hypercellular colloid nodule, follicular adenoma, FC, and FVPC. In the past, follicular carcinomas were the second most common thyroid cancer and accounted for 5-20% of thyroid malignancies.[54] However, current research has suggested true follicular carcinomas likely account for only 1-2% of all thyroid malignancies.[54] This suspected reduction in the incidence of FC is thought to be the result of the introduction of table salt fortified with iodine, as iodine deficiency is a known factor in the development of FC.[54] In a retrospective review of 197 surgically resected malignant thyroid nodules, DeMay identified 7 (4%) histologically diagnosed FC.[54] However the number of FC was reduced to 2 (1%) following intradepartmental and extradepartmental reviews and the reclassification of 5 FC to FVPC.[54] In our study, 6 FC and 2 HC were diagnosed out of 88 malignancies (9% of total carcinomas). Of particular interest is the number of FC (including Hurthle cells) identified in our FL and FN categories. According to DeMay and other authors, most of the carcinomas in this group will be FVPC, with only a small minority of FC.[54] However, in our classic indeterminate category, 7 of 22 carcinomas were of follicular type (31%). In our FN group 5 of 11 carcinomas were of follicular type (45%). All of these FC are of micro-invasive type. Overall, 21% of all cases diagnosed as FN were histologically confirmed as FC. The significant difference in FC rates between FN (21%) and FL (3%) is an additional finding in favor of the separation of the indeterminate category.

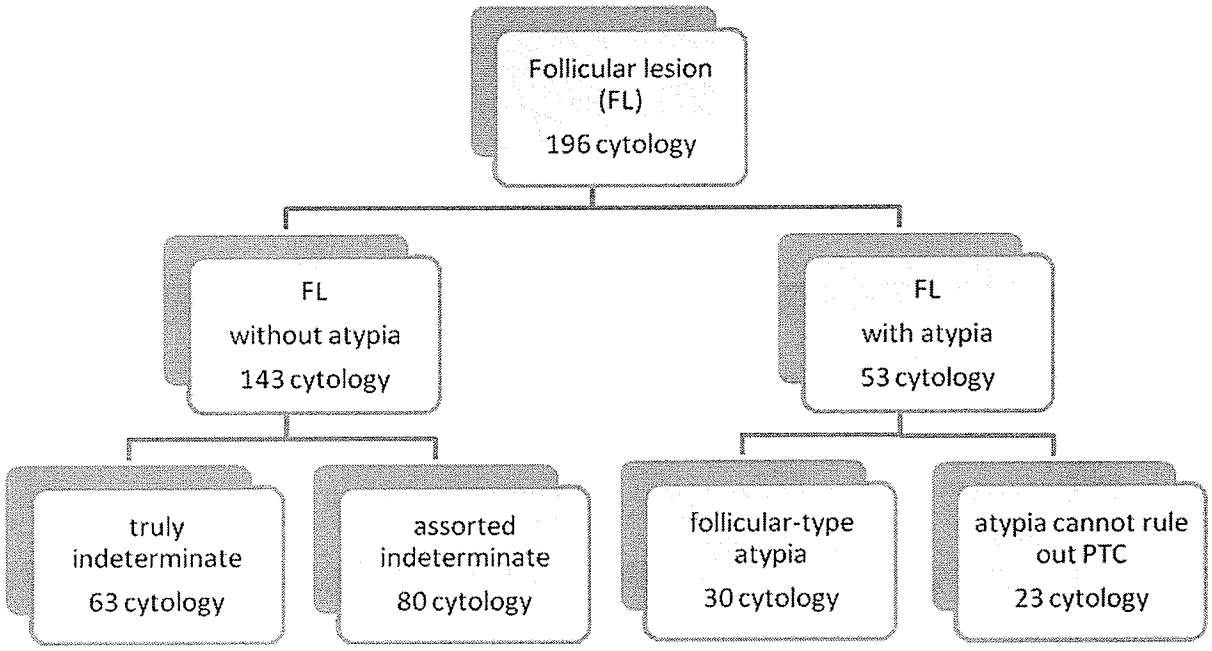
The classification of cystic FNA biopsies is problematic. The nature of cystic lesions (abundant watery fluid and a limited follicular cell population) makes precise cytologic diagnosis difficult.[42, 45, 50] Cystic thyroid nodules are associated with a malignancy risk of 10.7% compared to solid nodules which have a malignancy risk of less than 5%.[25, 26] The NCI recommends these lesions be placed in the unsatisfactory category as the underlying nature of the cyst cannot be determined in the absence of sufficient follicular cell populations.[46] Gharib and Goellner also recommend placement of cystic lesions in the unsatisfactory category, but based on the elevated risk of malignancy associated with these lesions, they suggest a repeat FNA biopsy and careful clinical follow-up of all cystic thyroid nodules.[47] Yang *et al.*, Sangalli *et al.*, and the Papanicolaou Society of Cytopathologists have classified cystic lesions in the benign category.[42, 53, 56]

The "atypia cannot rule out papillary carcinoma" category, in our study, contained very few cases (n=7). While there was a statistically significant difference (p=0.011) between the adjusted FL category and the atypia cannot rule out PTC category, a change in the diagnosis of a single case in this small population has the potential to alter the statistical significance. Thus, additional numbers of these cases would add greater justification not only to their separation from the FL category but also to their separate treatment protocol.

IX. CONCLUSION

In summary, the ultimate goal of thyroid FNA cytology is to provide a minimally invasive diagnostic procedure in which the cellular nature of thyroid lesions can be assessed and patients are accurately triaged for either surgical resection or clinical follow-up. In this regard, the separation of indeterminate lesions into FL and FN categories is advantageous, particularly when the more common FL (22%) diagnosis can be treated by repeat FNA, while FN (6%), aspirations are consulted to a surgeon. "Atypia cannot rule out papillary carcinoma" cytology has a malignancy risk of 56% and should not only be considered separately from the FL but should also be treated surgically. Based on the data presented, we recommend the use of a seven-tiered modified NCI thyroid classification scheme, in which thyroid FNA biopsies are classified as positive for malignancy, suspicious for malignancy, atypia cannot rule out PTC, FN, FL, benign, or unsatisfactory. Patients with malignant, suspicious for malignancy, atypia cannot rule out PTC, and FN aspirations (or diagnoses) are candidates for surgery. A repeat FNA biopsy and close clinical follow-up is recommended for FL and unsatisfactory aspirations. Benign nodules can be followed clinically and a repeat aspiration is recommended if there is a change in clinical features, particularly those features which suggest a malignant process.

Figure I: Classification Scheme for Winnipeg hospitals follicular lesion thyroid FNA biopsies



Follicular lesion (FL), papillary thyroid carcinoma (PTC)

Table I: Cytologic-histologic correlation of 197 patients						
		Histology diagnosis: No. of patients				
Cytology diagnosis	No. of patients	Benign	PTC/FVPC	FC	HC	Other
Benign	21	17	1 (3 PTC cysts)			
FL	59	48-49	8-9	2		
FN	24	13	6	4	1	
Suspicious for malignancy	40	8	30		1	1*
Positive for malignancy	29	2	21			6†
Unsatisfactory	24	21	3			

Papillary thyroid carcinoma (PTC), follicular variant papillary carcinoma (FVPC), follicular carcinoma (FC), and Hurthle cell carcinoma (HC), follicular lesion (FL), follicular neoplasm (FN)

* one sarcoma

† three anaplastic carcinomas, 2 non-Hodgkin lymphomas, and 1 medullary carcinoma

Table II: Cytologic-histologic correlation in the positive for malignancy categories						
		Histologic diagnosis No. patients				
Cytology diagnosis	No. of patients	Benign	PTC/FVPC	FC	HC	Other
Positive for malignancy	29	2	21			6*
Positive for PTC	24	2	21			1†
Positive for other	2					5‡

Papillary thyroid carcinoma (PTC), follicular variant papillary carcinoma (FVPC), follicular carcinoma (FC), and Hurthle cell carcinoma (HC)

*Three anaplastic carcinomas, 2 non-Hodgkin lymphomas and 1 medullary carcinoma

†One anaplastic carcinoma

‡ Two anaplastic carcinomas, 2 non-Hodgkin lymphomas and 1 medullary carcinoma

Table III: Cytologic-histologic correlation in the suspicious for malignancy categories						
		Histologic diagnosis No. patients				
Cytology diagnosis	No. of patients	Benign	PTC/FVPC	FC	HC	Other
Suspicious for malignancy	40	8	30		1	1*
Suspicious for PTC	38	8	29		1	
Suspicious for other	2		1			1*

Papillary thyroid carcinoma (PTC), follicular variant papillary carcinoma (FVPC), follicular carcinoma (FC), and Hurthle cell carcinoma (HC)

* One sarcoma

Table IV: Cytologic-histologic correlation in the follicular lesion category			
		Histological Diagnosis	
Cytology diagnosis	Number of Patients	Benign	Malignant
Follicular Lesion	59	48-49	10-11
FL without atypia	45	39-40	5-6
Truly indeterminate	15	13-14	1-2
Assorted indeterminate	30	26	4
FL with atypia	14	9	5
Follicular- type atypia	7	6	1
Atypia cannot rule out	7	3	4
PTC			

Follicular lesion (FL), papillary thyroid carcinoma (PTC)

X. References

1. Sherwood L. 2001. Human physiology: from cell to system. 4th ed. California: Brooks/Cole. p 669

2. Kumrow D, Dahlen K. 2002. Thyroidectomy: understanding the potential for complications. *MEDSURG Nursing* 11(5): 228-235.

3. DeMay RM. 1996. Thyroid. In: The art & science of cytopathology. Chicago, IL: ASCP Press. Chapter 17. p 703-715

4. Wang C, Crapo LM. 1997. The epidemiology of thyroid disease and implications for screening. *Endocrinol Metab Clin North Am* 26(1):189-218

5. De Vos Tot Nederveen Cappel RJ, Bouvy ND, Bonjer HJ, Van Muiswinkel JM, and Chadha S. 2001. Fine needle aspiration cytology of thyroid nodules: how accurate is it and what are the causes of discrepant cases? *Cytopathology* 12: 399-405

6. Bakhos R, Selvaggi SM, DeJong S, Gordon DL, Pitale SU, Herrmann M, and Wojcik EM. 2000. Fine needle aspiration of the thyroid: Rate and causes of cytohistopathologic discordance. *Diagn Cytopathol* 23(4): 233-237

7. Sclabas GM, Staerkel GA, Shapiro SE, Shapiro MS, Fornage BD, Sherman SI, Vassillopoulou-Sellin R, Lee JE, and Evans DB. 2003. Fine-needle aspiration of the thyroid and correlation with histopathology in a contemporary series of 240 patients. *The American Journal of Surgery*. 186: 702-710.

8. Bahar G, Braslavsky D, Shpitzer T, Feinmesser R, Avidan S, Popovtzer A, and Karl S. 2003. The Cytological and clinical value of the thyroid "follicular lesion." *American Journal of Otolaryngology*. 24(4): 217-220.

9. Mazeh H, Beglaibter N, Prus D, Ariel I, and Freund HR. 2007. Cytohistologic correlation of thyroid nodules. *The American Journal of Surgery*. 194: 161-163.

10. Yoder BJ, Redman R, Massoll NA. 2006. Validation of a five-tier cytodiagnostic system for thyroid fine needle aspiration biopsies using cytohistologic correlation. *Thyroid*. 16(8): 781-786.

11. Cramer H. 2000. Fine-needle aspiration cytology of the thyroid: an appraisal. *Cancer (cancer cytopathology)* 90(6): 325-329

12. DeLellis RA, Williams ED. 2004. Thyroid and parathyroid tumors: Introduction. In: DeLellis RA, Lloyd RV, Hertz PU, Eng C. World Health Organization classification of tumors pathology and genetics tumors of endocrine organs. Lyon France: IARC press. p 51-56

13. Canadian Cancer Society, National Cancer Institute of Canada. 2008. Canadian Cancer Statistics. Toronto: [publisher unknown]p 4, 12,14,15, 53, 55

14. Maitra A, Abbas AK. 2004. The endocrine system. In Kumar V, Abbas A, Fausto N. Robbins pathology basis of disease. 7th ed. Toronto: W.B. Saunders. p1166-1183

15. Gharib H and Goellner JR. 1993. Thyroid fine-needle aspiration biopsies: progress, practice, and pitfalls. *Endocrine practice* 9(2): 128-136

16. Lansford CD, Teknos TN. 2006 Evaluation of the thyroid nodule. *Cancer Control* 13(2): 89-98

17. Corrias A, Einaudi S, Chiorboli E, Weber G, Crino' A, Andreo M, Cesaretti G, De Sanctis L, Messina MF, Segni M, Cicchetti M, Vigone M, Pasquino AM, Spera S, De Luca F, Mussa GC, and Bona G. 2001. Accuracy of fine needle aspiration biopsies of the thyroid nodules in detecting malignancy in childhood: Comparison with conventional clinical, laboratory, and imaging approaches. *J Clin Endocrinol Metab* 86(10):4644-4648

18. World Health organization. 2006. Health effects of the Chernobyl accident: an overview. fact sheet No 303. [internet] Geneva: Retrieved March 2009 from: http://www.who.int/mediacentre/factsheets/fs303/en/index.html

19. Ron, E. 2007. Thyroid cancer incidence among people living in areas contaminated by radiation from the Chernobyl accident. *Health physics society* 93(5): 502-511

20. Parker SL, Tong T, Bolden S, Wingo PA. 1996. Cancer statistic, 1996. *CA Cancer J Clin* 65: 5-27

21. Vander JB, Gaston EA, Dawber TR. 1968. The significance of nontoxic thyroid nodules: final report of a 15-year study of the incidence of thyroid malignancy. *Annal of internal medicine* 69(3) 537-540

22. Uppal S, Mistry D, Coatesworth AP. 2007. Cowden disease: A review. *International Journal of Clinical Practice* 61(4): 645-652

23. Fotiadis C, Tsekouras DK, Antonakis P, Sfiniadakis J, Genetzakis M, Zografos GC. 2005. Gardner's syndrome: A case report and review of the literature. *World Journal of Gastroenterology* 11(34): 5408-5411

24. Yeung MJ, Serpell JW. 2008. Management of the solitary thyroid nodule. *The oncologist* 13:105-112

25. Gharib H, Papini E. 2007. Thyroid nodules: clinical importance, assessment, and treatment. *Endocrinol Metab Clin N Am* 36:707-735

26. Hegedus L. 2004. The thyroid nodule. *New England Journal of Medicine* 351(17): 1764-1771

27. Norman J. 2009, June 22. Thyroid gland function: common tests to evaluate. Retrieved July 7, 2009 from http://www.endocrineweb.com/test.html

28. Drucker. 2005. Hot nodule. Retrieved July 7, 2009 from http://mythyroid.com/hotnodule.html

29. Ravetto C, Colombo L, Dottorini ME. 2000. Usefulness of fine-needle aspiration in the diagnosis of thyroid carcinoma: a retrospective study in 37,895 patients. *Cancer (cancer cytopathology)* 90(6):357-363

30. Griggs W, Divgi C. 2008. Radioiodine imaging and treatment in thyroid disorders. *Neuroimag Clin N Am* 18: 505-515

31. American association of clinical endocrinologists and associazione medici endocrinology. 2006. American association of clinical endocrinologists and associazione medici endocrinology medical guidelines for clinical practice for the diagnosis and management of thyroid nodules. *Endocrine Practice* 12(1): 63-102

32. Elsheikh TM. Follicular lesions of the thyroid: Classification and criteria. papanicolaou society of cytopathology. Retrieved November 2008 from: http://papsociety.org/guidelines/diagnostic%20terminology%201.doc

33. Fewins J, Simpson CB, Miller FR. 2003. Complications of thyroid and parathyroid surgery. *Oncology Clin N Am* 36: 189-206.

34. Stalheim-Smith A, Fitch GK. 1993. Chapter 10: Introduction to the nervous system and chapter 15: Peripheral endocrine glands and. In: Understanding human anatomy and physiology. St. Paul: West Publishing Company p 332-78, 556-79

35. Asari R, Passler C, Kaczirek K, Scheuba C, and Niederle B. 2008. Hypoparathyroidism after total thyroidectomy: a prospective study. *Arch Surg.* 143(2): 132-137.

36. Prinz RA, Rossi HL, and Kim AW. 2002. Difficult problems in thyroid surgery. *Current Problems in Surgery* 39(1): 1-10

37. Moore KL, Agur AMR. 2002. Deep structures of the neck. In Essential clinical anatomy. 2nd ed. Lippincott Williams & Wilkins. p 612

38. Mathur R, Shiel WC. Fine needle aspiration biopsies of the thyroid. Retrieved July 7, 2009 from http://www.medicinenet.com/fine-needle_aspiration_biopsy_of_the_thyroid/article.html

39. Castro MR, Gharib H. 2003. Thyroid fine-needle aspiration biopsy: progress, practice, and pitfalls. *Endocrine practice* 9(2): 128-136

40. Tabaqchali MA, Hanson JM, Johnson SJ, et al. 2000. Thyroid aspiration cytology in Newcastle: a six year cytology/histology correlation study. *Ann R Coll Surg Engl* 82: 149-155

41. Giard RWM, Hermans J. 2000. Use and accuracy of fine-needle aspiration cytology in istologically proven thyroid carcinoma: an audit using a national pathology database. *Cancer (cancer cytopathology)* 90(6): 330-334

42. Yang J, Schnadig V, Logrono R, and Wasserman PG. 2007. Fine-needle aspiration of thyroid nodules: a study of 4703 patients with histologic and clinical correlations. *Cancer (cancer cytopathology)* 111(5): 306-315.

43. Hay ID, Crant CS, Bergstralh EJ, Thompson GB, van Heerden JA, and Goellner JR. 1998. Unilateral total lobectomy: is it sufficient surgical treatment for patients with AMES low-risk papillary thyroid carcinoma? *Surgery* 124(6):958-966

44. Redman R, Yoder BJ, Massoll NA. 2006. Perceptions of diagnostic terminology and cytolpathologic reporting of fine needle aspiration biopsies of thyroid nodules: a survey of clinicians and pathologists. *Thyroid* 16(10): 1003-1008

45. Wu HH, Jones JN, Osman J. (2006). Fine-needle aspiration cytology of the thyroid: ten years experience in a community teaching hospital. *Diagnostic cytopathology* 34(2):93-96

46. NCI Thyroid Fine Needle Aspiration State of the Science Conference. 2008. Review and Conclusions [internet]. Bethesda, Maryland. Retrieved March 2009 from http://thyroidfna.cancer.gov/pages/conclusions

47. Gharib H and Goellner JR. 1995. Fine-needle aspiration biopsy of thyroid nodules. *Endocrine Practice* 1(6): 410-417

48. Erik KA. 2008. Approach to the patient with a cytologically indeterminate thyroid nodule. *J Clin Endocrinol Metab* 93(11): 4175-4182

49. Wang HH. 2005. Reporting thyroid fine-needle aspirations: literature review and a proposal. *Diagnostic cytopathology* 34(1): 67-76

50. Gharib H and Goeliner JR. 1993. Fine-needle aspiration biopsy of the thyroid: an appraisal. *Annals of Internal Medicine* 118:282-289

51. Hamberger B, Gharib H, Melton LJ, Goellner JR, Zinsmeister AR. 1982. Fine-needle aspiration biopsy of thyroid nodules: impact on thyroid practice and cost of care. *The American Journal of Medicine* 73: 381-384

52. Sensitivity and specificity. (2009, June 4). In *Wikipedia, the free encyclopedia.* Retrieved March 21, 2009, http://en.wikipedia.org/wiki/Sensitivity_and_specificity

53. Sangalli G, Serio G, Zampatti C, Bellotti M, Lomuscio G. (2006). Fine-needle aspiration cytology of the thyroid: A comparison of 5469 cytological and final histological diagnoses. *Cytopathology* 17:245-250

54. DeMay RM. 2000. Follicular lesions of the thyroid: W(h)ither follicular carcinoma? *American Journal of Clinical Pathology* 114: 681-683

55. Ylagan LR, Farkas T, Dehner LP. (2004). Fine-needle aspiration of the thyroid: A cytohistologic correlation and study of discrepant cases. *Thyroid* 14(1): 35-41

56. The Papanicolaou Society of Cytopathologists task force on standards of practice. 1996. Guidelines of the papanicolaou society of cytopathology for the examination of fine-needle aspiration specimens from the thyroid. *Diagnostic cytopathology* 15(1): 84-89

57. Burgess JR, Tucker P. 2006. Incidence trends for papillary thyroid carcinoma and their correlation with thyroid surgery and thyroid fine-needle aspiration cytology. *Thyroid* 16(1) 47-53

58. Positive predictive value. (2009 9, June). In *Wikipedia, the free encyclopedia.* Retrieved March 21, 2009, from http://en.wikipedia.org/wiki/Positive_predictive_value

59. RET proto-oncogen. (2009, May 17). In *Wikipedia, the free encyclopedia*. Retrieved July 7, 2009, from http://en.wikipedia.org/wiki/RET_proto-oncogene

60. Iodine-131. (2009 13, June). In *Wikipedia, the free encyclopedia.* Retrieved August 18, 2009, from http://en.wikipedia.org/wiki/Iodine-131

61. Cooper DS, Doherty GM, Haugen BR, Kloos RT, Lee SL, Mandel SJ, Mazzaferri EL, McIver B, Sherman SI, and Tuttle RM. 2006. Management guidelines for patients with thyroid nodules and differentiated thyroid cancer: the American Thyroid Association Guidelines Taskforce. *Thyroid* 16(2) 117, 118