ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

DOTTORATO DI RICERCA IN

SCIENZE BIOMEDICHE E NEUROMOTORIE

Ciclo XXXVI

Settore Concorsuale: 06/A4 - ANATOMIA PATOLOGICA

Settore Scientifico Disciplinare: MED/08 - ANATOMIA PATOLOGICA

DEEP LEARNING AND SPATIAL TRANSCRIPTOMICS TO INVESTIGATE THYROID TUMORS

Presentata da:
Dott.ssa Thais Maloberti

Coordinatore Dottorato
Prof.ssa Matilde Yung Follo

Supervisore
Prof.re Giovanni Tallini

Esame finale anno 2024

Al mio eroe

Table of contents

Deep Learning applied to thyroid cytology

Spatial transcriptomic applied to thyroid carcinoma

Abstract

There are many diseases that affect the thyroid gland, and among them are carcinoma. Thyroid cancer is the most common endocrine neoplasm and the second most frequent cancer in the 0-49 age group.
This thesis deals with two studies I conducted during my PhD. The first concerns the development of a Deep Learning model to be able to assist the pathologist in screening of thyroid cytology smears. This tool created in collaboration with Prof. Diciotti, affiliated with the DEI-UNIBO "Guglielmo Marconi" Department of Electrical Energy and Information Engineering, has an important clinical implication in that it allows patients to be stratified between those who should undergo surgery and those who should not.
The second concerns the application of spatial transcriptomics on well-differentiated thyroid carcinomas to better understand their invasion mechanisms and thus to better comprehend which genes may be involved in the proliferation of these tumors.
This project specifically was made possible through a fruitful collaboration with the Gustave Roussy Institute in Paris.
Studying thyroid carcinoma deeply is essential to improve patient care, increase survival rates, and enhance the overall understanding of this prevalent cancer. It can lead to more effective prevention, early detection, and treatment strategies that benefit both patients and the healthcare system.

Cap 1 - Thyroid

The thyroid (from ancient Greek θυρεοειδής, *thyreoeidès*, "shaped like an oblong shield") is a butterfly-shaped endocrine gland (figure 1)[2].
It was first named after Thomas Wharton, a London physician, in 1656 because of its very shape.
The thyroid is the largest endocrine viscera and it's responsible for producing hormones essential for regulating metabolism and numerous bodily functions. Thyroid hormones, primarily thyroxine (T4) and triiodothyronine (T3), influence energy production, body temperature, heart rate, and more.
Proper functioning of the thyroid gland is critical for maintaining overall health, and the thyroid is a critical component of the endocrine system.
There are many diseases that affect the thyroid gland, and among them are carcinoma. Thyroid cancer is the most common endocrine neoplasm and the second most frequent cancer in the 0-49 age group [3].

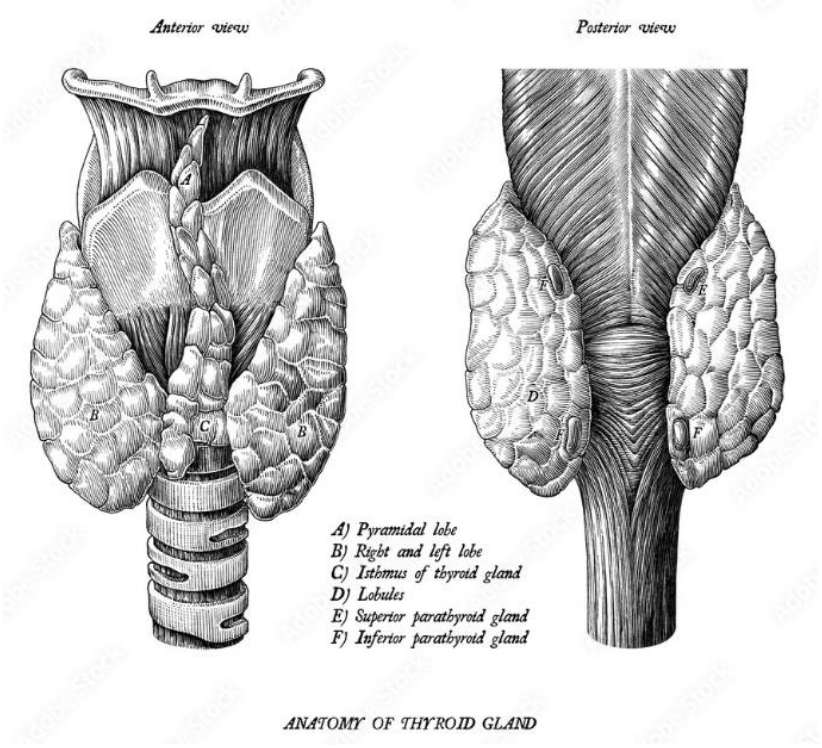


Figure 1 Anatomy of thyroid gland. Available at: https://www.istockphoto.com/it/vettoriale/tiroide-ormoni-gm487849500-73265933 (Accessed on 20 July 2023)

1. Anatomy

Thyroid has a brownish-red appearance, smooth surface, and soft texture. Age, sex and environmental factors influence its weight and size: it weighs about 1 g in an infant and increases by about 1 g/year until the age of 15 years. In the adult, the weight varies between 15 g and 20 g. Anatomically, it is located in the anterior median region of the neck, posteriorly to the sternothyroid and sternohyoid muscles, and it is in a lower position than the laryngeal thyroid cartilage, corresponding to the vertebral levels C5-T1 (figure 2). The thyroid consists of two lateral lobes that are connected by an isthmus. The two lobes are conical in shape with a lower base and extend, in the adult, about 3 cm from mid-height of the thyroid cartilage to about the 5th tracheal ring; they measure about 2 cm wide and between 0.5 cm and 2 cm thick. The isthmus connects the two lobes near their base and is attached to the first two tracheal rings; it measures about 2 cm wide, 2 cm high and 2 mm to 6 mm thick. In total, the maximum distance between the lateral margins of the two lobes is around 7 cm. It is partially covered by the sternocleidohyoid, sternotirohyoid, and homohyoid muscles, belonging to the suprahyoid muscles, and the middle cervical fascia surrounding them. Because of the characteristic oblique direction of these muscles, the gland is uncovered along the midline, where only the two cervical bundles, middle and superficial, separate it from the integument. Laterally, however, the gland is also covered by the two sternocleidomastoid muscles and, more superficially, by the bundles of the platysma muscle.

The thyroid gland is enveloped by the perithyroid sheath, a fibrous sheath from the posterior face of which connective tendrils extend to form true ligaments that secure the organ to the thyroid cartilage of the larynx (suspensory or median ligament), to the cricoid cartilage of the larynx and the first tracheal rings (internal lateral ligaments), and to the fibrous sheath that contours the vascular-nervous bundle of the neck (external lateral ligaments). The perithyroid sheath circumscribes an area within which is contained the gland lined by its own connective capsule, which is separated from the sheath by a vascular interlacing that is commonly referred to as the parathyroid danger space. In this space, along the posterior side of the organ, are located the four parathyroid glands [4, 5].

In about 50% of individuals, there is an anatomic variation of the thyroid named pyramidal lobe, the Lalouette or Morgagni Pyramid. It consists of a glandular extension, varying in shape and size, that departs from the upper margin of the isthmus and heads upward reaching the hyoid bone and sometimes passing it toward the root of the tongue [4, 5]. Sometimes there is a fibro-muscular formation that connects the hyoid bone to the isthmus or pyramidal lobe called Levator glandulae thyroideae muscle[6, 7].

The thyroid gland is supplied with blood by:

- the right and left superior thyroid arteries: arise from the external carotids and reach the organ from above;
- the right and left lower thyroid arteries: arise from the thyro-cervical trunk of the subclavian artery and reach the organ from below;
- the thyroid artery ima (when present): arises from the brachiocephalic artery or common carotid artery and reaches the organ from below at the level of the isthmus.

The thyroid is drained by:

- the superior thyroid veins, tributary to the internal jugular vein;
- the inferior thyroid veins, tributary to the brachiocephalic vein [8].

The middle and lower cervical ganglia of the sympathetic chain provide sympathetic innervation to the thyroid, while the vagus nerve provides parasympathetic innervation. An important anatomical structure of the thyroid is the recurrent laryngeal nerve (RLN). Damage to the RLN is the most common and serious complication of thyroidectomy. The RLN is a branch of the vagus nerve, transiting from the bottom to the top, either side of the trachea. From the perifollicular space it originates the lymphatic network that drains into the capsule lymph vessels, which in turn are tributary to the lymph nodes of the internal jugular chain, paratracheal lymph nodes, and pretracheal lymph nodes[6].

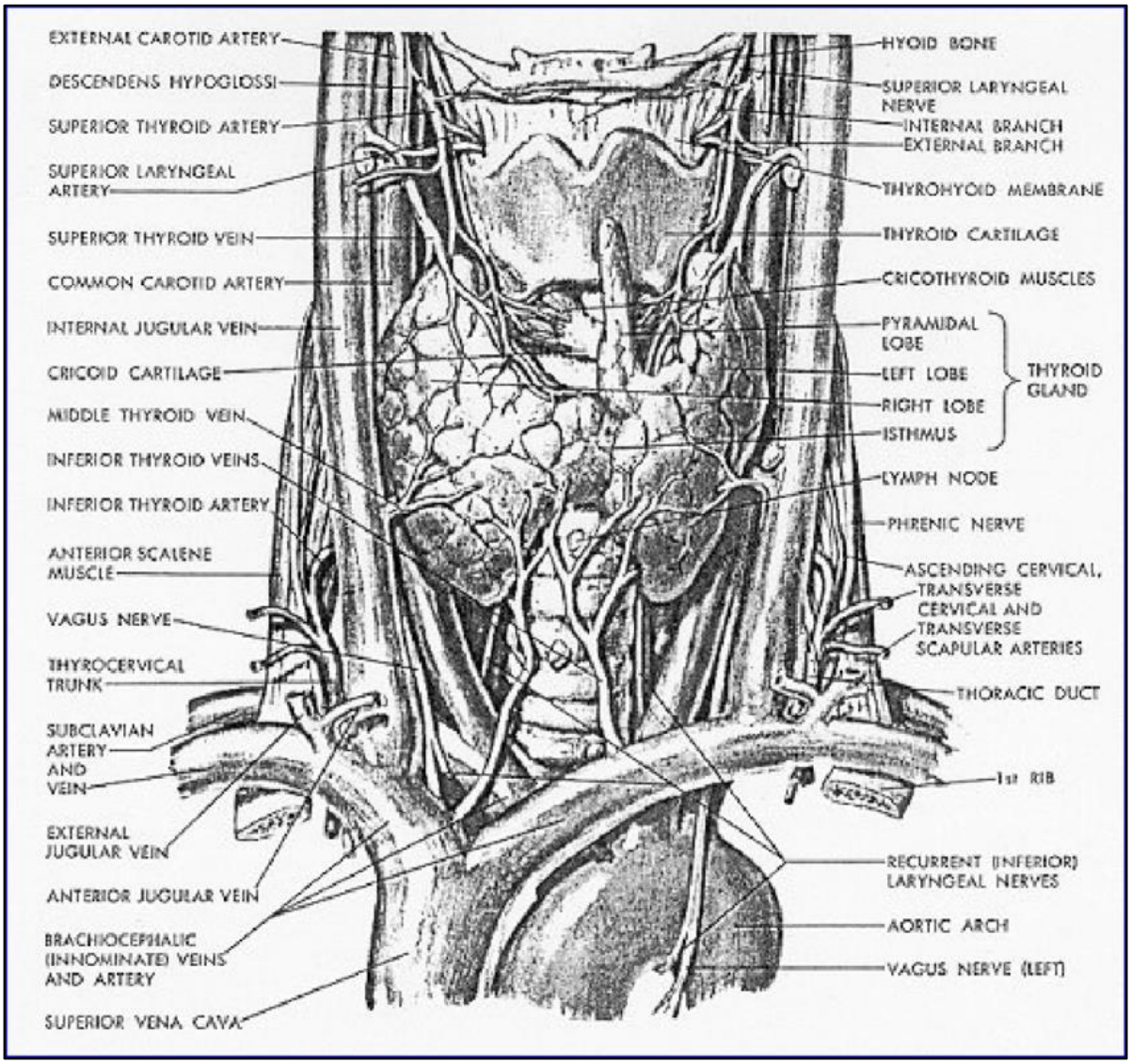


Figure 2 Anatomy of the thyroid and surroundings [Fh, N. 1965]

2.Histology

The thyroid gland is lined by a thin fibrous capsule of connective tissue and this projects within the parenchyma, forming a connective tissue framework called the stroma. This organization subdivides the gland into irregular areas called lobules. Each lobules consist of 20-40 follicles bound together by a thin sheath [9]. The thyroid follicles are the structural and functional units of thyroid gland (figure 3). They are spherical and are formed by a single layer of low cuboidal epithelium formed by follicular cells or called thyrocytes. The nucleus of the follicular cell is centrally located, round to ovoid in shape and with an inconspicuous nucleolus. The follicle is enveloped by a basal lamina and is surrounded by capillaries and lymphatics. They produce thyroglobulin, which is iodinated and stored in the colloid. Colloid is a gel-like substance contained in follicular lumen, and in adults it is common to find calcium oxalate crystals.

Thyroglobulin serves as a reservoir for the production of thyroid hormones in that, when thyroid-stimulating hormone (TSH) from the pituitary gland reaches the follicular cells of the thyroid gland, they reabsorb thyroglobulin from the colloid and release thyroid hormones into the bloodstream, where they exert their systemic effects [9, 10].

In addition to thyrocytes, the thyroid gland also has neuroendocrine-cell type called the parafollicular cells or C cells, which are intercalated between the thyrocytes or scattered in the interstitium between the follicles. C cells secrete mainly calcitonin, which through a negative feedback mechanism, regulates the concentration of calcium and phosphate in the blood. This occurs because calcitonin reduces the release of calcium and phosphate ions by osteocytes and inhibits their resorption process by osteoclasts.

The microscopic examination of thyroid tissue provides valuable insights into various thyroid disorders, from autoimmune conditions to thyroid cancer. The intricate relationship between the cellular components, the colloid, and the vascular supply ensures the proper functioning of this vital endocrine organ. Moreover, thyroid histology is instrumental in clinical diagnostics, guiding treatment strategies and ultimately improving patient outcomes in the context of thyroid diseases [9, 10].

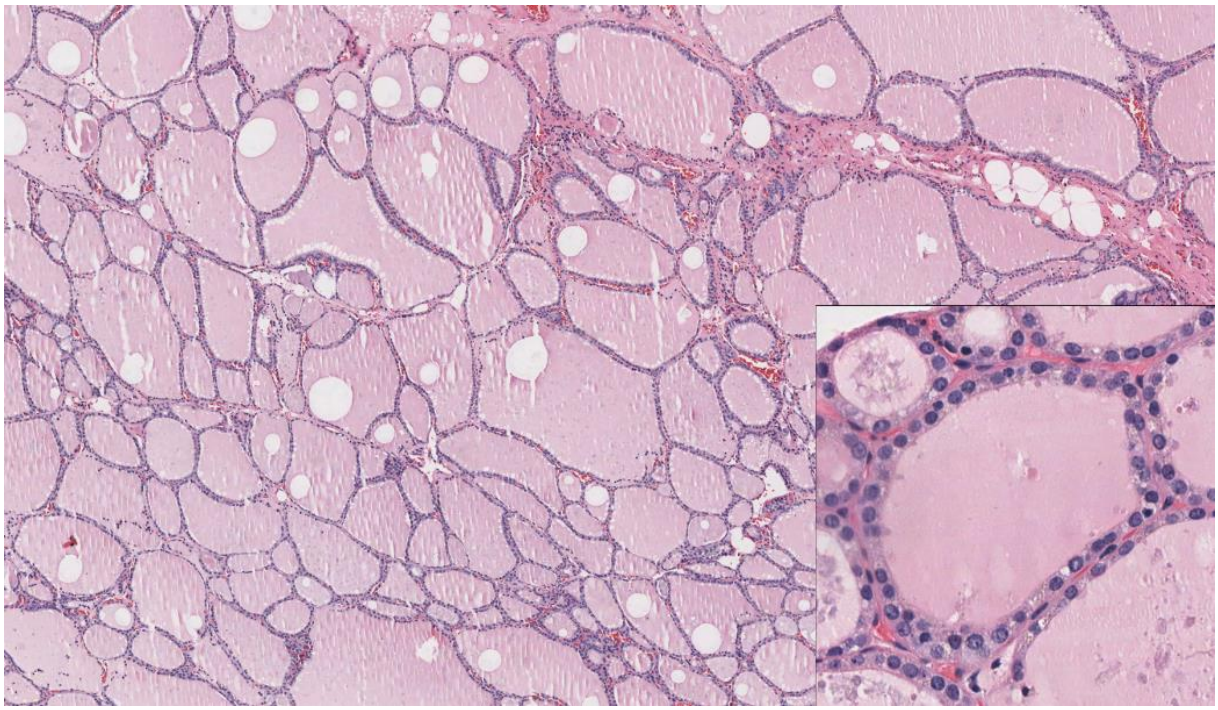


Figure 3 Histology of normal thyroid. In the ingrad, follicles with follicular cells and colloid in the center are clearly visible. Available at: https://medicine.nus.edu.sg/pathweb/normal-histology/thyroid/ (Accessed on 24 October 2023)

3. Embryology

To understand physiology and various abnormalities of the thyroid gland, knowledge of the embryological development of this gland is essential.
The thyroid gland is of endodermal origin. Around day 20-24 of gestation, endodermal cells of the primitive pharynx proliferate and create hollow embryological structure, called thyroid diverticulum or medial anlage. The thyroid diverticulum descends anteriorly to the pharyngeal sacs and migrates caudally along the midline beginning in the 5th week of gestation (figure 4).
At this point, the thyroid diverticulum is bilobed in fact, division into left and right lobes occurs, and an isthmus appears during the 6th week. It maintains an attachment to the tongue through the thyroglossal duct that usually disappears between the 6th and 8th week, but may persist and leads to abnormalities, including thyroglossal duct cysts, lingual thyroid or a pyramidal lobe.
Pharyngeal sacs, also known as ultimabranchial bodies or lateral anlagen, begin to form around the 4th/5th week. These are the lateral thyroid primordia and fuse with the medial anlage to form the complete thyroid gland. Ultimabranchial bodies contain parafollicular C-cells. After fusion and dissolution of Ultimabranchial bodies, these cells are spread through the surrounding follicular tissue, concentrating in the supero-lateral zone of the glans [10]. Parafollicular C-cells were always thought to derive from neural crest cells, but a 2016 study by Nilsson et al. challenged this thesis, suggesting that they derive from the endoderm. The ultimobranchial bodies fuse with the dorsolateral superior aspect of the developing thyroid gland and thus the tubercle of Zuckerkandl is formed [11].
The developing thyroid has continued to migrate, the solidification process also takes place at this stage, and reaches its final destination in the neck by the 7th week of gestation. From week 11th-12th, colloid is initially seen and thyroid hormone is detectable in the fetal serum. The follicular cells, derived from the medial thyroid anlage, are morphologically well developed at 13th week and distinct follicles begin to be observed throughout the thyroid gland.
The thyroid gland increases in size throughout gestation, reaching between 1 and 4 g at the end of 40 weeks of gestation.

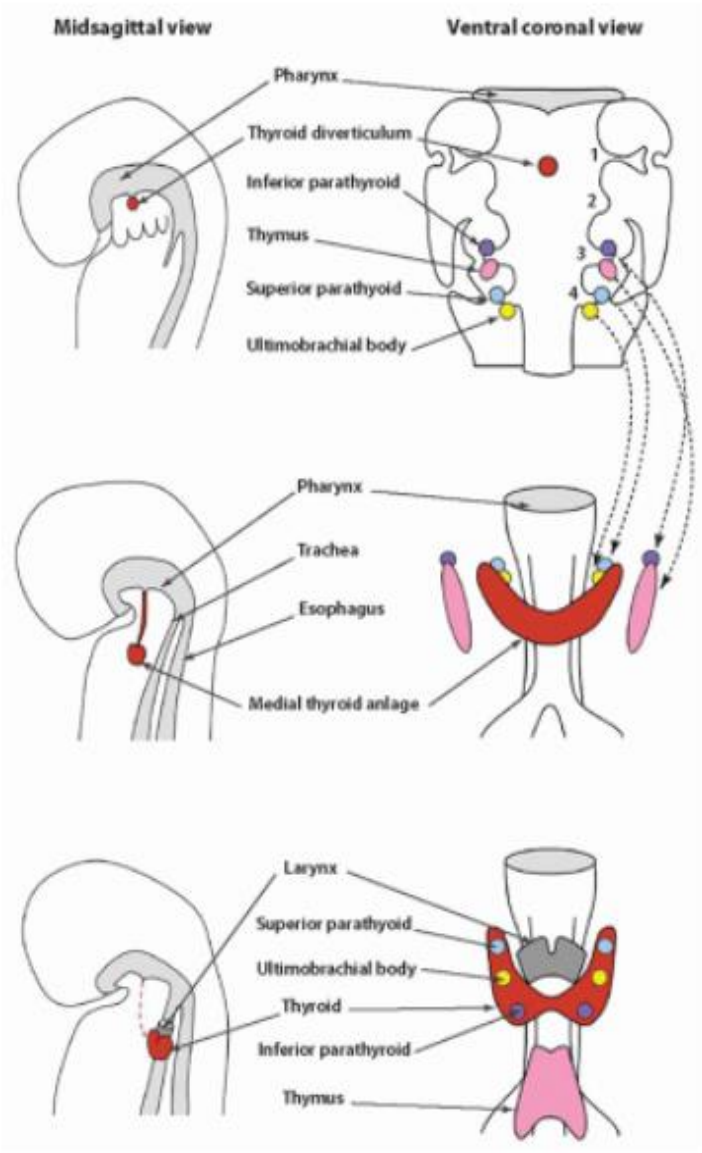


Figure 4 Embriology of Thyroid. Available on: https://basicmedicalkey.com/embryology-and-developmental-lesions]/ (accessed on 24 October 2023)

4. Physiology

The two primary hormones produced by the thyroid gland are thyroxine (T4) and triiodothyronine (T3), both of which contain iodine atoms. The process of thyroid hormone production begins with the uptake of iodine from the bloodstream into thyroid follicular cells. Inside these cells, iodine is combined with an amino acid called tyrosine to create monoiodothyrosine (MIT) or 3,5-Diiodothyronine (DIT). Combination of DIT and MIT creates the T3 hormone and combination of 2 molecules of DIT creates T4 hormone. These newly formed hormones are then stored in a substance called colloid within the thyroid follicles [12]. The release of T3 and T4 into the bloodstream is controlled by a feedback loop. When the body requires more thyroid hormones, the hypothalamus in the brain releases thyrotropin-releasing hormone (TRH), which prompts the pituitary gland to secrete thyroid-stimulating hormone (TSH) (figure 5). TSH then stimulates the thyroid gland to release T3 and T4. After the hormones are released into blood circulation, for the most part they are bind to Thyroxine-binding globulin TBG (70-75%), albumin and transthyretine [13]. This is important to regulate half-life of hormones and regulate their blood levels. Just 0,3% of T3 and 0,03% of T4 can circulate and have hormonal action [14].
Thyroid hormones are also regulated by negative feedback, because when T3 and T4 levels are high these block TRH production [15].
Thyroid hormones play a crucial role in numerous physiological processes, including [12]:

- Metabolism: T3 and T4 increase the body's basal metabolic rate, influencing energy production, heat generation, and overall metabolism,
- Growth and Development: Thyroid hormones are essential for normal growth and development, particularly in the brain and skeletal system, during infancy and childhood,
- Cardiovascular Function: Thyroid hormones affect heart rate, contractility, and blood pressure regulation,
- Temperature Regulation: They help control body temperature and response to cold stress,
- Reproductive Function: Thyroid hormones influence fertility and menstrual regularity,
- Neurological Function: Thyroid hormones are crucial for proper brain function, including cognition and mood regulation.

In summary, the thyroid gland and its hormones, are integral to regulating metabolism and various physiological functions. The intricate feedback mechanism involving the hypothalamus, pituitary gland, and thyroid gland ensures that the body receives the right amount of thyroid hormones to maintain overall health and well-being.

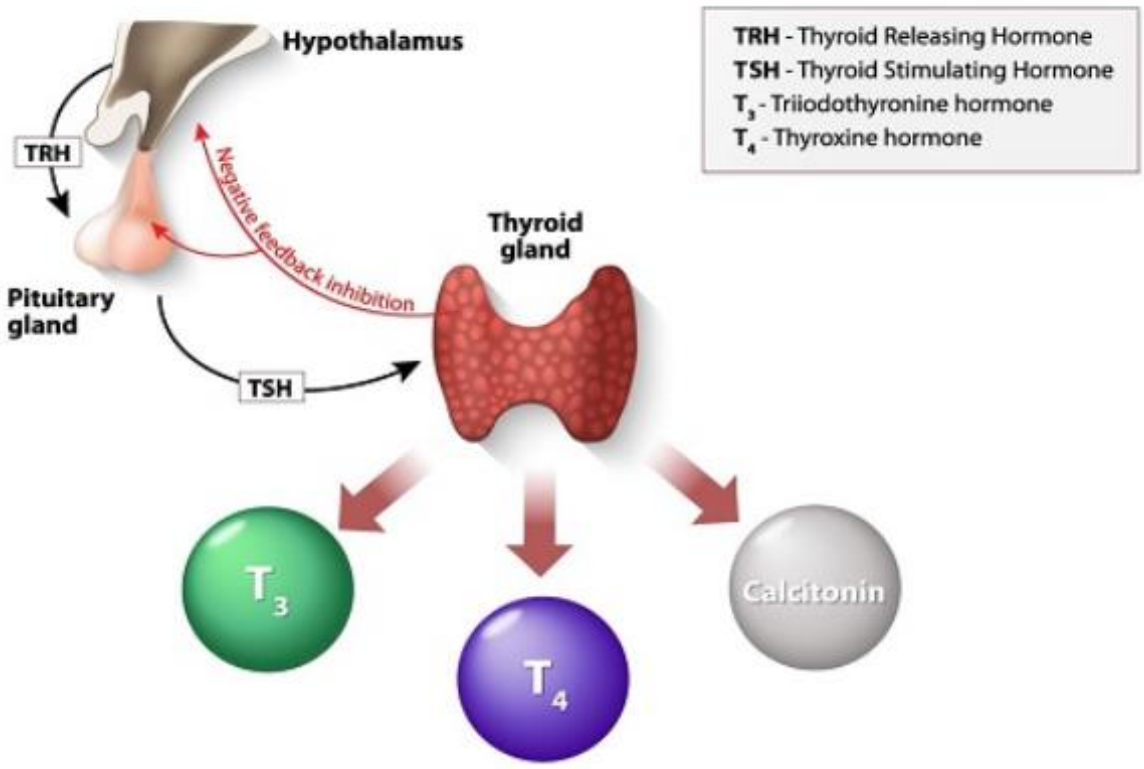


Figure 5 Thyroid hormones feedback loop. Available on: https://www.istockphoto.com/it/vettoriale/tiroide-ormoni-gm487849500-73265933 (accessed on 24 October 2023)

5. Non-cancerous pathologies

Thyroid disorders, excluding cancer which will be covered later in a dedicated chapter, mainly include hypothyroidism, hyperthyroidism and goiter. These conditions arise due to imbalances in thyroid hormone production and regulation, with each presenting distinct symptoms and underlying causes.

Hypothyroidism is a condition characterized by an underactive thyroid gland, resulting in insufficient production of thyroid hormones, mainly thyroxine (T4) and triiodothyronine (T3). This deficiency can lead to a range of symptoms, including fatigue, weight gain, cold intolerance, dry skin, and constipation. Hypothyroidism can develop due to various factors:

- Autoimmune Thyroiditis (Hashimoto's Disease): an autoimmune condition where the immune system mistakenly attacks the thyroid gland,
- Iodine Deficiency: Inadequate dietary iodine can hinder thyroid hormone synthesis,
- Thyroid Surgery or Radiation Treatment: removal of part or all of the thyroid gland or radiation therapy for thyroid cancer can result in hypothyroidism, requiring lifelong hormone replacement therapy with synthetic T4.

Treatment for hypothyroidism typically involves daily medication with synthetic thyroid hormones (levothyroxine or T4). Proper management helps restore normal thyroid hormone levels, alleviating symptoms and preventing complications [16].

Hyperthyroidism is the opposite of hypothyroidism, characterized by an overactive thyroid gland that produces excess thyroid hormones. This condition can lead to symptoms such as weight loss, rapid heart rate, anxiety, tremors, and heat intolerance. The most common cause of hyperthyroidism is Graves' disease, an autoimmune disorder that stimulates the thyroid to overproduce hormones [17]. Other causes may include thyroid nodules or inflammation (thyroiditis).

Treatment options for hyperthyroidism depend on the underlying cause and may include [17]:

- Antithyroid Medications: drugs like methimazole or propylthiouracil can reduce hormone production.
- Radioactive Iodine (RAI) Therapy: RAI destroys thyroid tissue, reducing hormone levels. This treatment may lead to hypothyroidism, necessitating thyroid hormone replacement.
- Thyroidectomy: surgical removal of part or all of the thyroid gland may be necessary, especially in cases of large goiters or when other treatments are ineffective.

Goiter is an enlargement of the thyroid gland, which can occur in both hypothyroidism and hyperthyroidism. It often presents as a visible swelling in the neck. Goiters may develop due to various reasons:

- Iodine deficiency: a common cause in regions with inadequate dietary iodine, leading to the thyroid's enlargement in an attempt to produce more hormones.
- Autoimmune conditions: conditions like Hashimoto's disease and Graves' disease can result in goiter as the immune system's activity affects the thyroid gland.
- Nodules: thyroid nodules, whether benign or cancerous, can lead to goiter formation.

Treatment for goiter depends on its cause and severity. In cases of iodine deficiency, iodine supplementation may be sufficient. For autoimmune-related goiters, treatment often involves addressing the underlying autoimmune condition. If nodules or tumors are responsible, surgical removal may be necessary [18].

6. Thyroid carcinoma

Thyroid neoplasms are distinguished according to the cells from which they originate, making a distinction between tumors that originate from non-epithelial cells and tumors that originate from the epithelial cells of the thyroid parenchyma.
Cancer originating from non-epithelial cells include: malignant lymphomas; benign mesenchymal tumors such as leiomyoma, angiomyolipoma and lymphangioma; malignant mesenchymal tumors such as angiosarcoma, leiomyosarcoma, liposarcoma and osteosarcoma.
As for neoplasms originating from epithelial cells, however, there are carcinomas.
Ninety-five percent of thyroid cancers are carcinomas, 3% are lymphomas, and 1% are sarcomas [19] and so thyroid carcinoma is the most frequent endocrine tumor. This thesis focused on the study of malignant neoplasms originating from thyroid epithelial cells. Therefore, different aspects of thyroid carcinoma will be discussed below.

6.1. Epidemiology

In 2020, the global incidence of thyroid carcinoma was approximately 586’000 people, and it is expected to increase in the next years to around 761’000 people in 2040 (figure 6) [20]. The increasing incidence of thyroid cancer has controversial explanations. Some experts argue that the increase in new cancers is due to increased diagnostic intensity, defined as over-diagnosis. Others, however, believe that the increase is attributable to changes in environment and lifestyle. The increase in thyroid cancer is most evident for small ones, but has actually occurred for all tumor sizes and stages. This suggests that the increase in diagnoses is not the only cause. If this is the case, the increase in small tumors at an early stage should be accompanied by a gradual decrease in larger and more advanced ones. Moreover, this increase of incidence has occurred almost exclusively in papillary carcinoma (PTC). Early diagnosis should have affected all histotypes. In contrast, incidence rates for follicular (FTC), anaplastic (ATC), and medullary (MTC) carcinoma have remained relatively stable over the past 30 years [21].

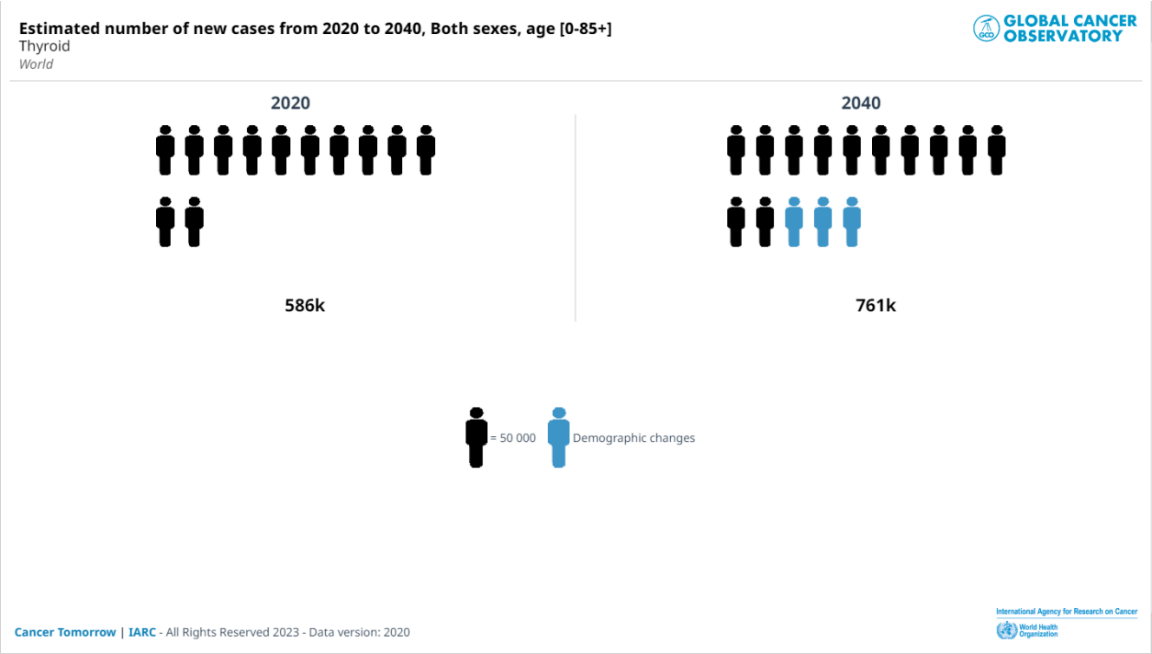


Figure 6 Estimated number of new cases from 2020 to 2040, both sexes, all ages. Available at: https://gco.iarc.fr/today (accessed on 23 May 2023)[Sung, H. 2021]

The incidence is not equal between the two sexes, in fact thyroid carcinoma is 2,9 times more common in women than men [22].
As for the incidence trend according to age, it starts to increase from adolescence to middle age, peaking at around 55 years in females and 65 years in males, and then decreasing with age [23].

Thyroid cancer often has a favorable prognosis. The mortality rate of thyroid cancers is very low and is 0,5 per 100’000 women and 0,3 per 100’000 men [24]. This carcinoma has excellent long-term survival. In US, the 5-year survival rate calculated between 2013 and 2019 is 98,5% [23].
This carcinoma is worldwide, although the incidence in some areas of the world is higher than others (figure 7). Higher-income countries, such as Austria, the Republic of Korea, Israel, Canada, Italy, France, Croatia, and the United States, have the greatest incidence, as do several middle- to upper-middle-income countries, such as Turkey, Brazil, Costa Rica, and China [25].

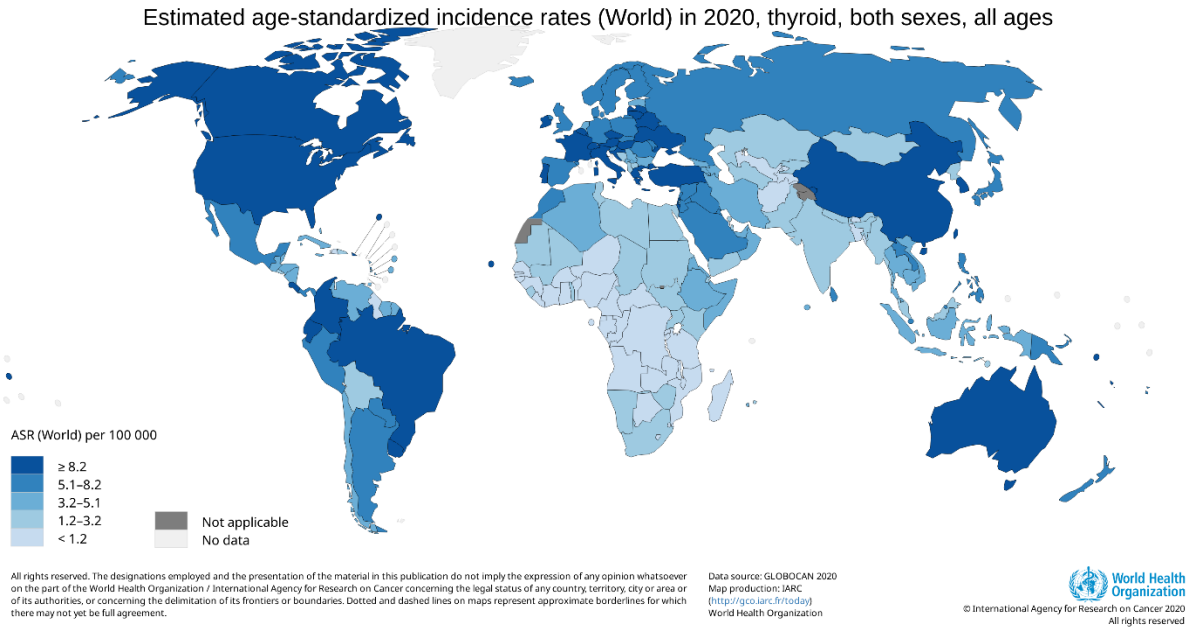


Figure 7 Estimated age-standardized incidence rates (World) in 2020, thyroid, both sexes, all ages. Available at: https://gco.iarc.fr/today (accessed on 23 May 2023)[Sung, H. 2021]

In the US, data from the Surveillance Epidemiology and End-Results Cancer Registries Program (SEER) show the largest annual percentage change (APC) in the period 1997 to 2009 with an APC of 6,7% [23]. Considering only the female population, the APC rises to 7%. Currently, one in every 55 American women and one in every 149 American males will be diagnosed with thyroid cancer over their lifetime [26].
Regarding Italy, thyroid cancer is the fifth most diagnosed cancer in women, preceded by breast, colorectal, lung and endometrial cancers [27]. Considering only women under 45, it is the second most diagnosed cancer [28].
Over the years, the incidence has followed the global trend, in fact it has increased. It should be noted that in recent years there has been a small decrease in incidence in the female population. Although the incidence has increased, however,

mortality rates have remained stable and at low levels (figure 8) [20].

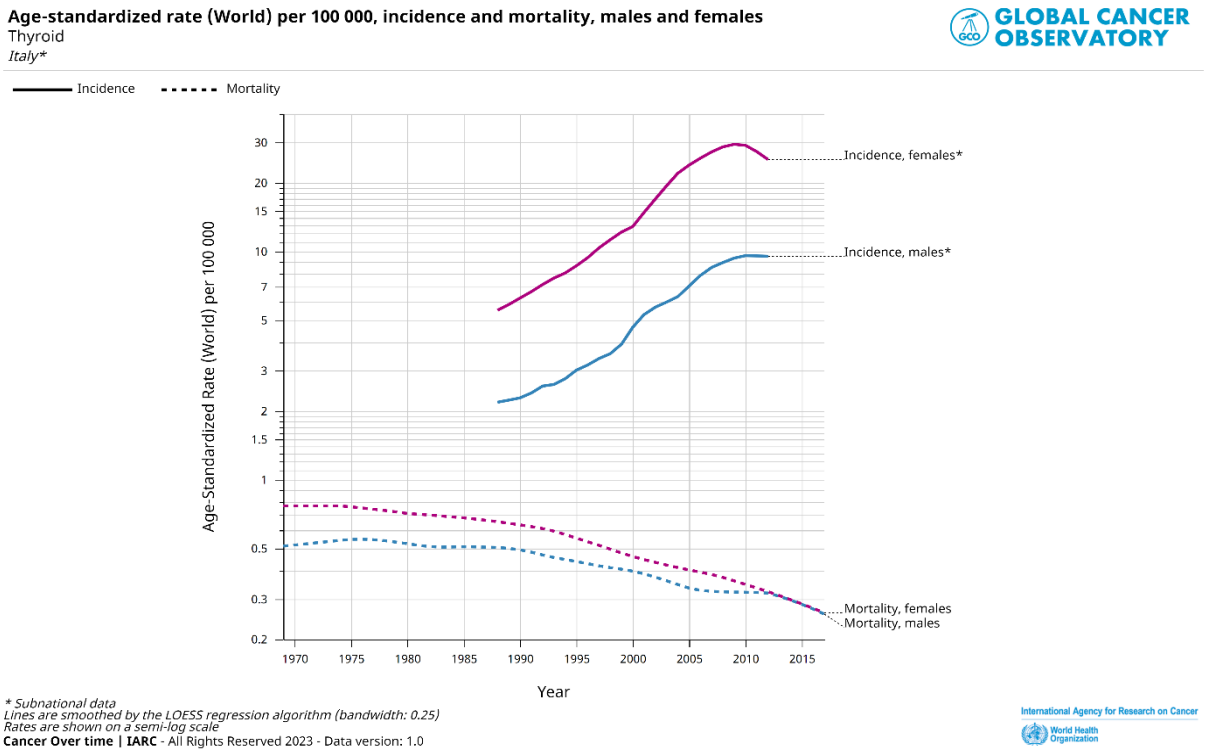


Figure 8 Age-standardized rate (World) per 100000, incidence and mortality, males and females, Thyroid, Italy. Available at: https://gco.iarc.fr/today (accessed on 23 May 2023) [Sung, H. 2021]

6.2. Risk factors

The etiology of thyroid cancer appears to be multifactorial, resulting from a complex interaction of genetic and environmental factors in at-risk individuals. This leads to an in-depth examination of the different risk factors.

6.2.1. Gender

As already mentioned, women have an approximately three times greater risk of developing thyroid carcinoma than men [22]. This gap could be explained by the different hormone balance present in the two genders. However, the issue is still a matter of debate. To date, we know that estradiol, the main female sex hormone, is a potent stimulator of both benign and malignant human thyroid cells. Hormone therapy, i.e., the introduction of exogenous estrogen, is not a risk factor for the development of thyroid carcinoma [29].

6.2.2. Exposure to ionizing radiation

The thyroid can be exposed and irradiated by ionizing radiation more than other tissues because of its location in the body, but mainly because of its ability to concentrate iodine. Specific sources of thyroid irradiation are:

- thyroid imaging with ^{131}I now replaced by ^{99m}Tc, although the therapeutic use of ^{131}I for hyperthyroidism continues and a small increase in thyroid cancer has been observed in these adult patients;
- radiotherapy for malignant neoplasms of the head and neck,
- dental X-rays.

A correlation between exposure to ionizing radiation and carcinoma in the thyroid gland has been demonstrated, especially development of papillary carcinoma. RET/PTC rearrangement is found in 80% of radiation-induced papillary carcinomas [30].
Thyroid is very radiosensitive especially at a young age. Ron et al. indicated that the risk of thyroid cancer in children exposed to head and neck radiation is inversely correlated with age [31]. This was also shown by the rise in thyroid cancer cases in kids and teenagers following the Chernobyl nuclear power plant accident. However, a carcinogenic effect of radiation on the adult thyroid population cannot be ruled out, as indicated by the increased incidence of thyroid cancer among female survivors of the atomic bomb in Japan who were over 20 years old at the time of the explosion [32].

6.2.3. Low intake of iodine

Iodine deficiency promotes the development of follicular thyroid carcinoma (FTC). It does not encourage a rise in overall instances; in fact, there is no evidence of a higher prevalence of FTC in iodine-deficient locations. Instead, an increased incidence of follicular carcinoma, has been seen to occur more frequently in iodine-deficient areas than papillary carcinoma (PTC). This was also demonstrated by the fact that the introduction of iodine prophylaxis produced an increase in the PTC/FTC ratio, but not an increase in the total number of cases. Regarding the relationship between iodine deficiency and PTC, it was seen that in cases of exposure to ionizing radiation, this deficiency is a predisposing factor for the development of PTC. This was confirmed by the fact that in the Chernobyl population, PTC cases were much more numerous in areas with iodine deficiency than in areas without iodine deficiency [33].

6.2.4. Lifestyle and diet

Studies aimed at identifying dietary and lifestyle risk factors for cancer have yielded controversial results because foods and beverages have a large number of different constituents and also because dietary and lifestyle intake can change significantly in the same individual over time. There are dietary factors that interfere with iodine and thyroid hormones that could increase the risk of thyroid cancer, but this possibility has never been proven. Among these dietary factors are nitrites and nitrates that inhibit iodine intake [34].
As far as physical activity is concerned, it appears to improve DNA repair capacity, reduce body fat content, lower insulin resistance and alter circulating inflammatory factors. It turns out to be a protective factor [35]. Another protective factor would appear to be the consumption of cruciferous plants (i.e., Brussels sprouts and cabbage). Cigarette smoking also seems to be a protective factor as it lowers TSH levels [36] and has an anti-estrogenic effect [37].

6.2.5. Body weight

Body mass index appears to be a risk factor exclusive for men. In a study by Zao et al. it was seen that a high BMI (body mass index) was significantly associated with an increased risk of developing papillary thyroid carcinoma. For every 5-unit increase in body mass, the risk-adjusted odds of malignancy increased by 37%. It was also significantly associated with larger tumor size and multifocal tumor. The potential mechanisms behind the link between obesity and cancer are complex and involve a number of factors [38].

6.2.6. Hashimoto's thyroiditis

Hashimoto's thyroiditis (HT), also known as chronic lymphocytic thyroiditis, is an autoimmune disease in which the thyroid gland is gradually destroyed. It is the most common cause of primary hypothyroidism. The relationship between Hashimoto's thyroiditis and thyroid carcinoma has been studied for decades. Its role as a risk factor for the development of thyroid carcinoma has been hypothesised through two mechanisms. First, the inflammatory response caused by thyroiditis fosters a favorable environment for malignant transformation. Damage to stromal cells by cytokines and growth hormones causes alterations in their reactivity and these alterations can lead to malignant transformation of epithelial cells [39]. Second, thyroiditis causes an increase in TSH, which is also a growth factor for thyroid cells. An increased TSH level in most patients with HT stimulates hyperplasia of the follicular epithelium promoting papillary thyroid carcinoma. Regarding the molecular aspect, the RET/PTC gene rearrangement was discovered to be involved in the change from HT to papillary thyroid carcinoma (PTC). The *BRAF* V600E mutation, usually mutually exclusive with the RET/PTC gene rearrangement, is more prevalent in PTC without HT [40].

6.2.7. Thyroid nodules

The presence of a thyroid nodule is a risk factor for thyroid carcinoma. To date, it is uncertain whether there is a difference in prevalence of thyroid carcinoma in uninodular versus multinodular subjects. It appears that patients with multiple thyroid nodules have the same risk of malignancy as those with solitary nodules [41].

6.2.8. Hereditary conditions

There are genetic risk factors in thyroid cancer that lead to diagnoses of thyroid carcinoma in different members of the same family, thus forming familial clusters. The neoplasms found in these individuals are called 'familial non-medullary thyroid carcinoma'[42]. This name highlights the difference between this neoplasm and medullary carcinoma, a thyroid neuroendocrine tumor, which in 25% of cases is caused by the manifestation of the hereditary syndrome Multiple Endocrine Neoplasia Type 2 (MEN 2)[43].

6.3. Clinical overview and diagnosis

Especially in the early stages, thyroid neoplasia presents little or no symptomatology. The carcinomatous nodule generally gives sign of its presence through occupation of space and compression of surrounding tissues. The period between its formation and the patient's first symptoms can range from a few months, in the case of the most aggressive tumors, to ten years. Patients usually experience a localized enlargement of the neck region, named goiter, or, in cases of multifocality, to a generalized enlargement of glandular tissue. The first manifestation may be a mass in a cervical lymph node, in this case it is an isolated cervical nodal metastasis that may not even have an impact on prognosis. The presence of thyroid nodule is not an indicator of cancer as hyperplasia and hypertrophy are physiological responses of cells under stressor.

After a palpation examination performed by an endocrinologist, the patient should undergo blood assay of TSH, FT3 and FT4 to check the function of the gland. This is followed by ultrasound examination (US), which provides key data for clinical framing, including possible lymph node involvement. At the same time as performing ultrasound, the guidelines advise a cytological examination using the Fine-Needle-Aspiration (FNA) technique in cases with US features of concern [44]. The combination of US to FNA is extremely effective for thyroid cancer screening. In that the former has high sensitivity and high negative predictive values, while the latter has high specificity and high positive predictive values [45, 46]. US-guided FNA should be performed on nodules >1 cm in diameter, and is discouraged in cases of smaller lesions. FNA is the gold standard in the differential diagnosis of thyroid nodules, allowing accurate identification of benign lesions from those with suspected potential for malignancy. Through FNA, cells are collected from the thyroid nodule and are spread on slides and stained with Romanowsky-type staining (May-Grunwald-Giemsa, Wright or Diff-Quik staining) and/or Papanicolaou staining (PAP). Covered by coverslips, the smears are then analyzed microscopically by a pathologist who makes a cytologic diagnosis. The cytologic diagnoses follow criteria established by guidelines, such as those of The Bethesda System for Reporting Thyroid Cytopathology, British Thyroid Association guidelines and Italian thyroid cytology classification system [44, 47, 48].

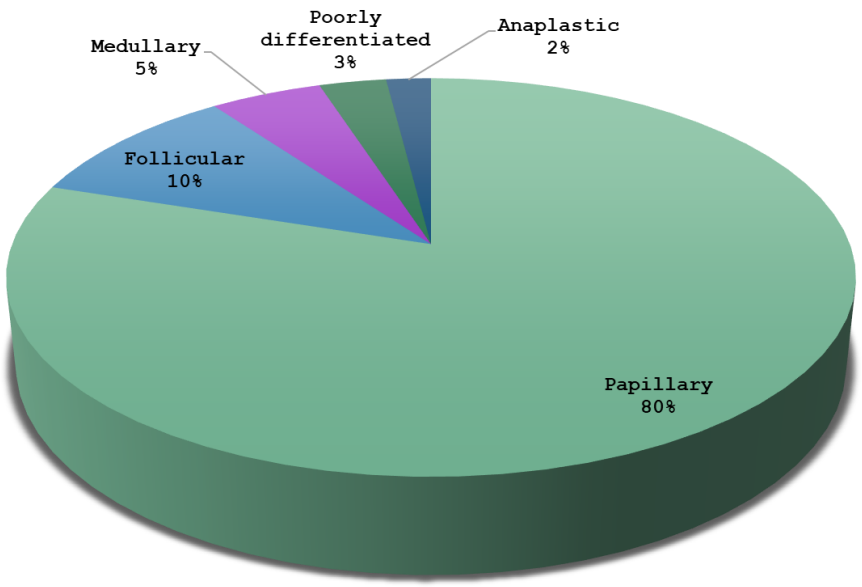
6.4. Thyroid carcinoma subtypes

There are different histotypes of thyroid carcinoma:

- Differentiated carcinoma,
  - Papillary thyroid carcinoma,
    - Papillary thyroid carcinoma follicular variant,
  - Follicular carcinoma,
- Poorly differentiated carcinoma,
- Anaplastic carcinoma,
- Medullary carcinoma.

Differentiated thyroid carcinomas, papillary and follicular, are so called because the cells can resemble normal thyroid tissue. On the other hand, undifferentiated carcinomas, poorly differentiated carcinoma and anaplastic, get their name from the fact that they lose the morphological characteristics of the tissue from which they originate.

Medullary carcinomas are a separate group because unlike the others that originate from follicular cells, they originate from C cells. Differentiated carcinomas constitute 90% of all thyroid carcinomas, followed by 5% medullary carcinomas, 3% poorly differentiated carcinomas, and 2% anaplastic carcinomas. The 90% incidence of differentiated carcinomas is due to 80% Papillary carcinoma and 10% Follicular carcinoma (Graphic 1)[49].



Graphic 1 Percentage composition of thyroid carcinomas: 80% Papillary carcinoma, 10% Follicular carcinoma, 5% medullary carcinoma, 3% poorly differentiated carcinomas, and 2% anaplastic carcinomas

6.4.2. Papillary thyroid carcinoma

The most common type of thyroid cancer is papillary thyroid carcinoma (PTC), a differentiated-type carcinoma arising from follicular cells (figure 9). That accounts for roughly 70% to 80% of thyroid cancer cases and 66,8% of deaths from endocrine tumors [26]. Commonly it is diagnosed between the ages of 25 and 50, although it can occur at any age. The development of this tumor is frequently associated with previous exposure to ionizing radiation. PTC can be defined as such if it is larger than 1 cm in size, otherwise it is referred to as papillary microcarcinoma, which is considered its precursor. Papillary carcinomas can be solitary or multifocal. In terms of morphology, most of papillary carcinoma show irregular borders with infiltration of thyroid parenchyma and only 8-13% are encapsulated. Tumors are frequently cystic and may have fibrotic or calcified regions (Psammoma bodies). The presence of papillae with a central fibrovascular core is one of the cardinal features for diagnosis. Regarding the microscopic distinctive features of papillary neoplasm, we can find: nuclear enlargement and elongation, chromatin clearing, margination and glassy nuclei (Orphan Annie eye nuclei), irregular nuclear contour with invaginations of the nuclear membrane and nuclear pseudoinclusion. As for invasion generally this involves the lymphatic vessels. Rarely there is invasion of blood vessels, especially for small lesions. Metastasis to adjacent cervical lymph nodes is seen in half of the cases [49]. In general, the prognosis of PTC is very good: the 5-year relative survival rate for regional carcinoma, spread only to nearby tissues or organs and/or regional lymph nodes, is 99% [50]. The frequency of local or regional recurrence is between 5% and 20% of patients and distant metastases between 10% and 15% [49].

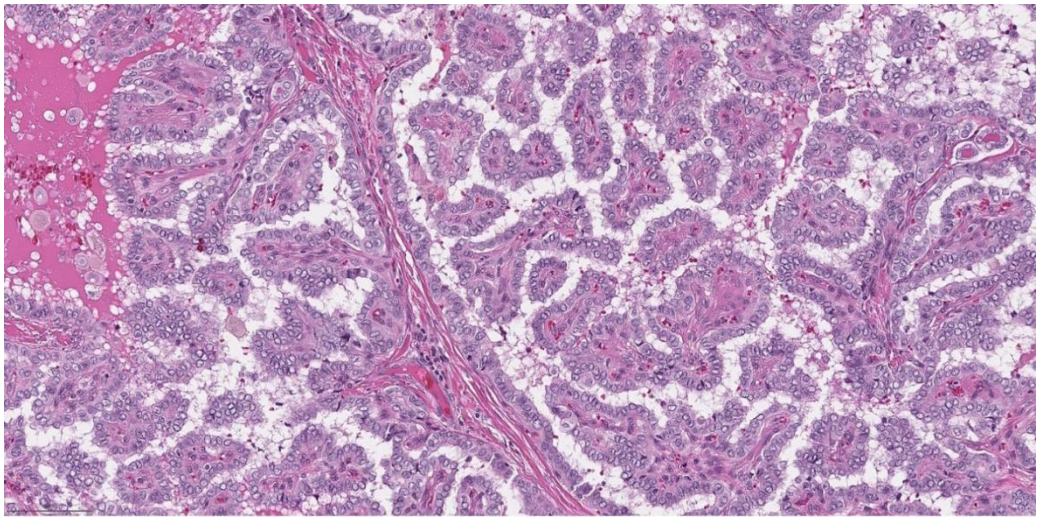


Figure 9 Papillary thyroid carcinoma, branching papillae and fibrovascular cores. Available at: https://www.pathologyoutlines.com/topic/thyroidpapillaryclassic.html (accessed on 24 October 2023)

Molecular signature of PTC

In papillary thyroid carcinomas, the mitogen-activated protein kinase (MAPK) pathway is commonly involved in genetic alterations. The most frequently mutated gene is *BRAF*, considered to be present mutated in 55-80% of cases, where it causes constitutive activation of this pathway [51, 52]. Generally, the mutation present is *BRAF* p.V600E, but in 1,2% of cases *BRAF* nonV600E mutations are present, the most frequent one being *BRAF* p.K601E [53]. In the thyroid of mice with expression of mutated *BRAF* V600E protein, thyroid carcinoma develops. This leads to the view that this mutation is initial transforming event. Considering the percentage of neoplastic cells and the mutation, however, it was seen that in some papillary carcinomas the *BRAF* mutation is not present in the entire population of neoplastic cells, but is present only in a subpopulation of cells. This means that the mutation was not the foundational event of carcinogenesis, but occurred during development [52]. Mutations in the *RAS* gene family are found in 13% of cases. Mutations in the *TERT* gene promoter were found in 9% of PTCs and is usually considered a marker of malignancy. TP53 gene is mutated in 9% e PTEN in 1% [54]. As for nonpoint variants, RET-PTC rearrangements is the molecular signature usually present in carcinomas resulting from exposure to ionizing radiation and are found in 40% of cases. 3% of PTC involve gene fusions involving an NTRK gene [55]. Mutually exclusive to *BRAF* and *RAS* mutation, EIF1AX mutation may occur in 10% of PTC [56].

6.4.2.1. Follicular variants of PTC

Only 50% of papillary thyroid carcinomas are classical variant, but the remaining 50% are formed by other variants, a total of 15 can be counted [57, 58]. Follicular variant of papillary thyroid carcinoma (FV-PTC) is the most prevalent subtype of PTC, account for about one-third of all PTC cases (Figure 10). As with papillary and follicular carcinoma, the age of onset is around 45 years [59]. These tumors have the architecture of a follicular carcinoma, in fact, at first glance they look like follicular carcinomas, but they present the nuclear features of papillary carcinomas [49]. We can therefore state that FV-PTC represents an intermediate entity between PTC and FTC. The mean tumor size of FV-PTC is somewhat higher than that of PTC, but lower than that of FTC.
The disease course is not homogeneous in all patients because a distinction in two subtypes must be made [58]:

- Encapsulated follicular variant (EFV-PTC): marked by the presence of a capsule or a well-defined border;
- Infiltrative follicular variant (IFV-PTC): associated with infiltrative growth.

EFV-PTC has a behavior more similar to FTC, in fact it can invade the capsule and blood vessels, but only in 5% of cases invade the lymphatic vessels. In general, EFV-PTC has a good prognosis. IFV-PTC have a higher rate of extrathyroidal extension and lymph node metastasis occurring in 65% of cases. IFV-PTC is a more aggressive type than EFV-PTC[60]. Noninvasive encapsulated FV-PTC has been reclassified to noninvasive follicular thyroid neoplasm with papillary-like nuclear features (NIFTP), a benign neoplasm that appears to be the precursor to the EFV-PTC [61].

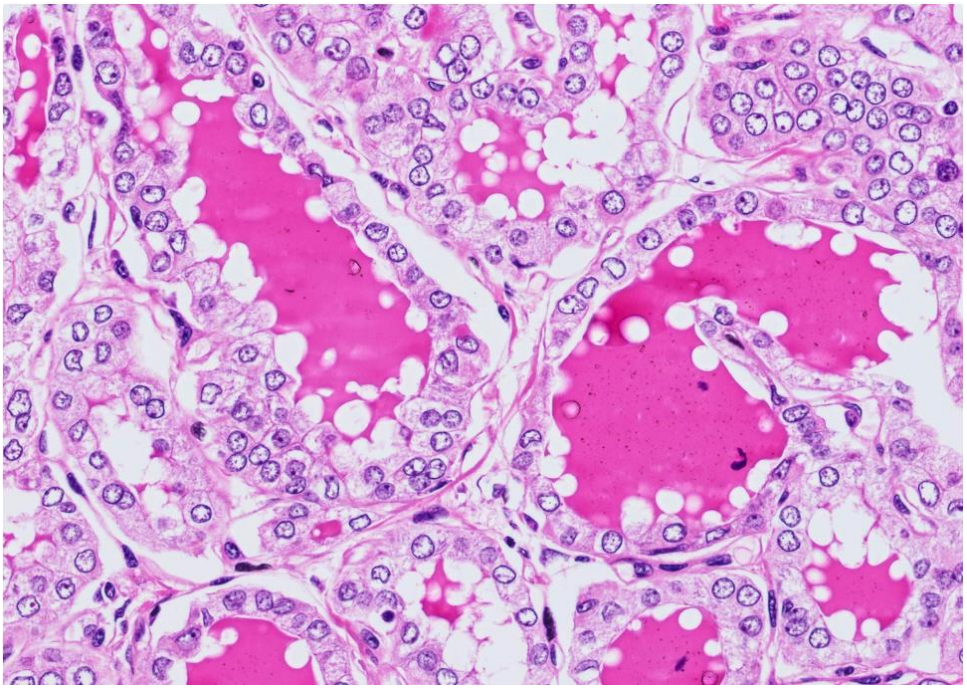


Figure 10 Infiltrative follicular variant of papillary thyroid carcinoma. (HE staining) [Kakudo, K. 2018]

Molecular signature of FV-PTC

Regarding the molecular aspect of the follicular variant of papillary thyroid carcinoma, there are differences between encapsulated and infiltrative carcinomas. Encapsulated carcinomas are mutated in 40% of cases in *RAS* gene family and do not carry mutations in *BRAF* gene, while infiltrative carcinomas are mutated *BRAF* in 26% of cases and *RAS* in 10%. Also in this case, *BRAF* and *RAS* mutations are mutually exclusive. In a study by Rivieri et al. considering tumors wild type for *RAS* and *BRAF* analysis, they found in one encapsulated follicular variant a *PAX8/PPARγ* (paired box 8/peroxisome proliferator activated receptor γ) rearrangement, while two infiltrating tumors had *RET/PTC* recombination [62, 63].

6.4.3. Follicular thyroid carcinoma

Follicular carcinoma, derived from follicular cells, accounts for 10% of primary thyroid carcinomas (figure 11). Unlike papillary carcinoma, it occurs in older patients, with a peak incidence between the ages of 40 and 60, but very rare in children [49]. A diet with low iodine intake is a risk factor for this cancer [64].These tumors usually have well-defined boundaries and are encapsulated, difficult to distinguish from follicular adenomas, but the differential diagnosis of carcinoma is determined if the neoplasm shows invasion, either of the capsule or blood vessels. In most cases, these are cold nodules on scintigraphy: nodules with medium to low radioactive iodine uptake. Better differentiated lesions, which are rarer, may appear warm as they are hyperfunctional and thus absorb more radioactive iodine [65]. Microscopically, most follicular carcinomas are composed of fairly uniform cells that form small follicles containing colloid, making them translucent, similar to normal thyroid tissue. Foci of calcification and fibrosis may be present. Follicular differentiation may also be less evident, and nests or sheets of cells without colloid may be present. The nuclei do not show the typical features of papillary carcinoma. Increased mitotic activity occurs. FTC could be divided in three classes:

- Follicular carcinoma with capsular invasion only;
- Encapsulated minimally angioinvasive: invasion <4 blood vessels;
- Widely invasive: extensive invasion of thyroid, extrathyroidal soft tissue and/or invasion more than 4 blood vessels. Prognosis is directly proportional with the degree of invasion: carcinomas with only capsule invasion have good prognosis while infiltrating carcinomas give worse prognosis and obviously an increased chance of distant metastasis to bone and lung tissue. The 10-year mortality of the widely invasive subtype is nearly 50 percent, in contrast with 10-year survival rate of 90% of minimally invasive [66]. They don’t metastasized through lymphatics vessels [49].

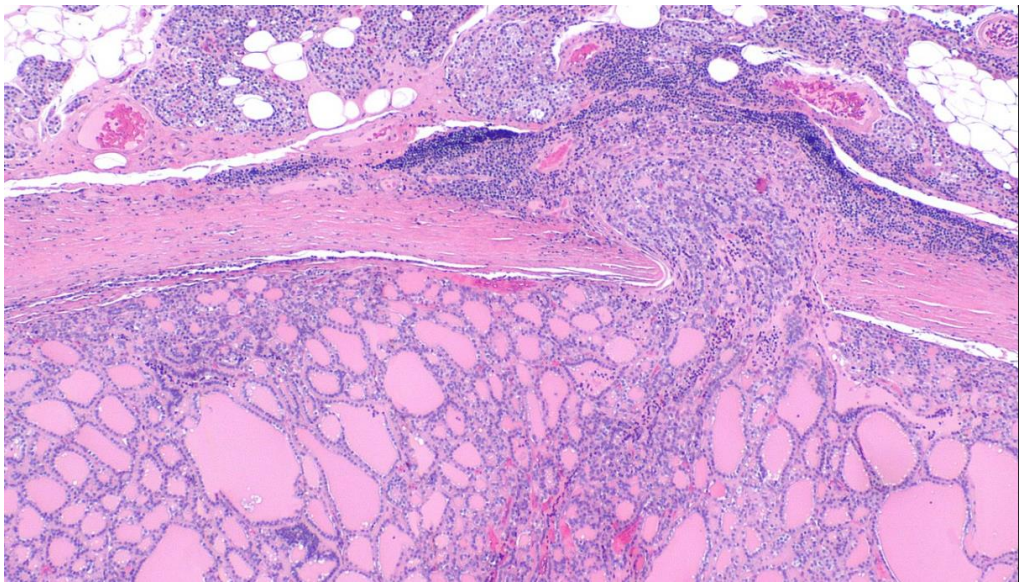


Figure 11 Follicular thyroid carcinoma with capsular invasion (H&E staining) Available at: https://en.wikipedia.org/wiki/Follicular_thyroid_cancer (accessed on 24 October 2023)

Molecular signature of FTC

Follicular carcinoma is characterized by the possible presence of mutations in *RAS* genes found in 50% of cases. Mutations in the *BRAF* gene are particularly rare, with a 2% of frequency. Mutations in *TERT* promoter are found in 17% of FTC and in this case it is also associated with poor prognosis. *PTEN* mutation is present in 7% of cases, *EIF1AX* and *PIK3CA* mutations in 5%. *TP53* is mutated in FTC with the same mutation frequency found in PTCs: 6%. As for rearrangements, *PAX8-PPARγ* fusion is present in 40% of cases. No *RET/PTC* rearrangements have ever been found [67, 68].

6.4.4. Poorly differentiated thyroid carcinoma

Poorly differentiated thyroid carcinoma (PDTC) is a follicular cell derived thyroid carcinoma, accounting about 3% of all thyroid carcinomas diagnosed (figure 12). It generally affects older patients, with a mean age at diagnosis of 59 years, and the incidence gap between women and men is about 1.5 times, much smaller than that for differentiated carcinomas [58]. Considering its morphology and clinical behavior, it is considered an intermediate type of carcinoma between differentiated carcinomas, particularly thyroid carcinoma, and anaplastic thyroid carcinoma. According to the Turin criteria, established in 2006 by an international group of pathologists, PDTC is characterized by a solid/trabecular/insular growth pattern with absence of conventional nuclear features of papillary carcinoma. It should also exhibit one of these features: twisted nuclei, mitotic activity ≥3/10 high-power microscopic fields, and tumor necrosis [69].
Prognosis is intermediate between indolent differentiated thyroid carcinomas and fast-growing, often fatal anaplastic carcinoma.
69% of patients have extrathyroidal extension [70], the risk of developing metastases in neck lymph nodes is very high (50-85%) and distant metastases occur in 85% of cases [71]. Five-year overall survival is estimated at 62-85% [70].

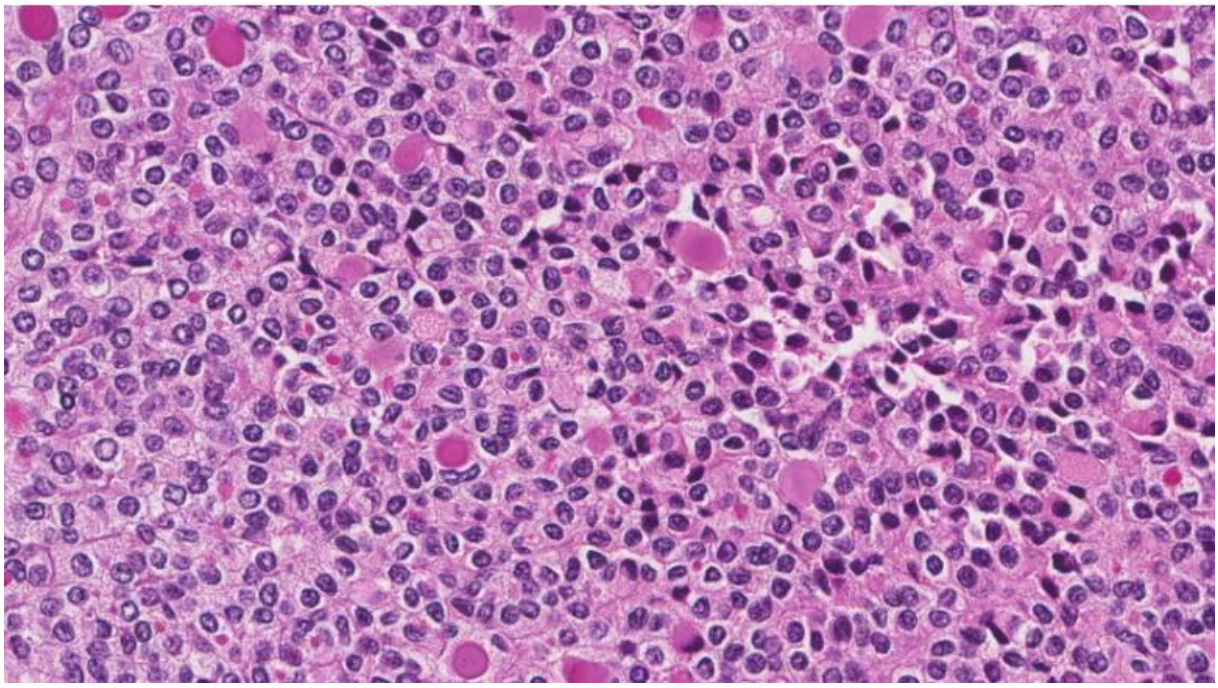


Figure 12 Poorly differentiated thyroid carcinoma (H&E staining). Available at: https://www.mypathologyreport.ca/it/diagnosis-library/poorly-differentiated-thyroid-carcinoma/(accessed on 09 June 2023)

Molecular signature of PDTC

Regarding the molecular signature, the most frequently found mutation is *TERT* with 40% of cases, and this is associated with a more aggressive phenotype and a higher mortality rate. They can present either *BRAF* mutation or a mutation of *RAS* genes with the same frequency: about 25%. Recall that these mutations are mutually exclusive [72]. Interestingly, *BRAF* mutated PDTCs are less differentiated and metastasize more frequently in regional lymph nodes, whereas *RAS* mutated PDTCs are more differentiated and show significantly higher rates of distant metastasis [73]. Mutations in the TP53 gene are detected in 16% and mutations in *ATM*, gene involved in cell cycle checkpoint control and DNA repair, in 8%. Mutations in *MED12, RBM10* and *EIF1AX* were found in 15%, 12% and 11% of cases, respectively. These mutations are predictive for poor survival. Mutations in the *PI3K-AKT-mTOR* pathway are uncommon: *PTEN* mutated in 4% and *PIK3CA* in 3% As for gene rearrangements, they were found in 14% of cases and included *RET/PTC, PAX8/PPARγ* and *ALK* fusions [72].

6.4.5. Anaplastic thyroid carcinoma

Anaplastic thyroid carcinoma (ATC) arises from follicular cells, but it is completely undifferentiated and has lost the normal architecture of the starting healthy tissue (figure 13). It's a very rare neoplasm, accounting for only 2% of thyroid carcinomas and occurring at an advanced age: 65-70 years [49]. This is a very aggressive carcinoma and only 10% of patients at diagnosis have the tumor localized only to the thyroid gland. Usually at diagnosis, it is a large tumor with a mean size of 5-6 cm [74]. Anaplastic carcinoma is characterized by variable morphology and could be present: large pleomorphic giant cells, including multinucleated osteoclast-like giant cells, spindle cells with sarcomatous appearance, epithelioid or squamoid histologic appearance. Is a very heterogeneous neoplasm so they can have more than one morphology and sometimes even foci of papillary or follicular differentiation. Compared with PDTC, ATC is more pleomorphic, has a higher mitotic rate, more atypical mitoses, and has a more infiltrative growth pattern [49].
The sites usually affected by locally invasion are strap muscles, trachea, esophagus, and larynx, while distant metastases are found in lungs, bone and brain [75].
ATC is a rapidly fatal disease based on median survival rates of 3 to 6 months. The 2-years survival rate of patients is 10%.

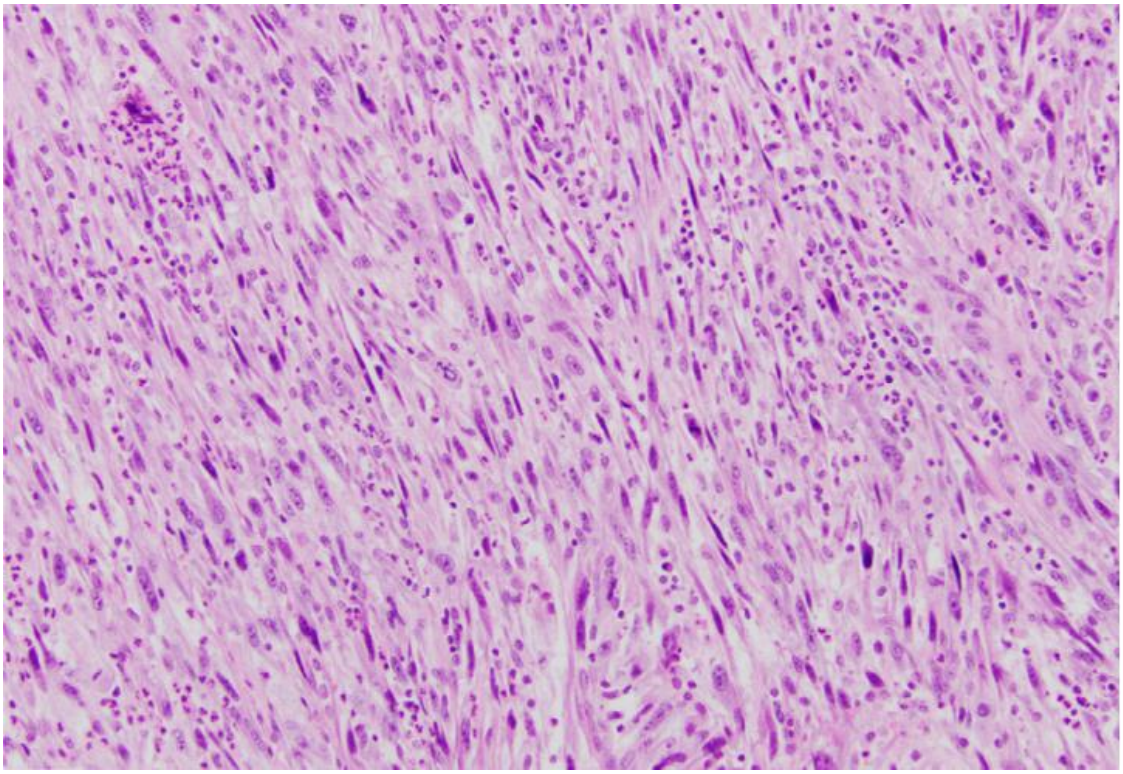


Figure 13 Anaplastic thyroid carcinoma (H&E staining) [Jiang, F. 2017]

Molecular signature of ATC

Regarding the molecular aspect, *TP53* is the most frequently mutated gene (70% of cases). Mutations in the *RAS* gene family are found in about 25% of cases [56] , while the frequency of mutations in the *BRAF* gene varies between study cohorts, from 19% to 45%, but generally found in elderly patients [76]. *TERT* mutations, usually the C228T mutation, C250T is rarely observed, occur in 55% of cases and are associated with a higher incidence of distant metastasis. The concomitant presence of mutation in *BRAF* or *RAS* genes with *TERT* is associated with a worse prognosis than the outcome for a single-mutation tumor [56]. *EIF1AX* mutations, 10% of ATC, are closely related to *RAS* mutations, while *PIK3CA* mutations,10% also, tend to be present with *BRAF* mutations. *NF1* and *NF2* are mutated in 10% of tumors, while *PTEN* in 12%. Generally do not present gene rearrangements [56].

6.4.6. Medullary thyroid carcinoma

Medullary thyroid carcinoma is a neuroendocrine tumor arising from parafollicular cells and accounts for 5% of thyroid cancers (figure 14). The incidence between men and women is the same [22]. Like the parafollicular cells from which it originates, it produces calcitonin, which is an excellent marker for diagnosis and patient follow-up. Cancer cells can elaborate other hormones such as serotonin, adrenocorticotropic hormone and vasoactive intestinal peptide[49]. It presents as a well-defined but unencapsulated carcinoma. There may be areas of necrosis and hemorrhage in some lesions. Microscopically, medullary carcinomas have spindle- or polygonal-shaped cells that can form follicles, nests, and trabeculae. There might be a small anaplastic cells. The stroma may include amyloid deposits made of calcitonin polypeptides. This carcinoma can be of two types: sporadic or familial. The sporadic form accounts for 70% of medullary carcinomas, diagnosed between the ages of 40 and 60, usually presenting as a single nodule. The familial form can be caused by MEN 2A or 2B syndromes, familial medullary thyroid carcinoma (FMTC) syndrome, von Hippel-Lindau disease or neurofibromatosis. In this form, the incidence is much higher in young patients and it is not usual for it to be multinodular. Histologically, unlike the sporadic form, they show areas of multicentric C-cell hyperplasia in the surrounding thyroid parenchyma. Both forms tend to metastasize to lymph nodes in 70-80% of case. The 5-year survival for medullary carcinomas is 65-89% and the familial form is more aggressive [77].

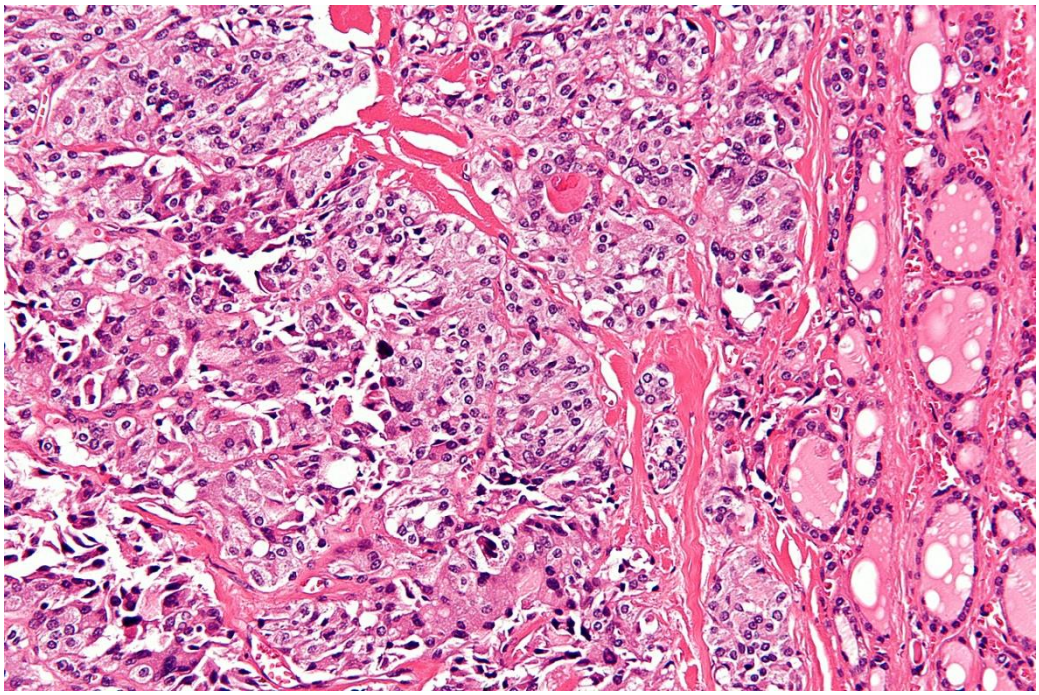


Figure 14 Medullary thyroid carcinoma (H&E staining). Available on: https://it.wikipedia.org/wiki/Carcinoma_midollare_della_tiroide. (Accessed on 24 October 2023)

Molecular signature of MTC

The principal gene in medullary carcinoma is RET, that encode a transmembrane receptor tyrosine kinase involved in several cell process such as proliferation, motility and apoptosis. This gene is mutated in 50% of sporadic forms and 95% of familiar forms. Genetic counselling for family members of patients with RET mutated medullary carcinoma is very important in order to rule out MEN syndrome [78]. In general, mutated *RET* medullary carcinomas have a more aggressive behavior. Mutations that may be present are those of the *RAS* gene family. Depending on the studies, they range in incidence from 0 to 40 per cent, and the most frequently mutated gene is *HRAS*. RET and *RAS* mutations are not mutually exclusive, but the *RAS* mutation rate is very higher in RET WT cases [79].

6.5. Prognosis and treatment

Malignant thyroid tumors are staged according to their mortality risks using the Tumor, Node, and Metastasis (TNM) classification system developed by the Union for International Cancer Control (UICC) (table 1). In this way, by assessing the primary tumor, regional lymph nodes and distant metastases, patients can be stratified.

TNM

T—primary tumor	
TX	Primary tumor cannot be assessed
T0	No evidence of primary tumor
T1	Tumor 2 cm or less in greatest dimension, limited to the thyroid
T1a	Tumor 1cm in greatest dimension, limited to the thyroid
T1b	Tumor >1 cm but 2 cm in greatest dimension, limited to the thyroid
T2	Tumor >2 cm but 4 cm in greatest dimension, limited to the thyroid
T3	Tumor >4 cm in greatest dimension, limited to the thyroid or with gross extrathyroidal extension invading only strap muscles
T3a	Tumor >4 cm in greatest dimension, limited to the thyroid
T3b	Tumor of any size with gross extrathyroidal extension invading strap muscles
T4a	Tumor extends beyond the thyroid capsule and invades any of the following: subcutaneous soft tissues, larynx, trachea, esophagus, recurrent laryngeal nerve
T4b	Tumor invades prevertebral fascia or encasing the carotid artery or mediastinal vessels from a tumor of any size
N—regional lymph node	
NX	Regional lymph nodes cannot be assessed
N0	No evidence of locoregional lymph node metastasis
N1	Regional lymph node metastasis
N1a	Metastasis to level VI (pretracheal, paratracheal and prelaryngeal/Delphian lymph nodes) or upper/superior mediastinum
N1b	Metastasis in other unilateral, bilateral or contralateral cervical compartments (levels I, II, III, IV or V) or retropharyngeal
M—distant metastasis	
M0	No distant metastasis
M1	Distant metastasis

Continue in next page

Stage papillary or follicular <55 years			
I	Any T	Any N	M0
II	Any T	Any N	M1

Stage papillary or follicular ≥55 years			
I	T1a/b	N0/NX	M0
	T2	N0/NX	M0
II	T1a/b	N1a/b	M0
	T2	N1a/b	M0
	T3a/b	Any N	M0
III	T4a	Any N	M0
IVa	T4b	Any N	M0
IVb	Any T	Any N	M1

Stage medullary			
Stage I	T1a, T1b	N0	M0
Stage II	T2, T3	N0	M0
Stage III	T1–T3	N1a	M0
Stage IVa	T1–T3	N1b	M0
	T4a	Any N	M0
Stage IVb	T4b	Any N	M0
Stage IVc	Any T	Any N	M1

Stage anaplastic			
Stage IVa	T1, T2, T3a	N0	M0
Stage IVb	T1, T2, T3a	N1a	M0
	T3b, T4a, T4b	N0, N1a	M0
Stage IVc	Any T	Any N	M1

Table 1 TNM staging [ESMO 2019]

The management of patients with thyroid carcinoma depends on the patient stratification mentioned above and the type of carcinoma they have. Treatment for thyroid carcinoma may involve a combination of surgery, radioactive iodine therapy, hormone therapy, targeted therapy, and, in some cases, external beam radiation therapy.

Surgical removal of the thyroid gland (lobectomy or total thyroidectomy) is the primary treatment for most thyroid carcinomas. In cases of localized disease, it may be curative. Lymph node dissection may be performed in certain cases to assess the spread of cancer.

After surgery is used Radioactive Iodine Therapy (RAI) to destroy any remaining thyroid tissue and any microscopic cancer cells that might have spread. This is particularly relevant for papillary and follicular carcinomas that can absorb radioactive iodine.

With regard to therapy, the following is currently available hormone therapy and targeted therapy. Thyroid hormone replacement therapy (e.g., levothyroxine) is often given to patients after thyroidectomy. It helps maintain the balance of thyroid hormones in the body and suppresses TSH, which can stimulate the growth of thyroid cancer cells. In cases of advanced or metastatic thyroid carcinoma, targeted therapies may be used. For example, tyrosine kinase inhibitors like Lenvatinib and sorafenib can be employed in patients with progressive, radioactive iodine-refractory differentiated thyroid carcinoma. As a case, molecular biology analysis is needed to identify targets for therapy.

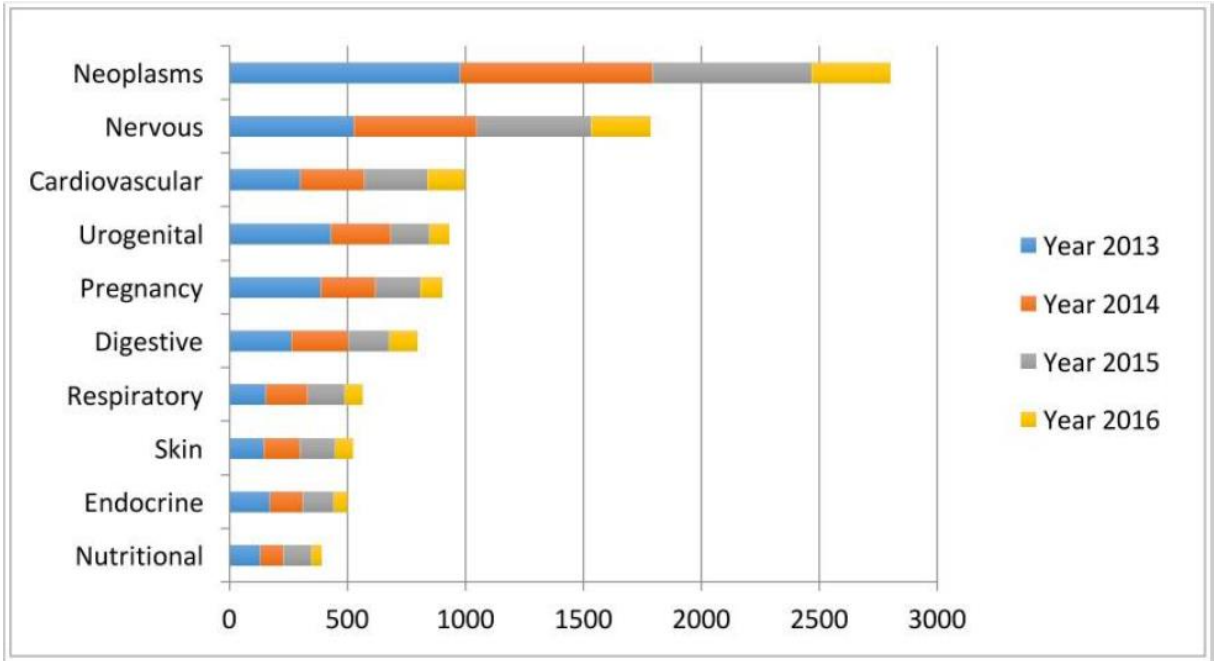
External beam radiation therapy may be used in certain situations, such as in the treatment of anaplastic thyroid carcinoma, where the cancer is less responsive to other treatments [80].

Cap 2 – Deep learning applied to thyroid cytology

Artificial intelligence (AI) has appeared in the healthcare field in recent years, and through various applications it has proven to be a valuable support in both research and clinical fields.
Through a project in collaboration with Prof. Diciotti, affiliated with the DEI-UNIBO "Guglielmo Marconi" Department of Electrical Energy and Information Engineering, we have developed an artificial intelligence model to support screening of fine-needle thyroid smears.
To do this, the first step was to collect images of thyroid cytology specimens, selected them according to study criteria, and then annotate all image information in a database, included diagnostic category according to The Bethesda System for Reporting Thyroid Cytopathology guideline. Next, engineers, led by Prof. Diciotti, trained a deep learning (DeepL) network so that it could discriminate between specimens diagnosed as follicular lesions of undetermined significance and others diagnosed as suspicious for malignancy or positive for malignancy.
For this study we precisely chose this aim due its relevance to patient management in fact patients with follicular lesions of undetermined significance generally do not undergo surgery but are followed in follow up, while others with thyroid nodules diagnosed as suspicious for malignancy or positive for malignancy undergo thyroid re-section (lobectomy or total thyroidectomy).
The development of this Deep Learning model can be a valuable support for pathologists' screening activities in thyroid FNA smears.

1. Artificial intelligence in medicine

One of the first to pose the question of whether machines are capable of simulating human behavior was, in the middle of the last century, Alan Turing who developed the Turing Test specifically to determine whether a machine was capable of intelligent thoughts [81]. Since then, the computational power of machines has increased exponentially such that the terminology "artificial intelligence" (AI) has become commonly used as AI is now integrated into daily life in many forms.
One of the fields where AI has been applied and can still be developed is the healthcare. Artificial intelligence offers valuable help in the fields of new drug development, big data management, support in clinical decision making, and more [82]. The fields where AI can potentially be involved are many, but there are diseases where the implementation of AI has stood out over others, such as cancer, diseases of the nervous system and cardiovascular system (graphic 2) [83].



Graphic 2 Considering the literature in Pubmed regarding AI, these are the 10 diseases where Ai has is most involved. [Jiang, F 2017.]

Pathology is also one of the fields involved in the introduction of AI and regarding cytopathology more specifically, the first publication is the one concerning development of computer-assisted Pap test screening [84].
Considering the level of complexity present in anatomy-pathology specimens there was a need to develop complex and novel imaging systems, such as the whole-slide imaging (WSI) scanner, which are capable of capturing the image of the entire slide in a couple of minutes, and the pathologists can examine it on their own computer monitor, even working remotely. This aspect was crucial during the COVID pandemic as it allowed continuity of work from remote. All this has enabled the transition from traditional pathology to digital pathology [85]. Digital pathology brings with it numerous advantages in both translational research and clinical practice. It simplifies processes through the integration of digital workflow software, reduced testing turnaround time, reduced inter-pathologist variability through the generation of accurate and highly reproducible tissue readings, storage of vast amounts of data, and sharing of image data anywhere in the world.
Deep Learning fits well in this context given its ability to recognize complex systems and build predictive models from large amounts of data. The DeepL is a valuable supporter because it improves diagnostic accuracy and objectivity and reduces the workload of pathologists who can focus on high-level decision-making tasks.
Considering that thyroid nodules are very common in the general population, millions of FNAs are performed each year worldwide [86].
Microscopic screening of thyroid cytology smears is a process that requires time, practice, and experience on the part of the pathologist. Therefore, thyroid cytology is a field in which the application of AI techniques appears to have strong potential to provide valuable diagnostic support. This support is becoming increasingly urgent given the worldwide shortage of pathologists, which is particularly dramatic for developing and low-income countries [87].

2. Thyroid cytology

Thyroid nodules are common in the population, and it is estimated that 65% are affected by them. However, the population investigated, whether or not they have risk factors, and the techniques employed to find nodules all affect the prevalence of nodular thyroid disease. The use of ultrasound greatly increases the probability of finding thyroid nodules compared with palpation alone. Thyroid nodules are mostly benign, but it is estimated that about 5-15% are malignant, typically low-risk carcinomas arising from follicular cells [45, 46].
As I have already described in the introductory section "6.3. Clinical overview and diagnosis," in the management of the patient with thyroid nodule, the recommended procedures by guidelines include ultrasonography (US) and eventually FNA if present alarming US characteristics (Figure 15). Thyroid cytology smears derived from FNA are examined under the microscope by pathologists who make a cytological diagnosis. The Bethesda System for Reporting Thyroid Cytopathology (TBSRTC) criteria are one of the most widely followed guidelines for diagnosing thyroid cytopathology.

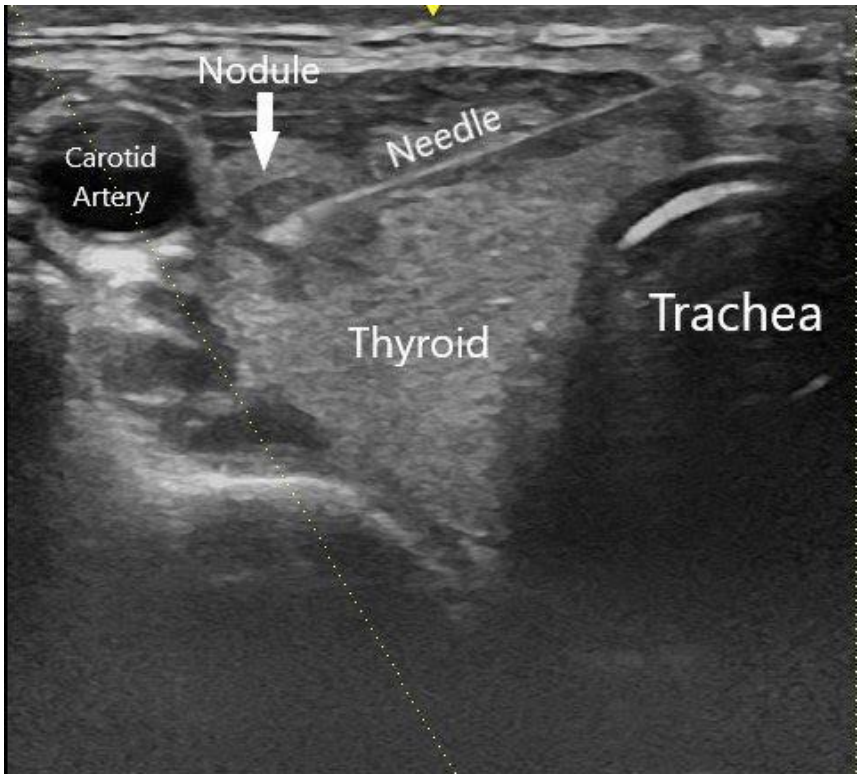


Figure 15 Image of an ultrasound-guided FNA. Available on https://www.entandallergy.com/blog-posts/details/head-and-neck-fine-needle-aspiration-biopsy (accessed on 24 October 2023)

2.1 The Bethesda System for Reporting Thyroid Cytopathology

The Bethesda System for Reporting Thyroid Cytopathology (TBSRTC) consists of six categories: TBSRTC I, TBSRTC II, TBSRTC III, TBSRTC IV, TBSRTC V, and TBSRTC VI. Each category is associated with risk of malignancy and a specific patient management (table 2, 3)[88].

TBSRTC	Criteria of inclusion for each category
I	**Nondiagnostic or Unsatisfactory** Cyst fluid only Virtually acellular specimen Other (obscuring blood, clotting artifact, drying artifact, etc.)
II	**Benign** Consistent with a benign follicular nodule (includes adenomatoid nodule, colloid nodule, etc.) Consistent with chronic lymphocytic (Hashimoto) thyroiditis in the proper clinical context Consistent with granulomatous (subacute) thyroiditis other
III	**Atypia of Undetermined Significance or Follicular Lesion of Undetermined Significance**
IV	**Follicular Neoplasm or Suspicious for a Follicular Neoplasm** Specify if oncocytic (Hürthle cell) type
V	**Suspicious for Malignancy** Suspicious for papillary thyroid carcinoma Suspicious for medullary thyroid carcinoma Suspicious for metastatic carcinoma Suspicious for lymphoma other
VI	**Malignant** Papillary thyroid carcinoma Poorly differentiated carcinoma Medullary thyroid carcinoma Undifferentiated (anaplastic)carcinoma Squamous cell carcinoma Carcinoma with mixed features Metastatic malignancy Non-Hodgkin lymphoma other

Table 2 The Bethesda System for Reporting Thyroid Cytopathology's inclusion criteria for several diagnostic categories [The Bethesda System for Reporting Thyroid Cytopathology 2018].

TBSRTC	Diagnostic categories	Usual management	Risk of malignancy (%)
I	Nondiagnostic/Unsatisfactory	Repeat FNA	5-10
II	Benign	Clinical follow-up	0-3
III	Follicular Lesion of Undetermined Significance/Atypia of Undetermined Significance	Repeat FNA, molecular testing*	10-30
IV	Follicular Neoplasm/ Suspicious for a Follicular Neoplasm	Molecular testing, lobectomy	25-40
V	Suspicious for Malignancy	Molecular testing, near-total thyroidectomy or lobectomy	50-75
VI	Malignant	Molecular testing, near-total thyroidectomy or lobectomy	99.

Table 3 Diagnostic categories based on The Bethesda System for Reporting Thyroid Cytopathology along with usual clinical management and associated risk of malignancy. [The Bethesda System for Reporting Thyroid Cytopathology 2018].

TBSRTC I – nondiagnostic

An FNA sample of a thyroid nodule must be representative of the underlying lesion, otherwise it cannot provide diagnostic information. The aid of ultrasound guidance when performing thyroid FNA is useful to confirm needle placement in the nodule. The adequacy of a thyroid FNA is defined by both the quantity and quality of the cellular components. Samples that do not reflect the quality or quantity criteria fall into this category (figure 16). In this category we find specimens in which there is only cystic fluid, only staining, only artifacts or other, but not a sufficient and representative amount of cells to be able to make a diagnosis. In general, the risk of malignancy in these cases is 5-10%. In this case an FNA should be repeated to collect representative material, because the malignancy of the lesion cannot be excluded by not having an optimal specimen.

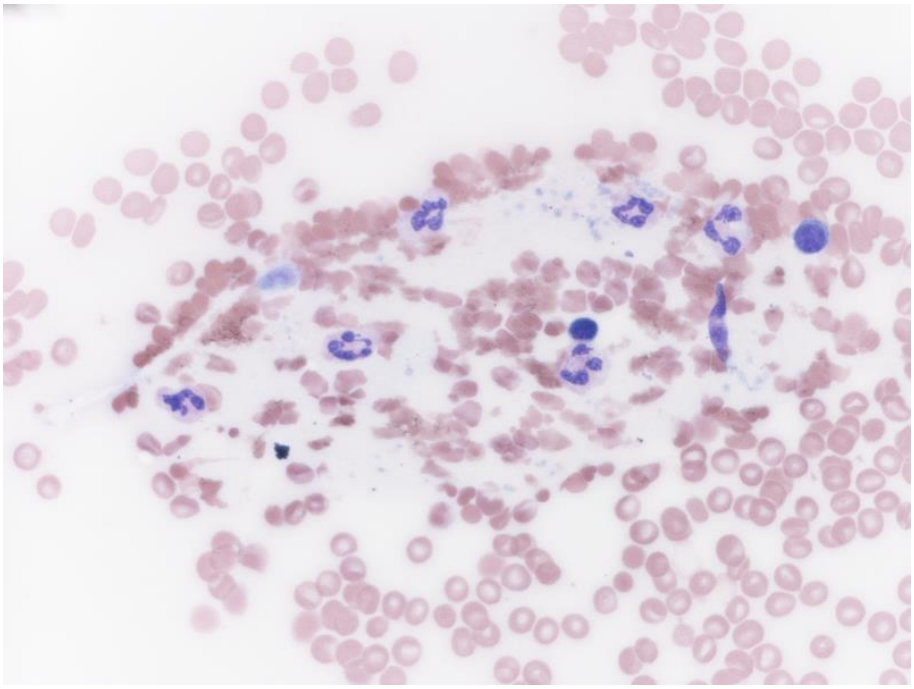


Figure 16 Image of TBSRTC I (smear, H&E staining)

TBSRTC II – benign

Representative samples showing only benign cells fall into this category (figure 17). Benign means that the specimen may or may not be diagnostic for disease but negative for malignancy, in fact in these samples there aren't malignant cells. This category includes benign follicular nodules, thyroiditis, and anything that is not suspicious of neoplasia. This category is the most frequent, as nodular goiter is the most common lesion sampled with FNA. The risk of malignancy of these specimens is obviously very low and falls in the 0-3 % range. Patients with a nodule TBSRTC II are not undergoing surgery, but follow-up is still recommended in these cases.

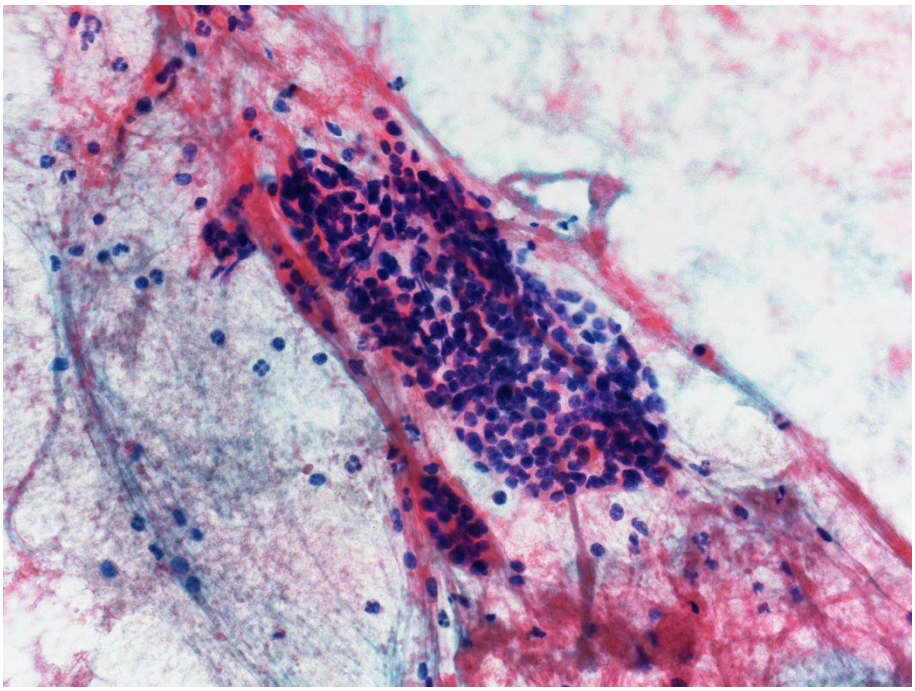


Figure 17 Image of TBSRTC II (smear, H&E staining)

TBSRTC III – slight atypia

In this category we find samples with atypia of undetermined significance or follicular lesion of undetermined significance (figure 18). This group is a challenging diagnostic category that includes a variety of specimens with mild cytological and/or architectural atypia that is insufficient to meet the criteria of TBSRTC IV, TBSRTC V or TBSRTC VI diagnostic categories. The degree of repeatability of TBSRTC III varies between institutions and ranges from 1% to 22% of all preoperative thyroid FNAs, which is a pretty low level. The separation of samples with mild degree of atypia from those with greater atypia is also justified by their degrees of malignancy. Minor atypia is associated with lower risk of malignancy and is around 10-30 %. For the patient management in this case, it is recommend repeating the FNA. Molecular analysis can be a valuable aid to diagnosis, especially multi-gene designed panel are useful for identifying malignancy marker [89].

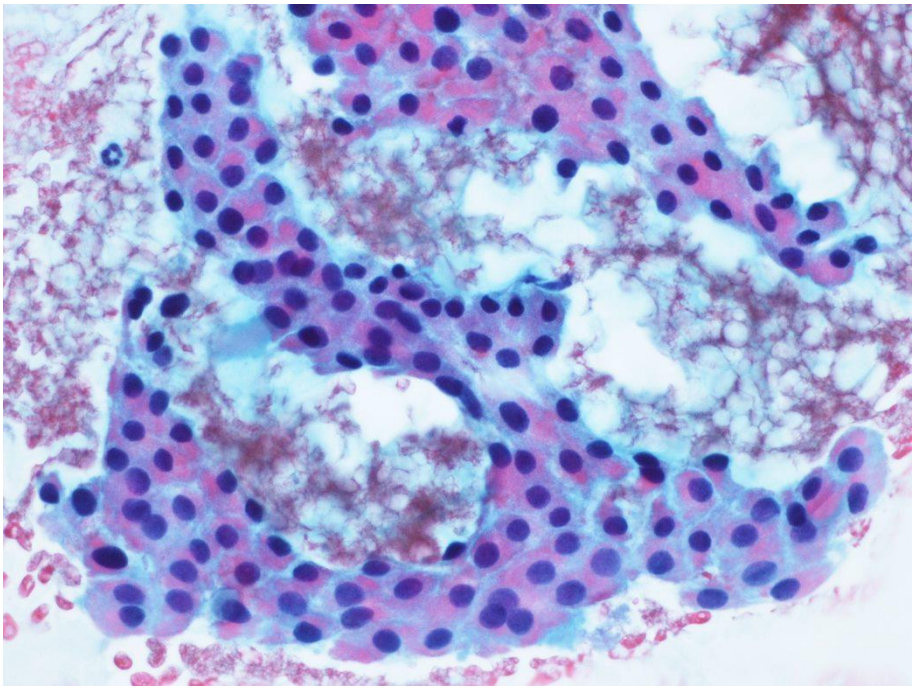


Figure 18 Image of TBSRTC III (smear, H&E staining)

TBSRTC IV – Suspicious for a Follicular Neoplasm

This category includes specimens that contain follicular cells, most of which are arranged in an altered architectural pattern or otherwise show more pronounced atypia (figure 19). Cases with suspected or definite nuclear features for papillary carcinoma should be classified as TBSRTC V or TBSRTC VI instead of falling into this group. In this category the risk of malignancy is raised to 25-40%. The patient should undergo "diagnostic" thyroid lobectomy, which is the removal of the thyroid lobe affected by the nodule.
In fact, only by histology analysis, performed in resected cases, can it be defined whether vascular and/or capsular invasion is present, which identifies the nodule as malignant and differentiates it from benign lesions including noninvasive follicular thyroid neoplasm with papillary-like nuclear features (NIFTP). However, molecular analysis of the sample is also recommended.

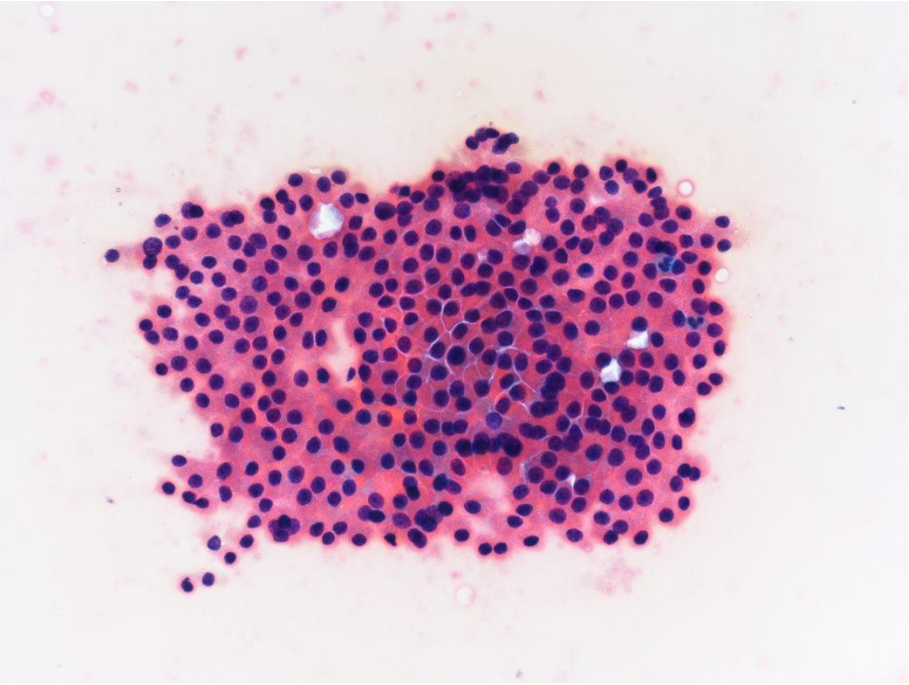


Figure 19 Image of TBSRTC III (smear, H&E staining)

TBSRTC V – suspicious for malignancy

This group includes specimens with cytomorphologic features that raise a strong suspicion for malignancy, but not sufficient to diagnose malignancy (figure 20). Cells with nuclear membrane irregularity, nuclear pallor, nuclear enlargement, nuclear grooves and so on may be present. Sometimes the uncertainty that leads to not diagnosing the nodule as malignant is given a specimen is suboptimal, an unusual variant of carcinoma or other. The risk of malignancy is around 50-70%. The patient will undergo total thyroidectomy or lobectomy not only to analyze by histological examination and molecular analysis the resected specimen, but also for treatment purpose.

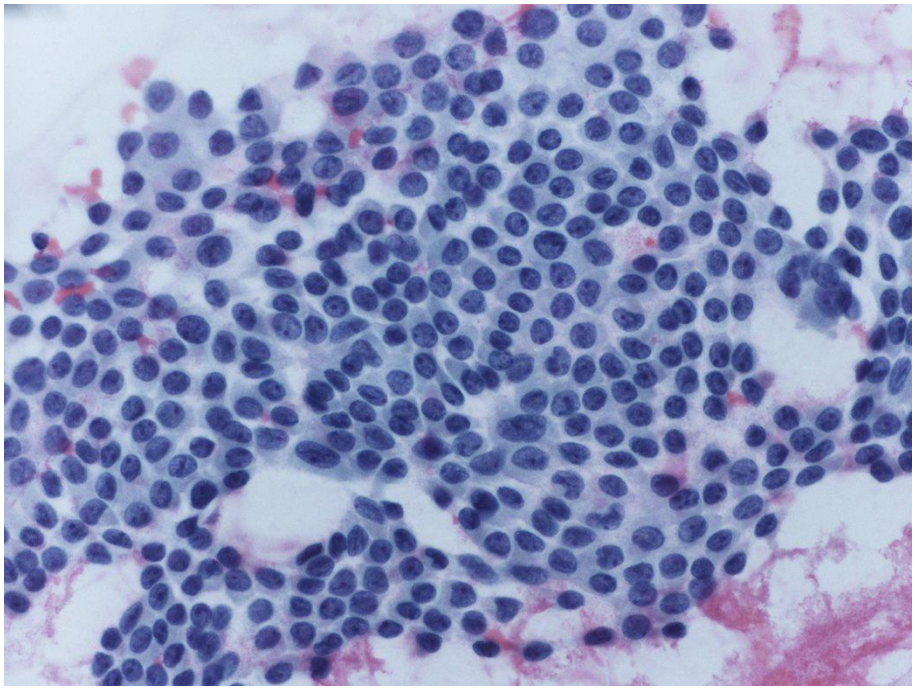


Figure 20 Image of TBSRTC V (smear, H&E staining)

TBSRTC VI – malignancy

This category includes specimens that clearly show features of malignancy, e.g., features of nuclear morphology that characterize papillary carcinoma, but cases of medullary, anaplastic, and other carcinomas also fall into this class (figure 21). It has a malignancy risk of 97-99% and as in TBSRTC V, lobectomy or thyroidectomy is not only diagnostic but has treatment purpose. Histological examination is performed to confirm the diagnosis and also molecular analysis.

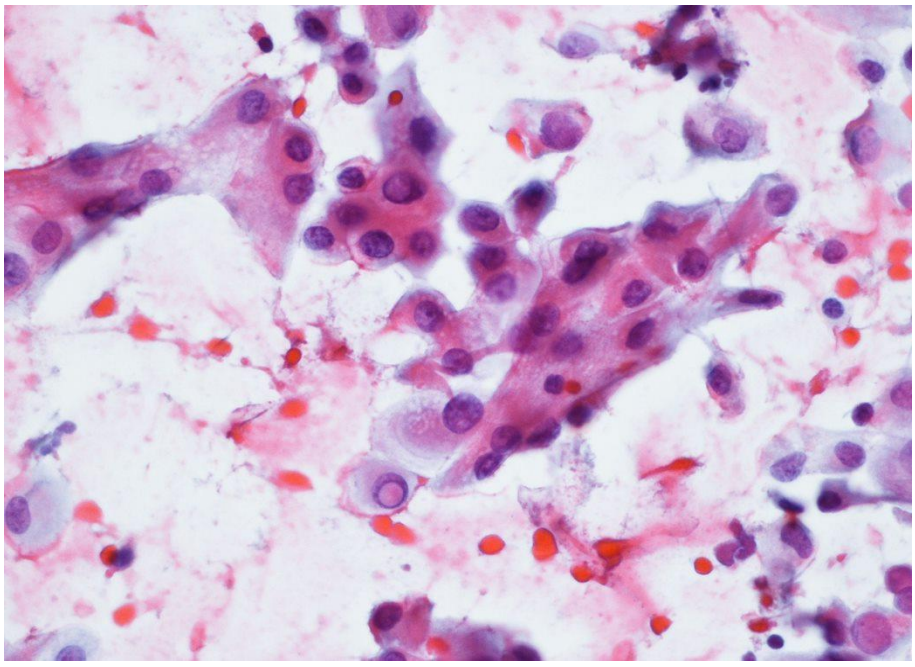


Figure 21 Image of TBSRTC VI (smear, H&E staining)

3. Aim of the project

Considering The Bethesda System for Reporting Thyroid Cytopathology, adequate cases that cannot be classified as benign, can be divided in two cases:

- cases for which surgery is not typically recommended and may be managed with a "wait and see" approach: TBSRTC III;
- cases typically require diagnostic and/or therapeutic surgery, lobectomy or total thyroidectomy: TBSRTC IV, TBSRTC V, and TBSRTC VI.

The purpose of this study is to develop and evaluate the performance of an artificial intelligence platform to assist pathologists' screening activities in distinguishing between these two groups.

To accomplish this, we have trained a Deep learning model capable of discriminating FNA specimens diagnosed as TBSRTC III from those diagnosed as TBSRTC IV, TBSRTC V, and TBSRTC VI.

Considering the literature this study appears to be the first to precisely evaluate AI's capacity to distinguish between these preoperative FNA diagnostic categories, which have an immediate influence on patient clinical treatment.

4. Materials and Methods

4.1 Specimens

For this study we collected images of thyroid FNA smears from the archives of the Endocrine Pathology program of the Department of Pathology of the University of Bologna Medical center-IRCCS Azienda Ospedaliero-Universitaria di Bologna, Bologna, Italy, sequentially acquired from 2011 to 2020. The image inclusion criteria we set for this study included only appropriate images derived from PAP-stained thyroid FNA smears that had been diagnostic TBSRTC III, TBSRTC IV, TBSRTC V, and TBSRTC VI according to The Bethesda System for Reporting Thyroid Cytopathology.

4.2 Acquisition of images and annotations

The professor Giovanni Tallini, an expert pathologist with interested in thyroid pathology and cytopathology, selected specific areas of the thyroid smears to examine using a bright-field microscope (BX51, Olympus Corporation, Tokyo, Japan). This selection process is critical to exclude artifacts such as clots and dye precipitates and identify only regions of interest-that is, regions useful for diagnostic purposes (Figure 22). After the selection of diagnostic fields, the next step was image acquisition by photography performed with a digital camera (Digital Sight DS-L1, Nikon Corporation, Japan) at 400X magnification (equivalent to a field diameter of 0.55 mm and a field area of ~0.237 mm2). The images generated were saved in .JPG format. Then a database was created where each image was systematically annotated by associating among other information the cytological diagnosis of each case.

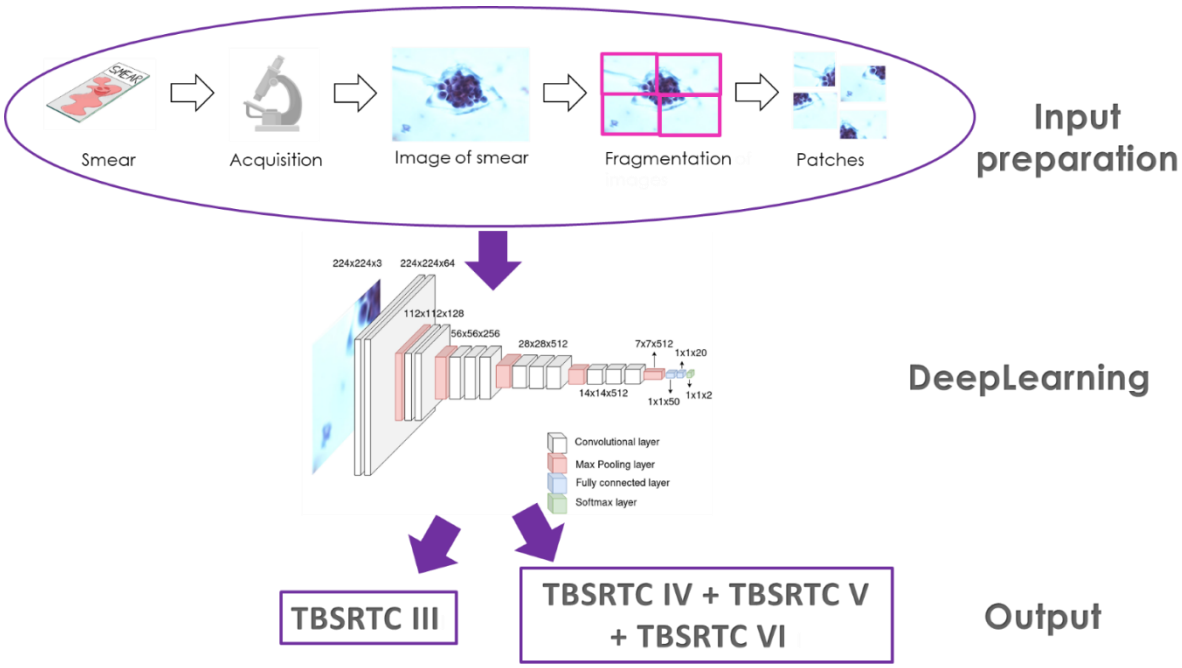


Figure 22 Study workflow. Images from the fine needle aspiration smear are acquired using a microscope and a digital camera, and then the images are divided into patches to prepare the input for the deep learning model. After processing each patch, the deep learning model outputs one of the two results listed below: TBSRTC IV+V+VI or TBSRTC III

4.3 Dataset creation

The dataset was made up of RGB photographs compressed into JPEG files, which were then subjected to the necessary preprocessing to be used as input for deep learning networks.
There are 381 images included in the dataset and they were subsequently divided into two categories according to their diagnosis and relative surgical implication (Table 4):

- TBSRTC III, surgery is usually not required
- TBSRTC IV+V+VI, surgery usually required.

Then several of patches of size 224x224 pixels as square tessellations of the original RGB image were extrapolated from each image (Figure 22). The total number of patches obtained was 9667 (table 4).

	TBSRTC III	TBSRTC IV + V + VI	Total
n° of images	141	240	381
n° of patches	3528	6139	9667

Table 4 The study's number of pictures and patches by diagnostic classifications

4.4 Model training

To address the project task, which involved a binary classification problem, pre-trained ResNet-50 neural network neural network was chosen by the engineers [90]. Transfer learning, a machine learning technique that entails retraining a piece of an existing neural network model using a different domain of analysis, was used to do this [91]. To do transfer learning they used a pre-trained ResNet-50 neural network network trained on ImageNet, a dataset comprising more than 14 million images from 1000 diverse classes [92].
They then used their particular dataset to retrain the network's last layer solely (Figure 23). They used repeated cross-validation (CV) as a validation approach to make sure a robust assessment of their model's performance [93]. The dataset was divided into K folds, with the first K-1 folds acting as the training set and the last fold acting as the validation set. The model's performance was averaged over all of the training instances after this process was performed for each fold. Ten folds were chosen for our investigation because they strike a good bias-trade-off [94]. To enable varied data splits within the folds, they conducted the CV method five times using different random seeds. To maintain the same number of patients in each fold, the K fold partition was stratified.
The Tensorflow framework was used to implement the neural network's training [95]. The model was trained for 25 epochs with early halting for each fold and repetition, using a patience parameter of 15 epochs as a regularization strategy to avoid overfitting [[95]. A single NVIDIA TITAN X12GB GPU was used for the training process, which required a total of about 10 hours of calculation time. We used Receiver operating curves (ROC) and calculated the area under the curve (AUC) for each fold and repetition to evaluate the classification performance.
The mean of the AUC values collected from each fold served as a measure of a repetition's performance. By computing the mean and standard deviation of the AUC values across all repeats, the experiment's overall performance was assessed.

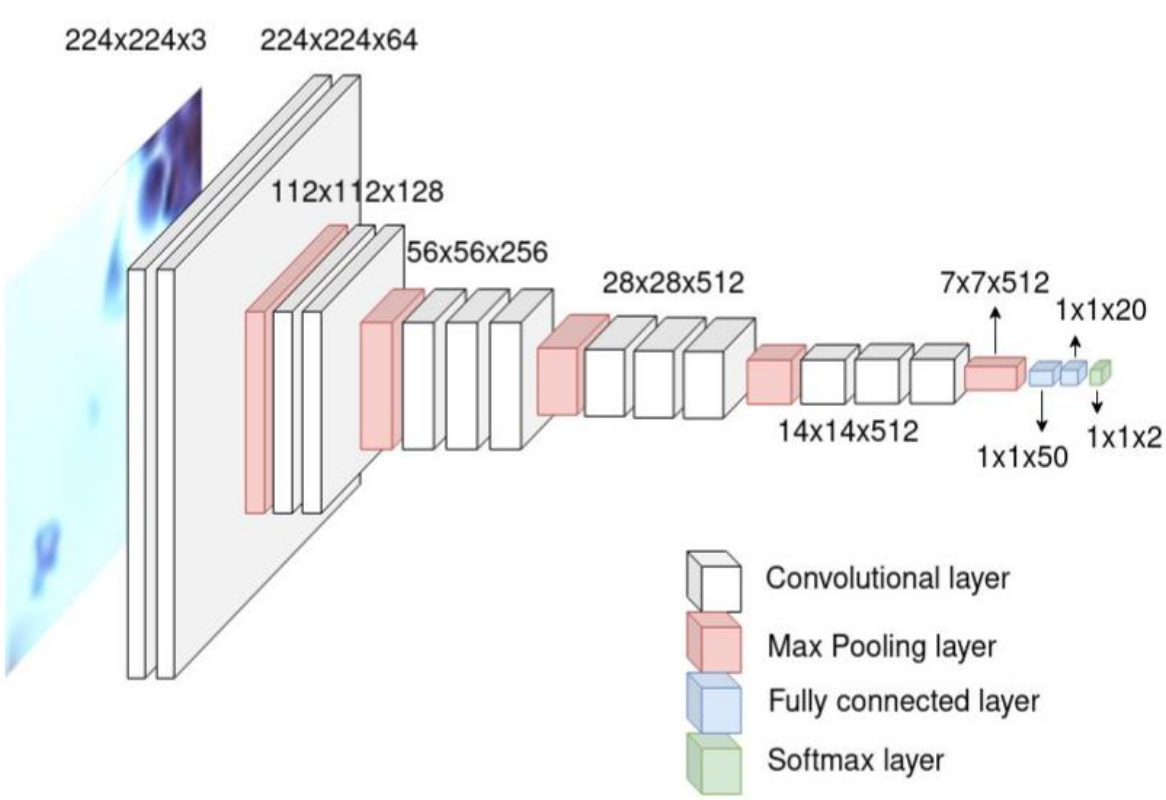
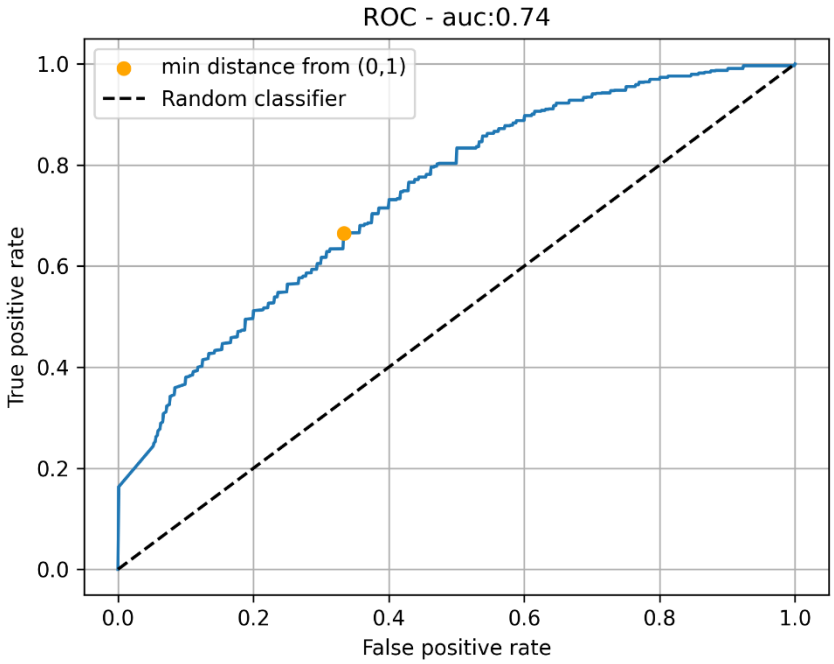


Figure 23 The ResNet-50-based neural network has 48 convolutional layers, one MaxPool layer, and one average pool layer in its architecture. The input image size is set to ResNet-50's standard 224x224x3 measurements. Every pooling layer cuts the size of the feature map in half. In this study, we concentrated on optimizing each neural network layer. Our study only needed two classes, TBSRTC III and TBSRTC IV+V+VI, compared to the original ResNet-50 architecture's final layer, which has 1,000 output classes. As a result, we changed the output channel number from 1,000 to 2 in the final layer (the green layer in the figure).

5. Results

The ROC curve of patches classification, presented in graphic 3, was used to evaluate the performance of deep learning. The cytologic classification task, TBSRTC III (diagnostic category not frequently followed by surgery) vs. TBSRTC IV- V- VI (diagnostic categories typically leading to thyroid resection), showed a favorable discriminative performance with an AUC of 0.74 0.08 (mean standard deviation). The operating point on the ROC curve, which represents the smallest separation from the ideal classifier at coordinates (0,1), produced a specificity and sensitivity of 0.667 and 0.665, respectively.



Graphic 3 Average ROC curve of the model developed using a 10-fold CV repeated five times for the classification problem. According to calculations, the area under the ROC curve is 0.74 0.08 (mean standard deviation), indicating a good performance in terms of discrimination. The operational point on the ROC curve with the smallest deviation from the ideal classifier is indicated in orange at coordinates (0.333, 0.665), which is where it is placed at coordinates (0,1).

6. Discussion

The aim of this study is to offer valuable assistance to pathologists in their clinical practice by introducing an effective AI-based method for analyzing thyroid cytological smears. Our deep learning model was used to analyzed a large dataset made up of 381 thyroid FNA smear images in an effort to distinguish between diagnostic categories that have significant therapeutic implications: TBSRTC IV, V, and VI, which are frequently followed by thyroid resection, TBSRTC III does not typically need thyroid surgery. The performance of our AI model was demonstrated by the ROC analysis, which produced an AUC of 0.74 ± 0.08.

Previous research has looked into the potential use of AI systems in digital pathology to assist pathologists in their work [96-98]. In order to distinguish between benign and malignant thyroid lesions in 51 FNA smears, Karakitsos et al. used an artificial neural network (ANN)[96]. Ippolito et al. used an ANN model to analyze 453 thyroid nodules identified as ambiguous during FNA and divided the patients into high-risk and low-risk groups according to predetermined standards such enlarged cell nuclei, altered nucleoli, and follicular architecture. Sensitivity of 0.857 and specificity of 0.588 were attained[97].

After the advent of deep learning, recent studies have explored its applicability to thyroid FNA. In order to distinguish papillary thyroid carcinoma (PTC) from benign thyroid nodules, Guan et al. used VGG-16 and Inception-v3 deep convolutional neural network (DCNN) models. He reported specificity of 94,91% for VGG-16 and 86,44% for Inception-v3 and sensitivity of 100% and 98,55 respectively [98]. Another research group that used DeepL was the one led by Hirokawa, who tried to categorize 393 thyroid FNA smears according to their histological diagnosis. To do this they used deep learning with EfficientNetV2-L. For practically all histology groups, the precision recall area under the curve was greater than 0.95. However, differentiating poorly differentiated thyroid carcinomas from other forms proved difficult, and medullary thyroid carcinoma may be misclassified as benign [99].

For our investigation, we selected 400X PAP-stained images as input because they are commonly used and compatible with laboratory operations. The choice of selecting images with a given coloration (PAP) and magnification (400X) is due to the fact that training a network of DeepL with highly variable images in the dataset is challenging and can reduce the statistical performance of the model. To seek the uniformity we needed, samples with different coloration and magnification were eliminated. We therefore opted for a "homogeneous" selection of images, although in principle it is possible to implement systems that accept a variety of image magnifications and cytological staining.

The application of deep learning in pathology involves a number of caveats. Most of these have to do with the high resolution of whole slide images. To partially overcome this problem, for training the deep learning model, we used a patch-based strategy, similar to that used by other researchers [100, 101]. We improved the model's performance by breaking the image into patches, which supplied less complex inputs for training the network. This approach takes advantage of the wealth of information contained in whole slide images of all cytology smears, but having smaller and more manageable inputs.

As previously stated, our investigation achieved a ROC AUC value of (0.791 ± 0.004), as well as a sensitivity of 0.740 and a specificity of 0.706.

Teramoto et al. used a similar approach to classify lung cytology pictures as adenocarcinoma, squamous cell carcinoma, or small cell carcinoma and they obtain AUC of 0.733 for original images and an improved AUC of 0.886 for images subjected to data augmentation [102]. Awan et al. explored urine cytology using two independent deep learning-based techniques (RetinaNet with ResNet followed by cell segmentation) to stratify individuals into low and high bladder cancer risk groups. The authors obtained an AUC of 0.81 when only malignant cell counts were considered, and an AUC of 0.83 when both malignant and atypical cell counts were considered [103].

Given the worldwide shortage of pathologists, particularly in low-income, developing nations, the need for diagnostic support for routine histopathology and cytopathology, particularly for cytology screening programs, is becoming an essential concern[87].

Following our study's goal, we divided samples into two categories: TBSRTC III and TBSRTC IV-V-VI. We chose these two classifications primarily because of their relevance for clinical care of patients following preoperative FNA diagnosis: TBSRTC III instances seldom require surgery, whereas TBSRTC IV-V-VI cases commonly necessitate thyroid excision. Our deep learning model's performance in distinguishing TBSRTC III from TBSRTC IV-V-VI is comparable to that of an expert pathologist routinely generating preoperative FNA diagnosis.

For future developments, the neural network could be implemented in such a way that it is capable of discriminating each diagnostic class and also introduce any diagnosis made on histology in order to further improve the specificity/sensitivity of the method.

Cap 3 – Spatial transcriptomic applied to thyroid carcinoma

Spatial transcriptomics, the method chosen for this study, has been elected as method of the year in 2020 by Nature. It is currently little used in Italy, while at the Gustave Roussy Cancer Institute it has been used for several years and can therefore offer good expertise. This study is the result of a valuable collaboration with the Gustave Roussy Institute, which hosted me for three months and accompanied me during the development of the project.

We chose to use this method because we had a will to examine gene expression in thyroid carcinomas in more detail, paying particular attention to genes that may be involved in tumor progression. In fact, what we want to investigate are the transcription gene profiles responsible for tumor proliferation and invasion, also turning a look at the invasion patterns that characterize papillary and follicular thyroid carcinomas. Indeed, these two histotypes of thyroid carcinoma show different patterns of invasion: papillary carcinoma tends to invade the lymph nodes, while follicular carcinoma tends to invade the capsule and/or blood vessels.

The goal of the project is to use spatial transcriptomics to better understand the mechanisms of invasion and progression of thyroid carcinoma.

1. Spatial transcriptomic

To better understand the complex world of biology, it is necessary to study not only the characteristics of individual components, but also the interaction between them. To do this, the spatial organization of tissues and organisms must be taken into account. Research in this area began many years ago to develop methods of spatial transcriptomics. "Spatial transcriptomics" refers to the analysis of mRNA expression of a large number of genes by considering the spatial context of cells and tissues.
Since the 1970s, the techniques that have followed over the years are manifold (figure 24), as described in a very detailed review by Lambda Moses and Lior Pachter published in Nature [104]. Typically, the origins of spatial transcriptomics can be related to laser capture microdissection (LCM) in the late 1990s, but in reality, the precise start date can be set in 1969 when, radioactive in situ hybridization (ISH) was introduced to visualize DNA [105] and ribosomal RNA [106] in *Xenopus laevis* oocytes, and in 1973, it began to be used to visualize transcription of specific genes, starting with globin [107]. Subsequently there was the advent of non-radioactive fluorescent or colorimetric ISH [108], with its subsequent development of this method on the whole slide (whole-mount ISH), which was widely used at the turn of the new millennium [109]. During this time, model organism databases were also created, like the Gene Expression Database (GXD) 30 and Zebrafish Information Network, which collect proliferating gene expression patterns from many sources.
Other methods have been developed over the years such as Enhancer trap, Gene Trap, Reporter in situ, and sequential FISH.
Obviously, Spatial transcriptomics has been applied in various fields of research such as neuroscience, developmental biology, and cancer biology. In cancer, spatial transcriptomics takes hold when the importance of the tumor microenvironment is recognized; in fact, through these technologies, the relationships present between neoplastic tissue with adjacent normal tissue, stroma, and immune system cells can be studied deeply [110].
Considering the method of acquiring spatial information, the technologies that are used to date fall into 5 categories:

- region of interest (ROI) selection,
- single-molecule FISH (smFISH),
- in situ sequencing (ISS),
- next-generation sequencing (NGS) with spatial barcoding,
- methods without a priori spatial locations.

We adopted for our study a next-generation sequencing technology with spatial barcoding: Visium Spatial Gene Expression by 10X Genomics.

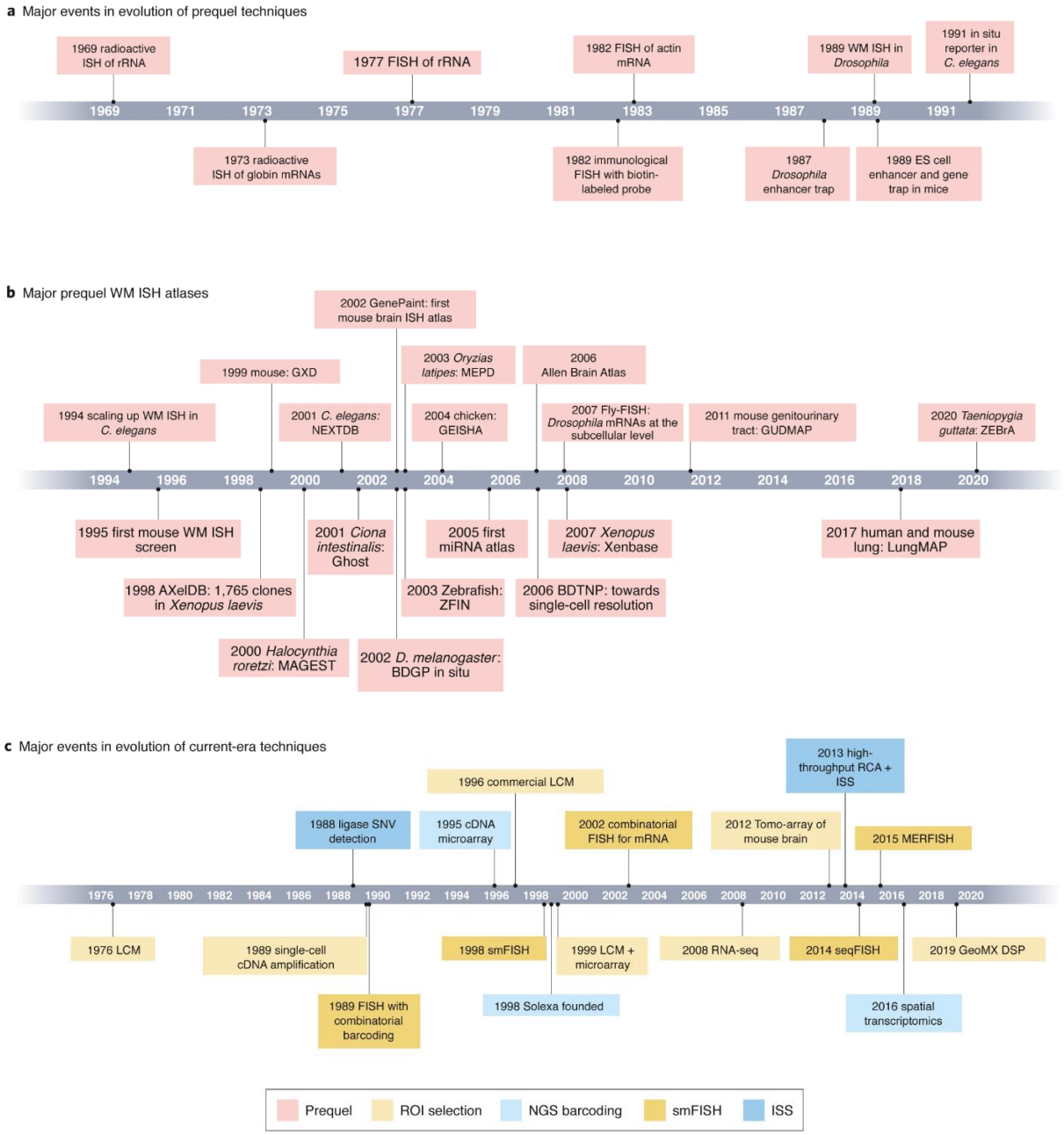


Figure 24 Timw line of technologies that have followed one another over the years. [Lambda, M. 2022]

2. Invasion pattern of well-differentiated thyroid carcinoma

Thyroid cancer encompasses various subtypes, two of the most prevalent being Follicular Thyroid Carcinoma (FTC) and Papillary Thyroid Carcinoma (PTC). These subtypes exhibit distinct invasion patterns, with FTC commonly involving capsule and vessel invasion, while PTC tends to display lymphatic vessel invasion [111]. Understanding these invasion patterns is crucial for diagnosis, treatment planning, and predicting the clinical course of these cancers.

Follicular Thyroid Carcinoma originates from the follicular cells of the thyroid gland and is known for its tendency for:

- Capsule Invasion (figure 25): FTC often infiltrates the thyroid capsule, a fibrous structure that surrounds the thyroid gland. The capsule invasion is typically characterized by tumor cells penetrating the capsule's outer layer, allowing the cancer to extend into surrounding tissues. This invasion can make complete surgical resection more complex, as surgeons must carefully navigate the boundaries of the thyroid gland to remove all cancerous tissue[112].

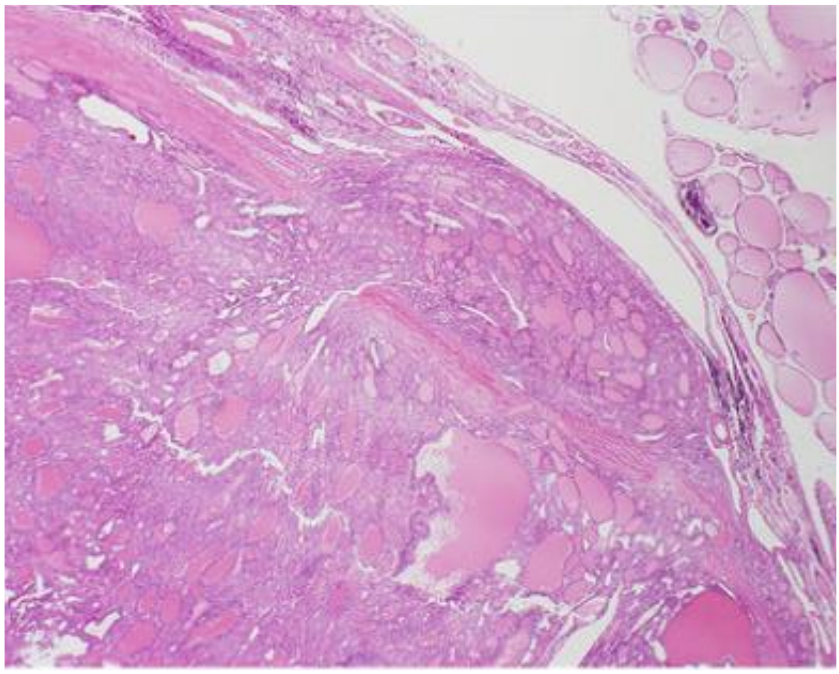


Figure 25 Follicular thyroid carcinoma invading the fibrous capsule. (H&E staining) [Kakudo, K. 2018]

- Vessel Invasion (figure 26): one of the defining features of FTC is its propensity to invade blood vessels. Tumor cells infiltrate the walls of blood vessels within the thyroid, leading to intravascular tumor growth. This behavior can result in tumor embolism, where cancer cells break free and enter the bloodstream, facilitating metastasis to distant organs such as the lungs and bones. Vessel invasion is a key factor in the aggressiveness of FTC and often necessitates additional treatments like radioactive iodine therapy or external beam radiation [112]
.

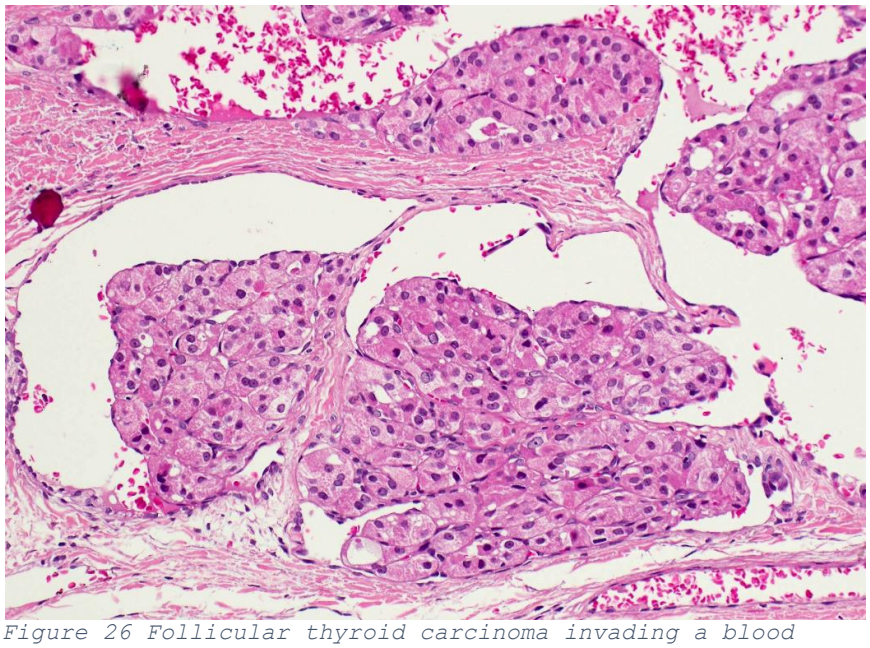


Figure 26 Follicular thyroid carcinoma invading a blood vessel (H&E). Available on: https://www.pathologyoutlines.com/topic/thyroidfollicular.html (Accessed on 25 May 2023)

In contrast to FTC, PTC exhibits a different invasion pattern, primarily involving lymphatic vessel invasion (figure 27). While PTC is the most common type of thyroid cancer and is generally associated with a better prognosis than FTC, its invasion into lymphatic vessels carries its own set of clinical implications [111].

PTC frequently spreads to nearby lymph nodes through the lymphatic vessels in the neck. This lymphatic vessel invasion is a hallmark of PTC and is a key aspect of the diagnostic criteria for the disease. The presence of cancer cells within lymphatic vessels and regional lymph nodes is a critical factor in staging PTC and determining the extent of the disease. Lymphatic vessel invasion often results in the spread of PTC to cervical lymph nodes, particularly those in the central and lateral compartments of the neck. The extent of lymph node involvement plays a significant role in treatment planning and prognosis assessment [113].

Unlike FTC, PTC typically does not invade the thyroid capsule or blood vessels to the same extent. This relatively limited local invasion is one of the reasons why PTC generally carries a more favorable prognosis compared to FTC [113].

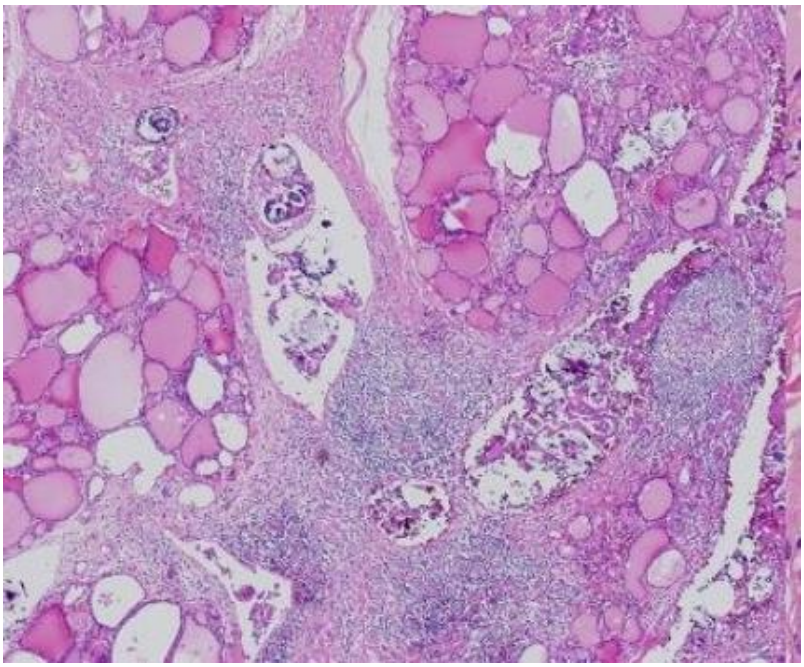


Figure 27 Lymphatic invasion in PTC (diffuse sclerosing variant). (H&E stain, x4 magnification) [Kakudo, K. 2018].

An important histological subtype is follicular variant papillary thyroid carcinoma, which in turn, depending on the type of invasion pattern by which it is characterized is divided into Invasive Follicular Variant of Papillary Thyroid Carcinoma (IFVPTC) and Encapsulated Follicular Variant of Papillary Thyroid Carcinoma (EFVPTC) [114].
IFVPTC is characterized by the presence of invasive features, meaning that it exhibits the capability to infiltrate the surrounding thyroid tissue or even extend into blood vessels or nearby structures (figure 28). This invasive behavior can make IFVPTC more aggressive and potentially challenging to treat. It may also carry a higher risk of recurrence and metastasis. On the other hand, EFVPTC is characterized by a more contained growth pattern. The tumor cells in EFVPTC form a well-defined capsule, which encases the malignant tissue (figure 29). This encapsulated nature makes EFVPTC less likely to invade surrounding tissues and structures, typically resulting in a more favorable prognosis. To diagnose EFVPTC, however, there must be invasion of the capsule otherwise it is called non-invasive follicular thyroid neoplasm with papillary-like nuclear features (NIFTP) [114].
Both IFVPTC and EFVPTC may have overlapping microscopic features, making accurate diagnosis critical. Distinguishing between the two is essential for treatment planning and prognostication. Treatment for both subtypes typically involves surgical removal of the affected thyroid tissue, and additional therapies such as radioactive iodine (RAI) may be considered based on the specific characteristics of the tumor and the patient's individual risk factors. In summary, the key difference lies in the invasive nature of IFVPTC versus the encapsulated, less invasive nature of EFVPTC, which significantly impacts their clinical management and outcomes.

In summary, the invasion patterns of FTC and PTC are distinct and have important clinical implications. FTC tends to invade the thyroid capsule and blood vessels, contributing to its aggressiveness and potential for distant metastasis. In contrast, PTC primarily exhibits lymphatic vessel invasion, leading to regional lymph node involvement, which is a crucial factor in staging and treatment planning. Understanding these invasion patterns is essential for clinicians to make informed decisions regarding diagnosis, treatment, and prognosis in patients with thyroid cancer, ultimately improving patient outcomes and quality of life.

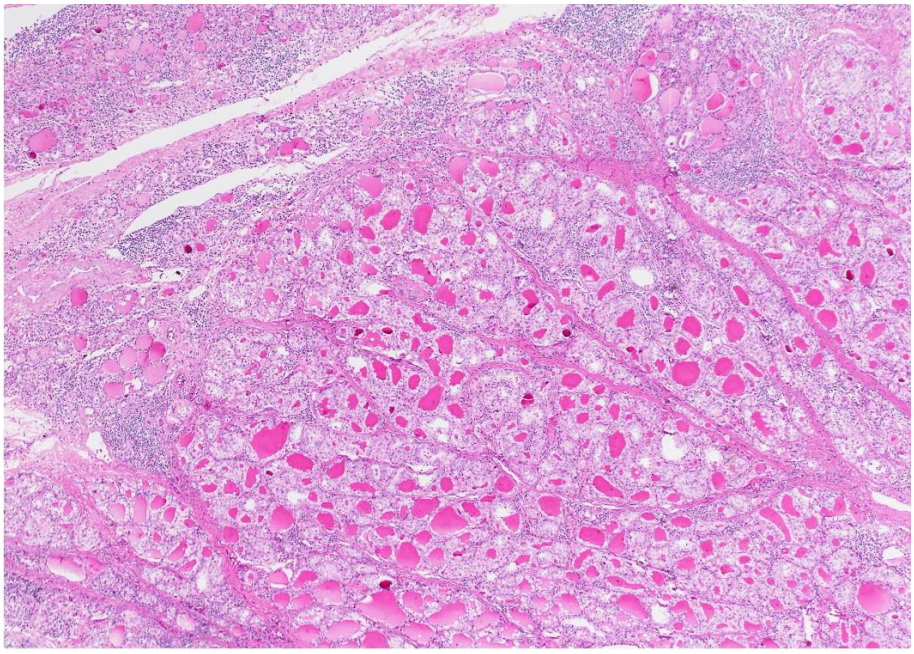


Figure 28 Infiltrative follicular variant Follicular Variant of Papillary Thyroid Carcinoma (H&E). Available on https://www.pathologyoutlines.com/topic/thyroidfollicularvariant.html (Accessed on 25 May 2023)

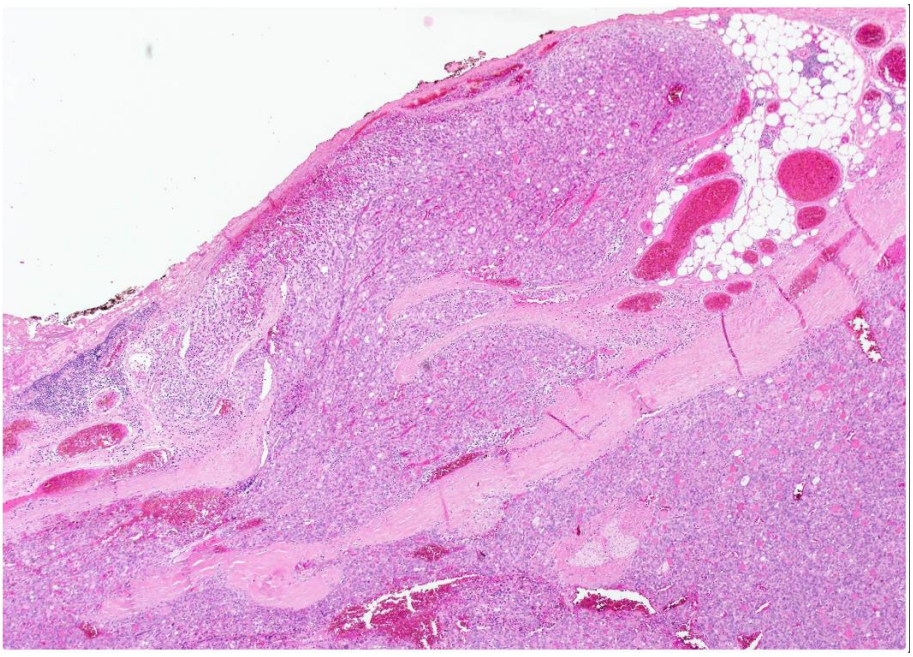


Figure 29 Encapsulated Follicular Variant of Papillary Thyroid Carcinoma: invasion of capsule is seen in the image (H&E). Available on https://www.pathologyoutlines.com/topic/thyroidencapsulatedfollicularvariant.html (Accessed on 25 May 2023)

3. Aim of the project

The goal of the project is to use spatial transcriptomics to better understand the differences of two types of well-differentiated thyroid carcinomas: follicular and papillary. To better study the correlations present, we also wanted to include an intermediate subtype in the study: infiltrative follicular variant papillary thyroid carcinoma (IFVPTC).
The differences we want to investigate in these carcinomas are mostly related to their different patterns of invasion, as papillary carcinomas tend to invade lymphatic vessels, while follicular carcinomas tend to invade the capsule and blood vessels. IFVPTCs have an invasion pattern more similar to PTCs, in fact they tend to invade lymphatic vessels.
The project will focus on the analysis by spatial transcriptomics, through the use of Visium 10X Genomics, which allows transcriptomic data integration with histology data.

4. Materials and Methods

4.1 Case Selection and area of interest

For Spatial Transcriptomic, 3 formalin-fixed paraffin embedded (FFPE) samples of well-differentiated thyroid human carcinoma (Case 1-3) were selected (table 5).
Case 1 is a follicular thyroid carcinoma of an 80-year-old woman. The carcinoma was in the right lobe, with a size greater than 2,5 cm, has a portion with invasion of the fibrous capsule and a portion with invasion of a blood vessel. The patient at diagnosis had lymph node metastasis.
Case 2 is an infiltrative follicular variant of papillary thyroid carcinoma of a 50-year-old woman. The carcinoma was located in the right lobe and isthmus, with a size greater than 0,8 cm, has a portion with lymphatic invasion. The patient at diagnosis had lymph node metastasis.
Case 3 is a tall cell papillary thyroid carcinoma of a 44-year-old woman. The carcinoma was located in the right lobe, with a size greater than 3,5 cm, has a portion with lymphatic invasion. The patient at diagnosis had lymph node metastasis.

	Male/ Female	Age at diagnosis	Type of carcinoma	Size (cm)	Thyroid lobe	Invasive front	Extrathyroidal invasion	Metastases
Case 1	F	80	FTC	2,5	Right	VI and CI	no*	Lymph node metastasis
Case 2	F	50	IFVPTC	0,8	Right/ isthmus	lymphatic invasion	yes	Lymph node metastasis
Case 3	F	44	PTC	3,5	Right	lymphatic invasion	yes	Lymph node metastasis

Table 5 Characteristics of the cases included in the study
**Iatrogenic capsule rupture*

For the spatial transcriptomics of these three cases, we selected 4 areas of interest (figure 30,31,32).
Case 1 had two areas of interest Sample 1 and sample 2. Sample 1 near capsule invasion and Sample 2 near blood vessel invasion. An area of interest near lymphatic vessel invasion was also selected in Cases 2 and 3: Sample 3, and Sample 4 respectively.

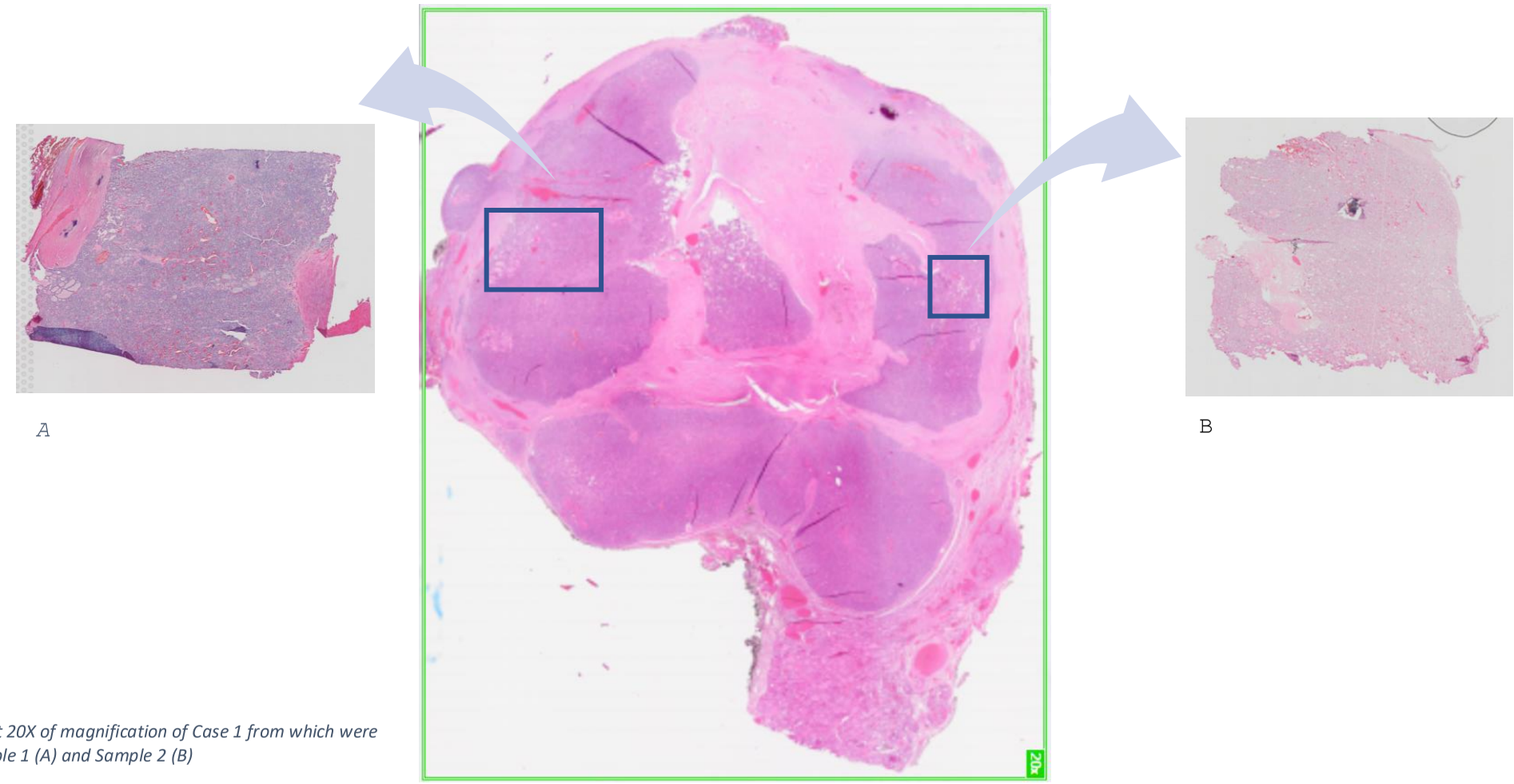


Figure 30 H&E at 20X of magnification of Case 1 from which were selected at Sample 1 (A) and Sample 2 (B)

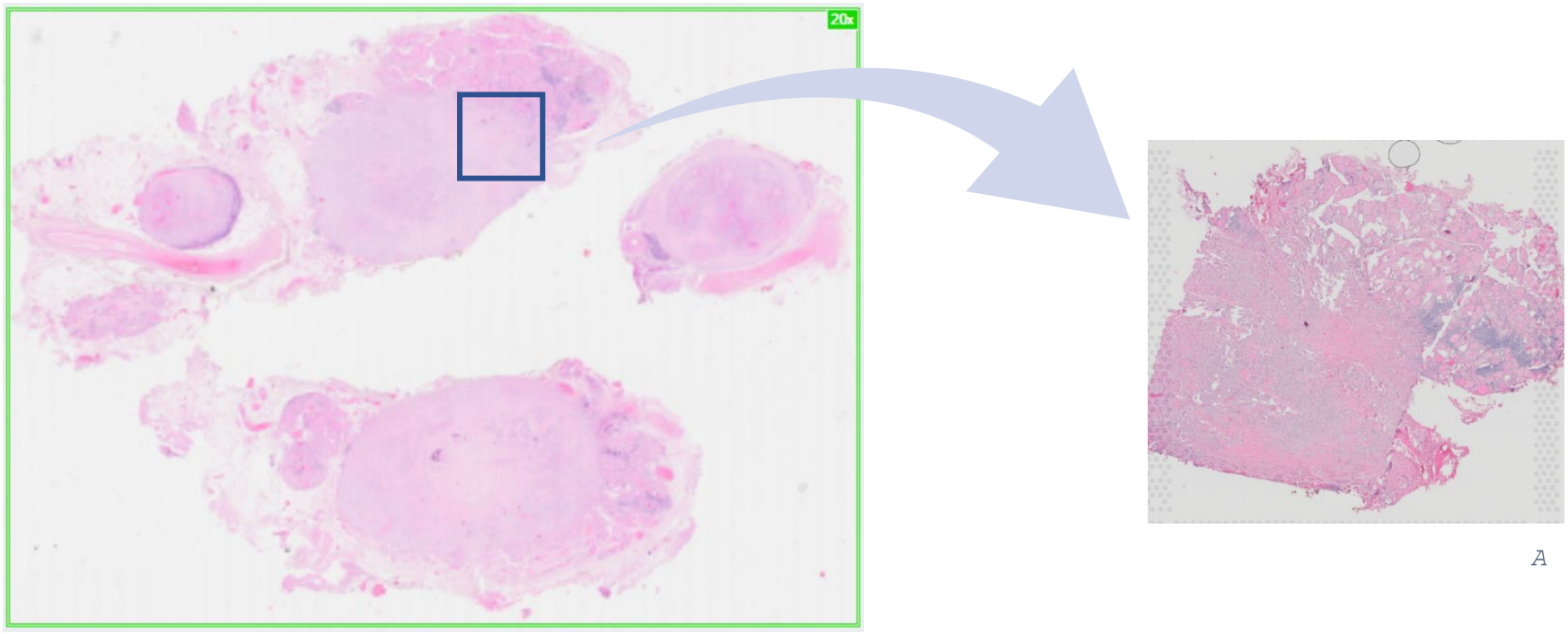
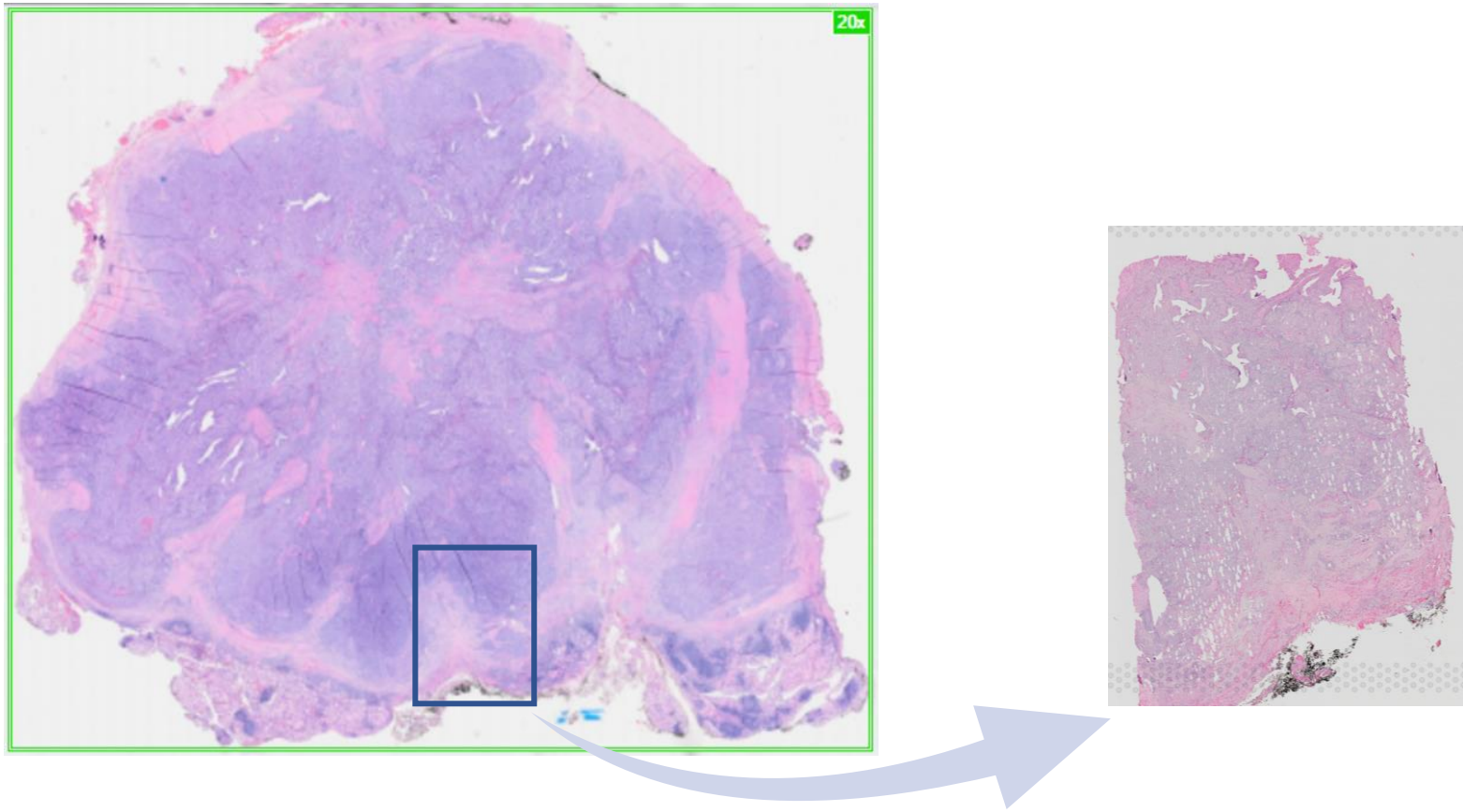


Figure 31 H&E at 20X of magnification of Case 2 from which were selected at Sample 3 (A)



A

Figure 32 H&E at 20X of magnification of Case 3 from which were selected at Sample 4 (A)

4.2 Next Generation sequencing

DNA from FFPE specimens was extracted starting from 2 to 4 consecutive sections of 10 µm, using the QuickExtract FFPE DNA extraction kit (LGC Biosearch Technologies, Berlin, Germany), taking into account the areas of interest identified on the control slide stained with hematoxylin and eosin (H&E) by a pathologist. DNA was quantified using Qubit dsDNA BR assay kit (Thermo Fisher Scientific, Waltham, MA, USA). The next generation sequencing (NGS) was performed using a multi-gene panel developed in the Molecular Pathology Laboratory of IRCCS Policlinico di S.Orsola [89]. This panel allows the analysis of the entire coding regions (CDS) or hot-spot regions of 28 genes, for a total of 330 amplicons (88.73 kb, human reference sequence hg19/GRCh37, 21.77 kb). The genes included in the panel are *BRAF*, c-KIT, CTNNB1, DICER1, DPYD, EGFR, EIF1AX, GNA11, GNAQ, GNAS, H3F3A, H*RAS*, IDH1, IDH2, *KRAS*, MED12, MET, N*RAS*, PDGFRα, PIK3CA, PTEN, RET, RNF43, SMAD4, TERT, TP53, TSHR, VHL.
The Gene Studio S5 Prime sequencer (ThermoFisher Scientific Inc) was used for NGS. In brief, approximately 30ng of DNA were used per panel for amplicon library preparation using the AmpliSeq Plus Library Kit 2.0. An Ion Chef Machine was used to prepare templates, and an Ion 530 chip was utilized to sequence them. Sequences were analyzed with the IonReporter tool (v. 5.18 – Thermo Fisher Scientific). According to the previously reported publication [89],for mutational calls, only nucleotide changes observed in both strands and at least 5% of the total number of reads evaluated were included. The Varsome tool (https://varsome.com/, accessed in April 2023) [115] was used to evaluate the pathogenicity of each variant.

4.3 Visium spatial gene expression

The protocol we applied, using Visium spatial gene expression by 10X Genomics, involved several steps (figure 33).

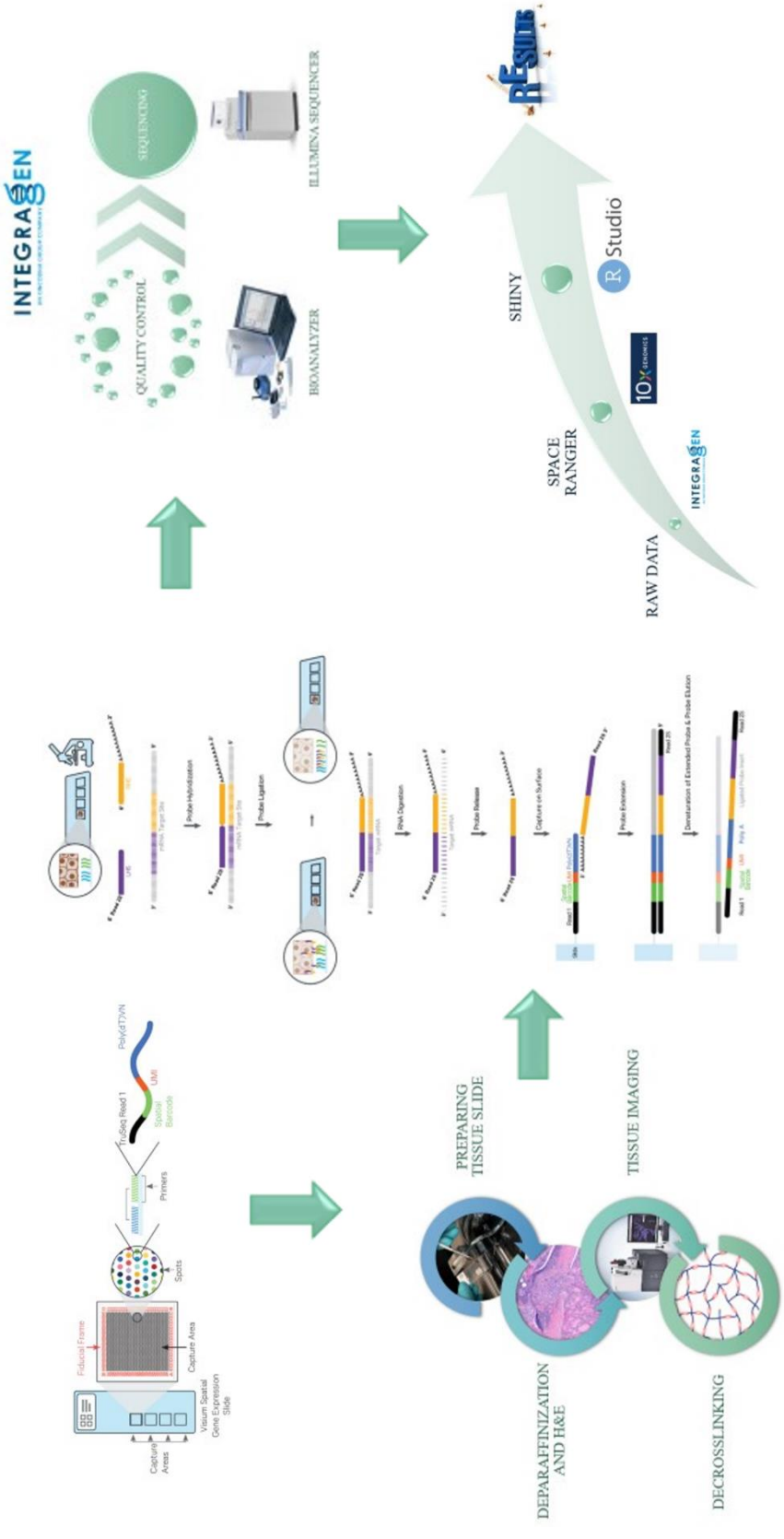


Figure 33 Overview workflow applied for the study

4.3.1 Visium Spatial Gene Expression Slide

The kit "Visium Spatial Gene Expression for FFPE assays RNA level" involves the use of "Visium Spatial Gene Expression Slide". This is a slide with 4 capture areas, where our tissue will be placed. Therefore, with a single slide, we can analyze four sections (figure 34). The capture area dimensions are 6,5 x 6,5 cm, thus the regions of interest of our samples must fall within these limits. The Capture area is surrounded by a Fiducial frame for a total area of 8 x 8 cm. This frame aids in the placing of the tissue on the slide, but should never be covered. Each capture area contains the spots arranged in a hexagonal grid, 100 µm apart from center to center and 55 µm in diameter. Every spot contains primers, required to capture the probes, to detect the expression of about 5'000 genes. Every primer is composed by:

- Illumina TruSeq Read 1(partial read 1 sequencing primer,
- 16 nucleotides Spatial Barcode (all primers in a specific spot share the same Spatial Barcode),
- 12 nucleotides of unique molecular identifier (UMI),
- 30 nucleotides poly(dT) sequence (captures ligation product).

Visium Slide spot's Poly(dT) sequence catches the kit probes' Poly(dA) sequence. The probe panel to be used consists of probes that bind to the entire transcriptome of the samples based on what is expressed, mapping the sites where gene activity occurs.

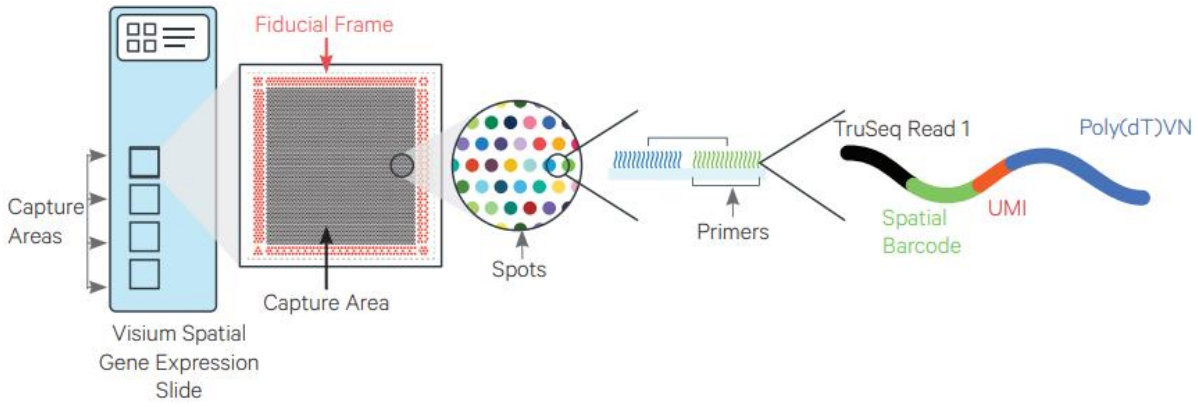


Figure 34 Representation of Visium Spatial Gene Expression Slide with capture area and fiducial frame, spots and primer portions (Visium Spatial Gene Expression for FFPE User Guide)

4.3.2 Tissue preparation

To prepare our sample, the first step is to place the FFPE tissue block in a microtome and cut a few slices to expose the tissue (figure 35). Considering our region of interest, the FFPE block is cut with a razor blade to generate sections suitable for the Capture area of the Visium Slide (6,5 x 6,5 cm). The block is incubated for 1 hour in an ice bath for rehydration. Thereafter, a 5 µm slice is cut, left floating in the water bath with Milli-Q water at 42°C, and then carefully placed in the Captured area. The slide is placed in an oven at 42°C for 3 hours, then place in a desiccator and keep overnight at room temperature to ensure proper drying of the samples. The tissue must be deparaffinization. We use two 10-minute xylene passes to do this. Next, we rehydrate the tissue with ethanol passes at various concentrations (100%, 96%, 85% and 70%) before immersing the slide in Milli-Q water. The deparaffinized slide is subjected to H&E staining and protected by coverslip. Using an Imaging System (Olympus VS120 scanner), the image of the capture area is acquired at 10X magnification. Therefore, the coverslip must be removed using Milli-Q water after the slide is air-dried. The slide is placed in Visium Cassette and decrosslinked with three passes in 0.1 N HCL, one pass with TE buffer and followed by a 60-minute incubation at 70°C in a thermal cycler.

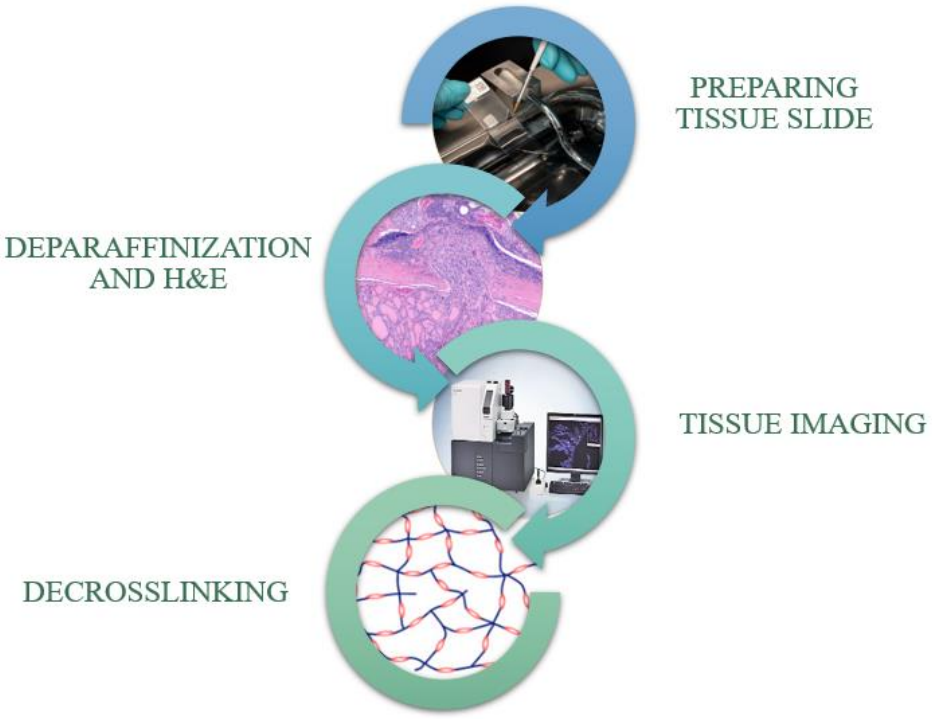


Figure 35 Workflow of Tissue preparation (Visium Spatial Gene Expression for FFPE User Guide)

4.3.3 Probe management and library construction

Once the samples are prepared, they will undergo the probe management and library construction phase, which involves many steps (figure 36)

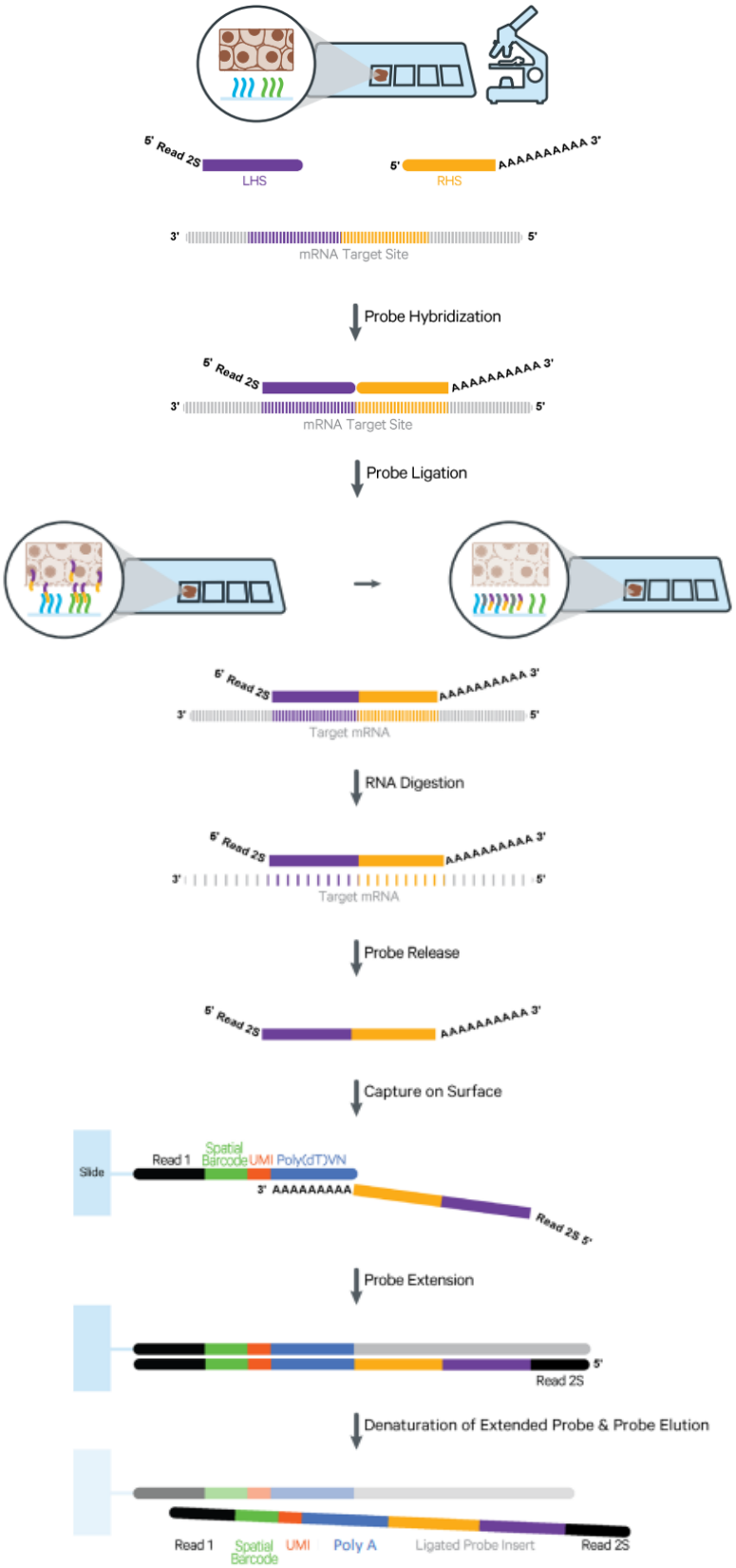


Figure 36 Probe management and library construction Workflow (Visium Spatial Gene Expression for FFPE User Guide)

The first step is to add to our sample the probe panel, consisting of pair of specific probes for each targeted gene. If the target RNA is present, the probe pairs hybridize to it (figure 37).

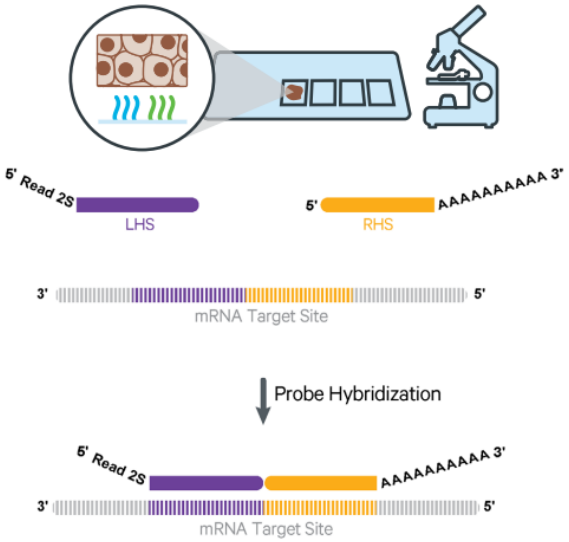


Figure 37 Hybridisation of the probe with its complementary RNA (Visium Spatial Gene Expression for FFPE User Guide)

To do this, we prepare the Pre-hybridization mix, according to the following table (table 6).

Pre-hybridization Mix *Add reagents in the order listed. Maintain at room temperature.*	10x PN	1X (µl)	4X+ 10% (µl)	8X + 10% (µl)
Nuclease-free Water	-	89.0	391.6	783.2
10X PBS	-	10.0	44	88
Diluted Perm Enzyme B Dilute Perm Enzyme B 200X in Buffer EB before using. DO NOT use concentrated stock directly.	3000602/ 3000553	0.5	2.2	4.4
10% Tween-20	-	0.5	2.2	4.4
Total	-	**100.0**	**440.0**	**880.0**

Table 6 Protocol to prepare Pre-hybridization mix (Visium Spatial Gene Expression for FFPE User Guide)

When the thermal cycler's program for decrosslinking is over, add 100 µl Pre-hybridization mix to samples and incubate for 15 minutes at room temperature.

At the same time, prepare Probe Hybridization Mix, according to the following table (table 7).

Probe Hybridization Mix *Add reagents in the order listed.* *Maintain at room temperature.*	10x PN	1X (µl)	4X + 10% (µl)	8X + 10% (µl)
Nuclease-free Water	-	10.0	44.0	88.0
FFPE Hyb Buffer	2000423	70.0	308.0	616.0
Human WT Probes –RHS **Or** **Mouse WT Probes – RHS**	2000453/ 2000449 Or 2000457/ 2000455	10.0	44.0	88.0
Human WT Probes – LHS **Or** **Mouse WT Probes – LHS**	2000454/ 2000450 Or 2000458/ 2000456	10.0	44.0	88.0
Total	-	**100.0**	**440.0**	**880.0**

Table 7 Protocol to prepare Probe Hybridization mix (Visium Spatial Gene Expression for FFPE User Guide)

After incubation, remove Pre-hybridization mix and add 100 µl Probe Hybridization Mix. Every time you need to add a reagent, you must remove the one previously inserted. Incubate at 50°C for 16 to 24 hours in a thermal cycler. Next, remove the unbound probe by washing it with FFPE Post-Hyb wash buffer preheated to 50°C. Incubate the Visium Cassette with the Buffer at 50°C for 5 minutes. Repeat two more times. Add 150 µl 2X SSC Buffer to each well and incubate at room temperature for 3 minutes.

4.3.4. Probe ligation

At this step, the ligase enzyme is added to seal the junction between the probe pairs that have hybridized to RNA, forming a ligation product (figure 38).

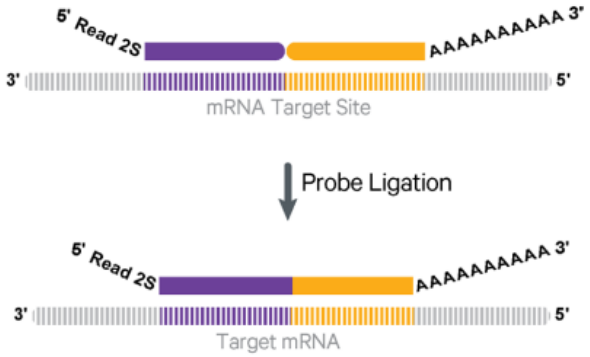


Figure 38 Ligation of probe pairs by ligase (Visium Spatial Gene Expression for FFPE User Guide)

After incubation, remove 2X SSC Buffer and add 60 µl Probe Ligation Mix, prepared according to the following protocol (table 8).

Probe Ligation Mix *Add reagents in the order listed. Maintain on ice.*	10x PN	1X (µl)	4X + 10% (µl)	8X + 10% (µl)
Nuclease-free Water	-	24.0	105.6	211.2
◯ 2X Probe Ligation Buffer	2000446/ 2000445	30.0	132.0	264.0
◯ Probe Ligation Enzyme	2000426/ 2000425	6.0	26.4	52.8
Total	-	**60.0**	**264.0**	**528.0**

Table 8 Protocol to prepare Probe Ligation Mix (Visium Spatial Gene Expression for FFPE User Guide)

Incubate at 37°C for 1 hour in a thermal cycler. After that, remove Probe ligation Mix, add 100 µl Post Ligation Wash Buffer and incubate at 57°C for 5 minutes in thermal cycler. Wash again with Post Ligation Wash Buffer, but this time heated to 57°C. Incubate at 57°C for 5 minutes. Remove Post Ligation Wash Buffer and wash twice with 150 µl 2X SSC buffer.

4.3.5 RNA digestion

In this phase, the target mRNA linked to the probe is destroyed, allowing the probes to be released (figure 39).

Figure 39 Target RNA digestion by RNase Mix (Visium Spatial Gene Expression for FFPE User Guide)

Remove 2X SSC buffer, add 75 µl RNase Mix, prepared according to the following table (table 9).

RNase Mix *Maintain on ice*	10x PN	1X (µl)	4X + 10% (µl)	8X + 10% (µl)
Nuclease-free Water	-	33.0	145.2	290.4
2X RNase Buffer	2000411/ 2000410	37.5	165.0	330.0
RNase Enzyme	3000605/ 3000593	4.5	19.8	39.6
Total	-	**75.0**	**330.0**	**660.0**

Table 9 Protocol to prepare RNase Mix (Visium Spatial Gene Expression for FFPE User Guide)

Incubate at 37°C for 30 minutes in thermal cycler.

4.3.6 Probe release and capture on the surface

The probes that are now free of the target RNA can be released into the tissue. The latter is made permeable so that the probes come out and bind to the spot of slide. This linkage is allowed between the Poly(dT) tail of the primer, present in the Visium Slide spots, and the Poly(dA) tail of the probe (figure 40).

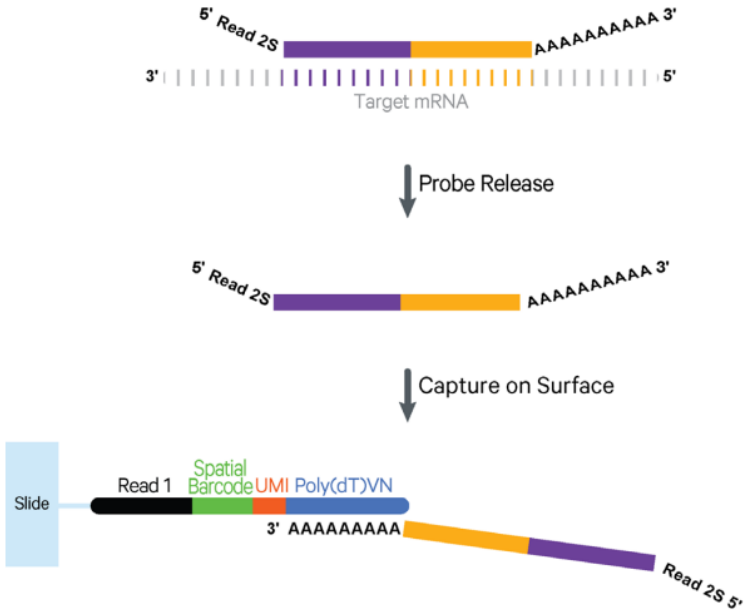


Figure 40 Probe release and capture of it on surface (Visium Spatial Gene Expression for FFPE User Guide)

Remove RNase Mix and add 75 µl Permeabilization mix, prepared according to the following table (table 10).

Permeabilization Mix *Maintain at room temperature*	10x PN	1X (µl)	4X + 10% (µl)	8X + 10% (µl)
Perm Buffer B	2000413/ 2000412	70.3	309.4	618.8
Perm Enzyme B	3000602/ 3000553	4.7	20.6	41.2
Total	-	**75.0**	**330.0**	**660.0**

Table 10 Protocol to prepare Permeabilization Mix (Visium Spatial Gene Expression for FFPE User Guide)

Incubate for 40 minutes at 37°C. Do three washes with 2X SSC Buffer.

4.3.7 Probe extension, denaturation and elution

Once the probe is bonded to the glass, the missing parts are synthesized to form a double-stranded chain. Once this is done, the extended probe must be denatured so that it is eluted and released into the solution (figure 41).

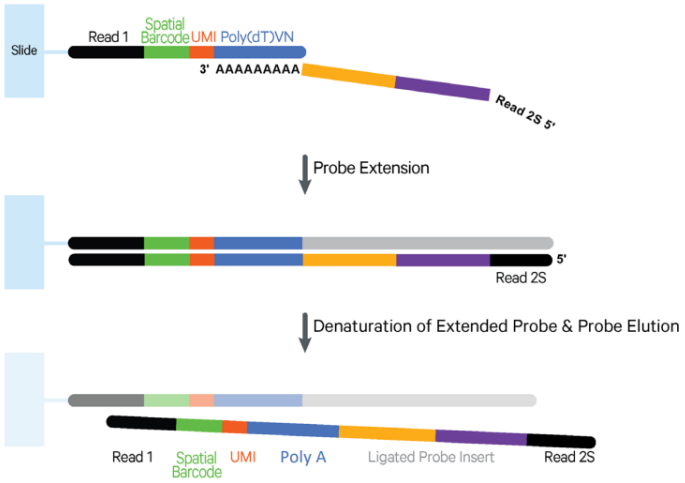


Figure 41 Probe extension, denaturation and elution (Visium Spatial Gene Expression for FFPE User Guide)

Add 75 µl Probe Extension Mix, prepared according to the following table (table 11).

Probe Extension Mix *Maintain on ice*	10x PN	1X (µl)	4X + 10% (µl)	8X + 10% (µl)
Extension Buffer	2000409/ 2000408	73.5	323.4	646.8
Extension Enzyme	2000427/ 2000389	1.5	6.6	13.2
Total		**75.0**	**330.0**	**660.0**

Table 11 Protocol to prepare Probe Extension Mix (Visium Spatial Gene Expression for FFPE User Guide)

Incubate for 15 minutes at 45°C in thermal cycler. Remove Probe Extension Mix and wash with 100 µl 2X SSC Buffer. Add 40 µl KOH Mix, prepared according to the following table (table 12). Incubate at room temperature for 10 minutes.

KOH Mix *Maintain at room temperature*	Stock	Final	1X (µl)	4X + 10% (µl)	8X + 10% (µl)
KOH	8 M	0.08 M	0.4	1.8	3.5
Nuclease-free Water	-	-	39.6	174.2	348.5
Total	-		**40.0**	**176.0**	**352.0**

Table 12 Protocol to prepare KOH mix (Visium Spatial Gene Expression for FFPE User Guide)

4.3.8 Library preparation

After 10 minutes of incubation, transfer the solution containing the ligation product to an 8-tube strip. Add 5 µl 1 M Tris-HCL pH 7.0 to solution in the 8-tube strip and place on ice (figure 42).

Figure 42 Transfer ligation product from Visium Cassette to 8-tube strip (Visium Spatial Gene Expression for FFPE User Guide)

Prepare qPCR mix, according to the following table (table 13).

qPCR Mix *Add reagents in the order listed. Maintain on ice.*	Stock	Final	1X (µl)	5X* + 10% (µl) *Includes 1 negative control	9X* + 10% (µl)
Nuclease-free Water	-	-	4.0	22.0	39.6
KAPA SYBR FAST qPCR Master Mix *Minimize light exposure*	2X	1X	5.0	27.5	49.5
TS Primer Mix A (PN-2000447)	-	-	1.0	5.5	9.9
Total			**10.0**	**55.0**	**99.0**

Table 13 Protocol to prepare qPCR mix (Visium Spatial Gene Expression for FFPE User Guide)

In a new PCR plate add 9 µl qPCR Mix and 1 µl sample/water. Add two negative controls. Set up the following PCR program, place the plate on the thermal cycler, and start the program (table 14).

Lid Temperature	Reaction Volume	Run Time
105°C	10 µl	35 min
Step	**Temperature**	**Time**
1	98°C	00:03:00
2	98°C	00:00:05
3	63°C	00:00:30
	Read signal	
4	Go to step 2, 24x (total 25 cycles)	-

Table 14 qPCR program to calculate the number of cycles of the next step (Visium Spatial Gene Expression for FFPE User Guide)

Once the PCR is complete, set the threshold along the exponential phase of the curve at 25% of the fluorescence peak value for each sample. Calculate the corresponding Cq with the threshold set; this will be the number of amplification cycles for the following reaction (figure 43).

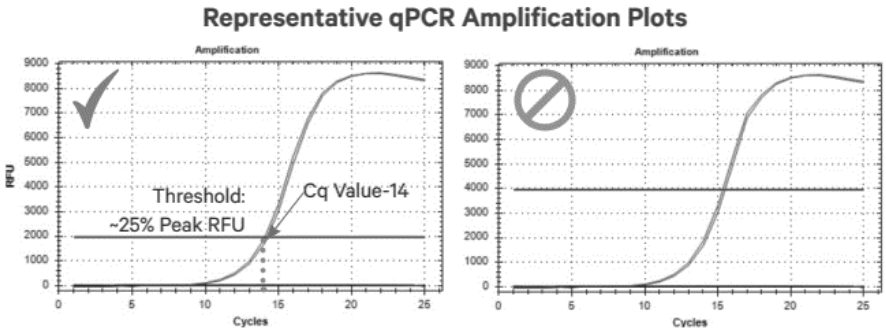


Figure 43 qPCR amplification plots which show how to calculate Cq (Visium Spatial Gene Expression for FFPE User Guide)

Choose appropriate sample index sets so that each sample has a unique index and this allows for recognition during post-pooling analysis. Add 50 µl Amp Mix to your samples, approximately the volume will be 45 µl. Add 5 µl of an individual Dual Index TS Set A to each well and
note which index you used for each sample. Incubate in a thermal cycler with the following protocol (table 15).

Lid Temperature	Reaction Volume	Run Time
105°C	100 µl	~25-40 min
Step	**Temperature**	**Time**
1	98°C	00:01:00
2	98°C	00:00:15
3	63°C	00:00:20
4	72°C	00:00:30
5	Use the Cq value as the total X of cycle	
6	72°C	00:01:00
7	4°C	Hold

Table 15 qPCR program (Visium Spatial Gene Expression for FFPE User Guide)

Proceeding with clean up. Add 85 µl paramagnetic beads (SPRIselect Reagent 0,85X) to each sample and incubate for 5 minutes at room temperature. Place on the magnet until the solution clears and remove the surnatant. Add 200 µl ethanol 80% to the pellet, so as to wash it, wait 30 seconds and remove the ethanol. Repeat this passage, so that the pellets are washed twice. Centrifuge briefly and remove as much ethanol as you can and air dry for 2 minutes. Remove from the magnet, add 25,5 Buffer EB and incubate for 2 minutes at room temperature. Place on the magnet until the solution clears and then transfer 25 µl to a new tube strip.

4.3.9 Sequencing and bioinformatics analysis

The libraries thus set up were sent by the company Integragen for sequencing. Before sequencing, they provided quality control using Agilent Bioanalyzer High Sensitivity chip, as required by the protocol. All four samples passed quality control and then were sequenced by Illumina Platform Sequencing, NovaSeq™ 6000 system. (figure 44).

Figure 44 Quality control was performed using BioAnalyzer and sequencing with Illumina platform.

With the help of bioinformaticians at the Gustave Roussy Institute, raw data from Integragen sequencing were processed with Space Ranger and then with CellsFromSpace (its user interface is called Shiny) (figure 45).

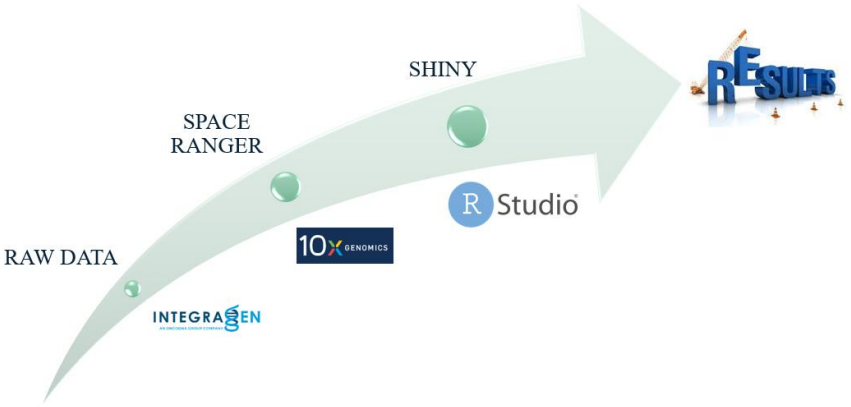


Figure 45 Workflow of bioinformatics analysis

Space Ranger is a software package, a set of analysis pipelines, developed by 10x Genomics that is used for the analysis of Visium Spatial Gene Expression. It helps researchers process, align, and analyze raw sequencing data to obtain gene expression information by considering the spatial organization of the tissue. Space Ranger uses raw FASTQ files from the sequencing phase as input. These files contain the RNA sequences. The next step that Space Ranger performs is the alignment of sequencing reads to the reference genome using a specialized alignment algorithm, this step determines which genes are expressed. Then barcode demultiplexing occurs: the barcodes used to label individual spots are demultiplexed to assign each read to a specific spot. After low-quality reads are filtered out and removed from further analysis, Space Ranger quantifies the gene expression of each spot. Data visualization can be performed using UMAP.

In addition to Space Ranger, a reference free decomposition framework for spatial transcriptomic dataset, called CellsFromeSpace (CFS), was used for the analysis. CFS was implemented by the bioinformaticians of the research group hosting me at Gustave Roussy.
CFS is based on independent component analysis (ICA). ICA allows the data obtained from bulk RNA sequencing to be decomposed to identify the independent components (ICs) that are responsible for gene expression. This is possible through biological interpretation and an expert-supervised process that performs precise manual annotation of the ICs and removes background noise and artifacts, without altering the signal of interest. CFS is supported by an easy-to-use user interface called Shiny, which makes it accessible even to non-bioinformaticians and allows the analysis of various ST datasets from different technologies besides Visium, such as Slide-seq, MERSCOPE, and CosMX. CFS thus allows a biologically relevant deconvolution of the spots into its components that are responsible for the expression observed within the tissue, such as cell type signatures, biological processes, and tissue organization.
In the specific case of this study, bioinformaticians at the Gustave Roussy Institute introduced me to CFS and subsequently I personally annotated the various ICs of the samples. Once the annotation is done, the relevant ICs are used to calculate clusters, which are a set of spots grouped according to the ICs present.

4.4 Real time PCR and digital PCR

For sample 3, the pathologist selected the area of normal tissue adjacent to the tumor (an area adjacent to the tissue with no identifiable tumor cells). Based on this selection, DNA was extracted from two consecutive 10 µm sections using the QuickExtract FFPE DNA extraction kit (LGC Biosearch Technologies, Berlin, Germany). DNA was used for Real-time PCR and digital PCR. Real-time PCR was performed using CE-IVD automated real-time polymerase chain reaction (RT-PCR) analysis, kit EasyPGX ready BRAF Diatech Pharmacogenetics, Jesi, Italy. The kit can identify the following *BRAF* gene variants: p.V600E (c.1799T>A; c.1799_1800delinsAA; variants not distinguishable from each other; analytical sensitivity 0.5%): p.V600K (c.1798_1799delinsAA; analytical sensitivity 1.0%); p.V600D/p.V600E (c.1799_1800delinsAT; c.1798_1799delinsAG; variants not distinguishable from each other; analytical sensitivity 1.0%). The input DNA was 20 ng for each well containing primers and probes in anhydrous format. The thermal profile provided by the protocol was used for the reaction (table 16).

Step	**Temperature**	**Time**
Hot start (1 cycle)	95°C	5 minutes
Amplification (40 cycles)	95°C	15 seconds
	56°C	15 seconds
	60°C	45 seconds

Table 16 Thermal profile of real time PCR

In the other hand, Digital PCR was performed with QuantStudio Absolute Q Digital PCR system. This system combines digitalization, scanning, thermal cycling, and data collecting into one instrument, which is all that is necessary for dPCR. The extracted unquantified sample was diluted 1:10 as our laboratory's standardized procedure for performing dPCR so as to avoid overloading the sample. The diluted sample was mixed according to the data sheet (table 17) with Absolute Q™ DNA Digital PCR Master Mix (5x) and probe Absolute Q™ Liquid Biopsy dPCR Assays *BRAF* V600E 40X, with *BRAF* V600E marked with FAM and *BRAF* WT marked with VIC.

Reagent	**Volume**
Absolute Q™ DNA Digital PCR Master Mix (5x)	2 ul
Absolute Q™ Liquid Biopsy dPCR Assays *BRAF* V600E 40X	2,5 ul
DNA	5,5 ul

Table 17 Protocol for the preparation of the reaction for QuantStudio Absolute Q Digital PCR system

The 9 ul volume of the reaction was pipetted into microfluidic array plate (MAP) and then 15 ul of isolation buffer was added. The thermal profile provided by the protocol was used for the reaction (table 18)

Step	**Temperature**	**Time**
Hot start (1 cycle)	96°C	10 minutes
Amplification (40 cycles)	96°C	5 seconds
	60°C	15 seconds

Table 18 Thermal profile of digital PCR

The system enables consistent and automated sample separation in 20’480 microcameras, with a coefficient of variation (CV) of less than 1 percent. Data were analyzed using QuantStudio Absolute Q Digital PCR 6.1.0 software.

5. Result

5.1 Next generation sequencing

The cases were analyzed by Next Generation Sequencing (NGS). Case 1 harbors pathogenic DICER1 mutation and cases 3 and 4 have the canonical *BRAF* V600E mutation. In the table below (table 19) are the details of the mutations found in the samples.

	GENE	EXON	AMINO ACID CHANGE	CODING	FREQUENCY (%)	ACMG CLASSIFICATION
Sample 1&2	*DICER1*	27	p.Asp1810Val	c.5429 A>T	41	Pathogenic
Sample 3	*BRAF*	15	p.Val600Glu	c.1799 T>A	20	Pathogenic
Sample 4	*BRAF*	15	p.Val600Glu	c.1799 T>A	26	Pathogenic

Table 19 Mutations found in samples by NGS with descriptions of gene, exon, amino acid change, coding, allele frequency and ACMG classification.

5.2 Spatial transcriptomic

To better visualize the results of the spatial transcriptomics test, each sample was divided by CellFromSpace into IC related-clusters according to the spatial transcriptomics profiles, i.e., each cluster corresponds to a set of ICs sharing similar expression profiles.
For each sample, we displayed the plot regarding the degree of separation present between each cluster (figure 46). For samples 1, 2, 3 and 4 we obtained 12, 6, 11 and 10 clusters, respectively. Each dot within the graph represents a spot, and the color of the dot represents which cluster it belongs to; in fact, clusters are denoted by a number and a color.
Each spot is graphed according to its expression, and the closer the spots of the same color, that is, belonging to the same cluster, are to each other, the more reliable the results are. The more the clusters are separated from each other, the more they have different gene expression from the others.
Samples 1, 3 and 4 have clusters with a good degree of separation, spots belonging to the same cluster are close to each other and well-spaced from the rest. Sample 2, on the other hand, has a strong overlap of signals, many spots belonging to different clusters are overlapping. This problem is due to the presence of different cell types within the same spot and the difficulty of having good signal separation.

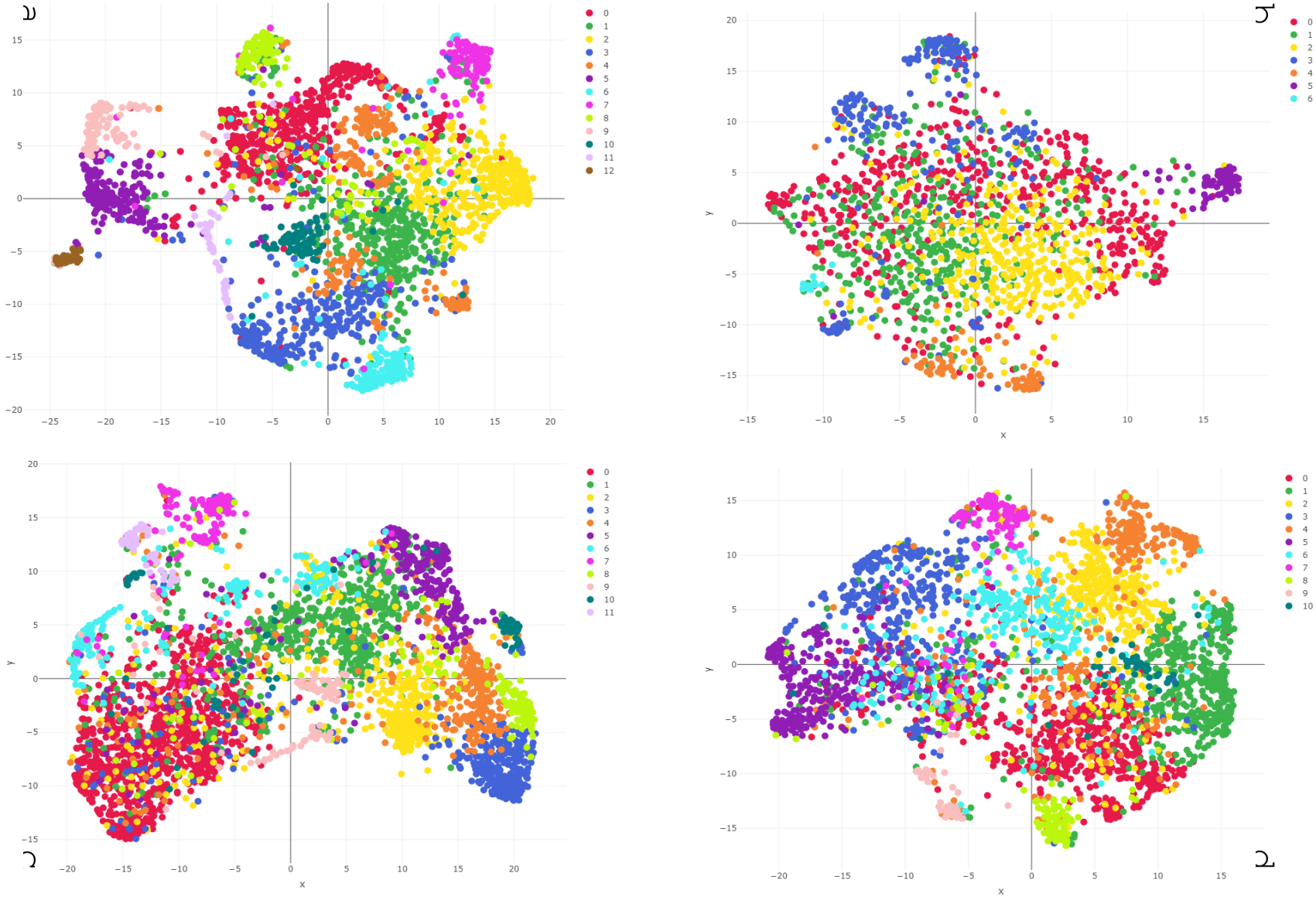


Figure 46 UMAP projection of spots, colored by cluster identity. (a)sample 1, (b) sample 2, (c) sample 3, (d) sample 4.

Considering the spatial component of the study, in the figures 47 and 48 we see on the right the graphs related to the spatial organization of the clusters identified in the different samples and on the left the H&E related to the sample. On the H&E there are white arrows indicating the direction of tumor progression and they were included in order to be able to discriminate which clusters are associated with invasion.

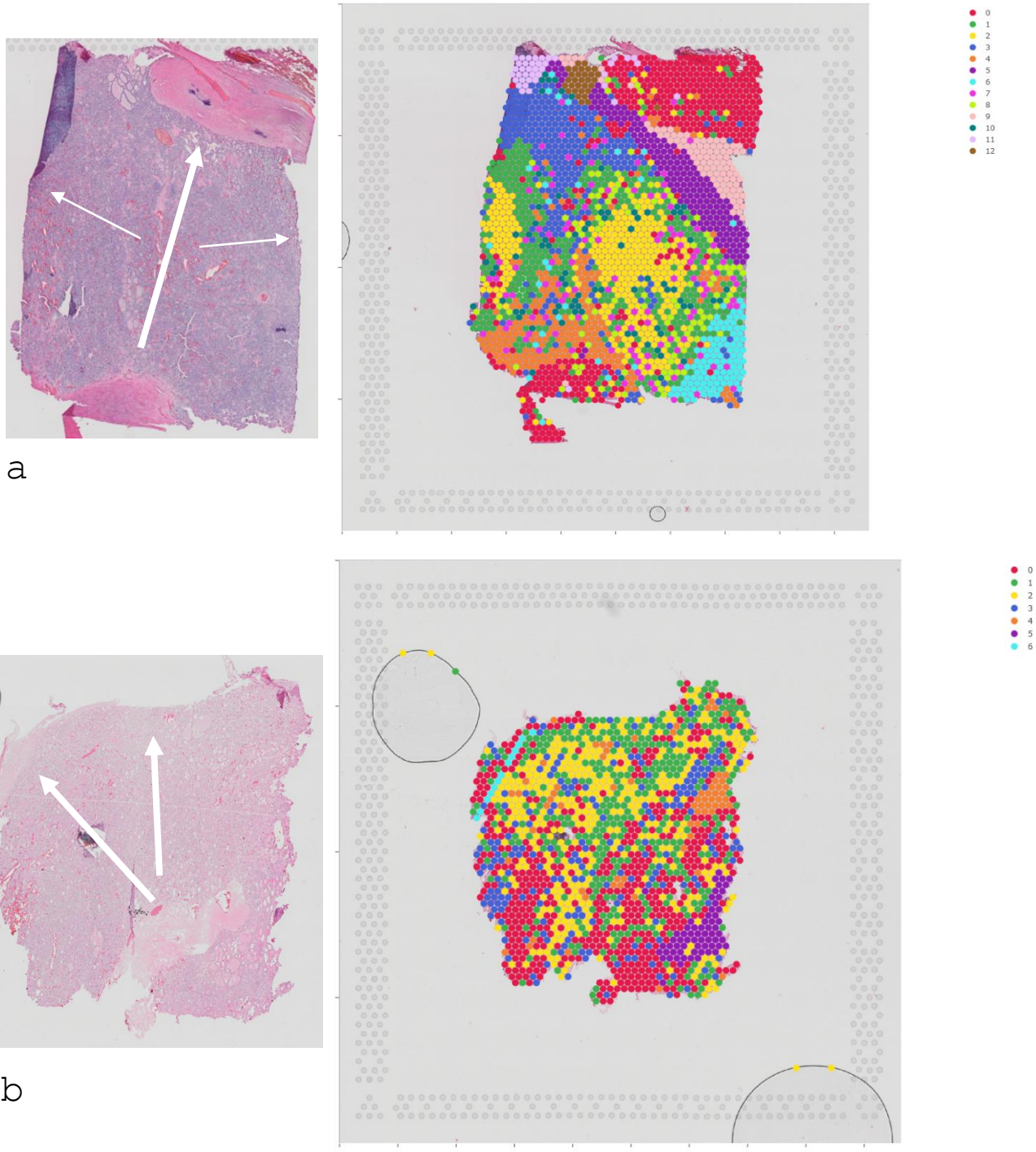
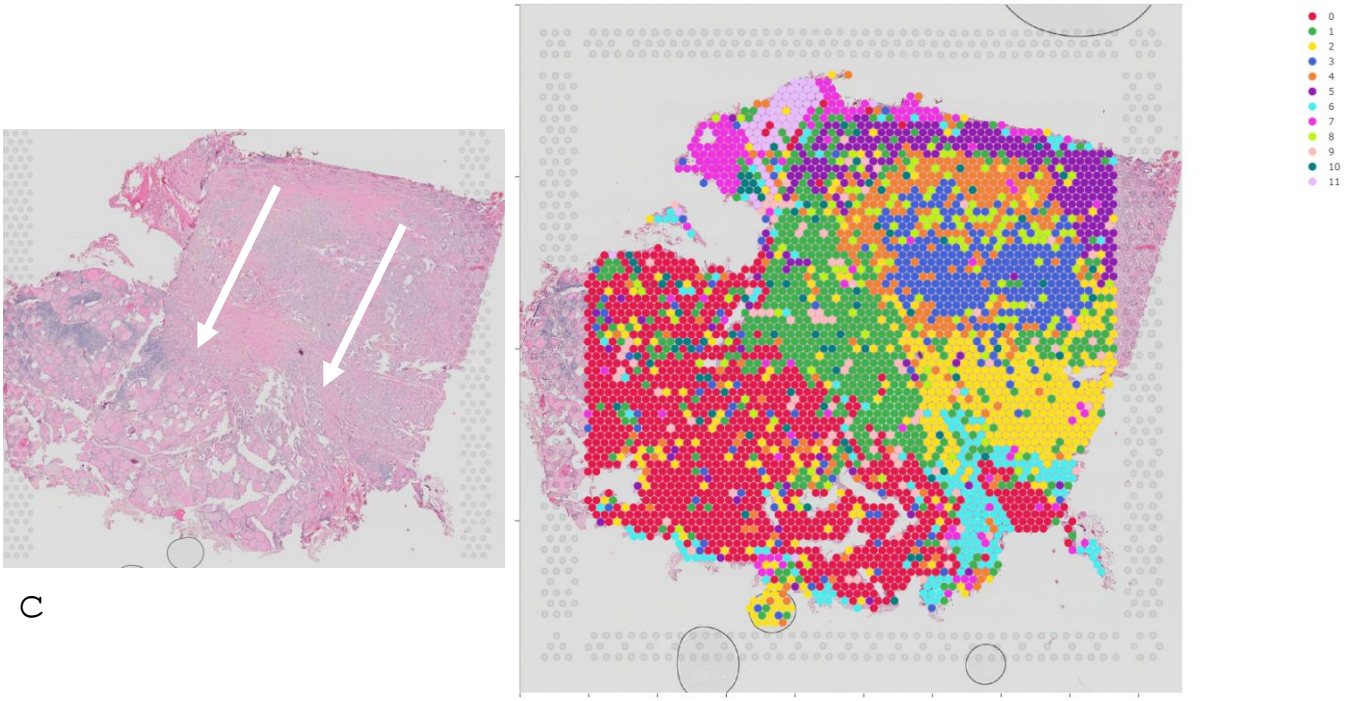
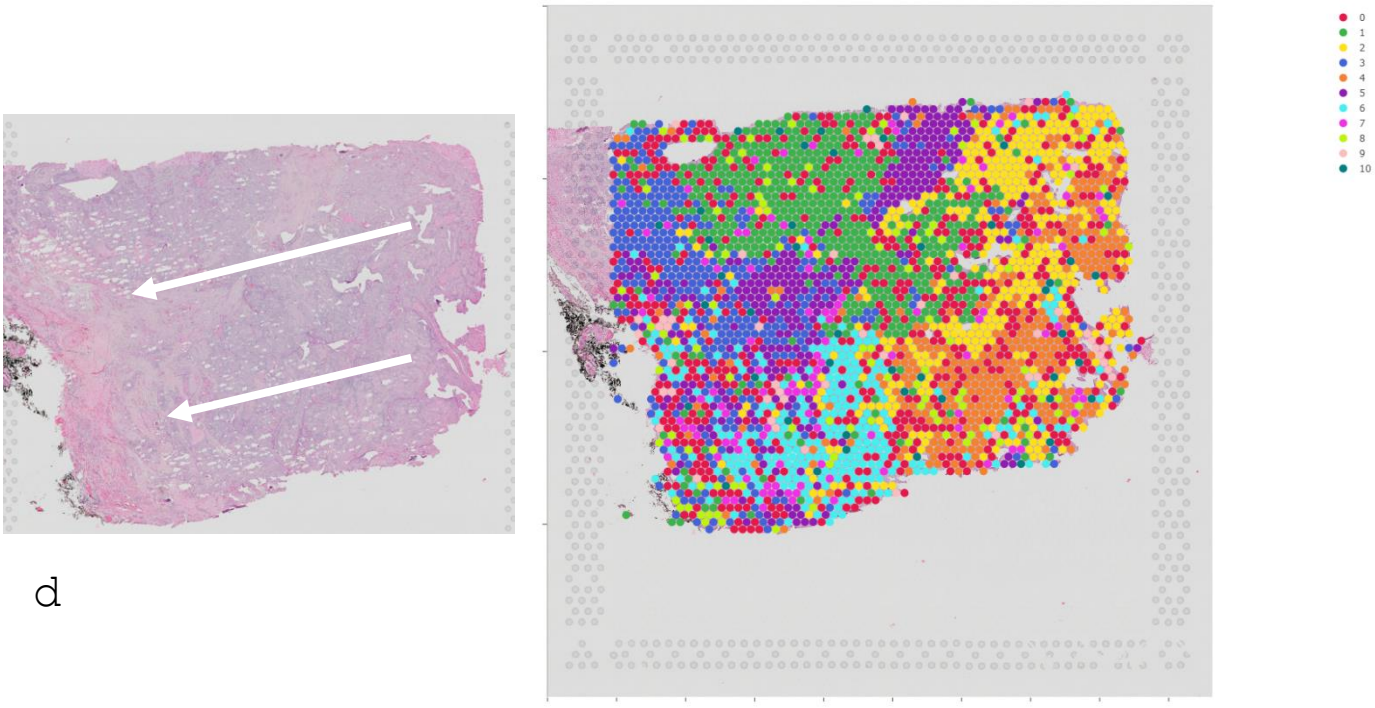


Figure 47 Cluster-related spatial plot of sample 1 (a) and 2 (b), and their H&E with white arrow indicating the directionality of tumor progression (clustering resolution 1,2)



c



d

Figure 48 Cluster-related spatial plot of sample 3 (c) and 4 (d), and their H&E with white arrow indicating the directionality of tumor progression (clustering resolution 1,2)

One of the tasks we had set ourselves was the identification of genes related to tumor invasion. Primarily with Prof.re Tallini, the directionality of tumor progression was identified for each sample by microscopic analysis of the H&E. The respected directionality was shown on the H&E image using white arrows (Figures 47, 48). Based on this, I identified the clusters associated with tumor progression and they are as follows:

- Sample 1: cluster 3, 5, 9, 11, 12;
- Sample 3: cluster 1, 2, 6;
- Sample 4: cluster 1, 3, 5, 6.

Using an algorithm found in Shiny based on internal standards, I identified which genes were upregulated in these clusters and selected only the upregulated genes in at least three clusters of those examined (table 20). In the table shown are both clusters in which the gene is upregulated and clusters in which the gene is downregulated.

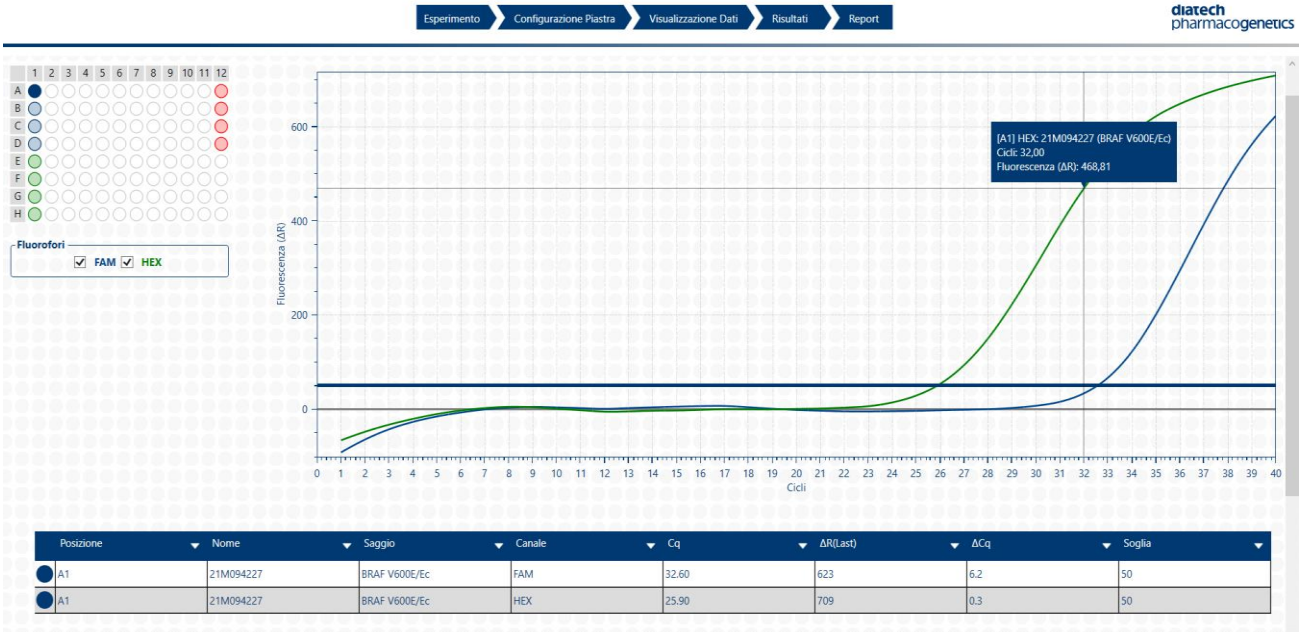
Genes in at least 3 invasion zones	UP							DOWN									
UP in different samples																	
APOE	**S1 CL9**	**S1 CL3**	S3 CL8	**S3 CL1**	**S3 CL6**	**S4 CL6**		S3 CL9	S3 CL0	S3 CL3	S3 CL7	S3 CL8					
COL1A2	**S1 CL3**	**S1 CL9**	S3 CL11	**S3 CL2**	**S4 CL3**			S1 CL0	S1 CL1	S1 CL2	S3 CL0	**S3 CL1**	S4 CL2	S4 CL4			
LUM	**S1 CL5**	**S3 CL2**	S3 CL11	**S4 CL3**	**S1 CL9**	**S4 CL3**		S1 CL0	S1 CL1	S1 CL2	S1 CL8	S3 CL0	**S3 CL1**	S3 CL7	S3 CL8	S4 CL2	S4 CL4
DCN	**S1 CL5**	**S1 CL9**	**S4 CL3**					S1 CL1	S1 CL2	S1 CL8	S1 CL0	**S3 CL1**	**S3 CL6**	S3 CL8	S4 CL2	S4 CL4	
COL3A1	**S1 CL9**	**S1 CL3**	**S1 CL5**	S2 CL6	**S3 CL2**	S3 CL11		S1 CL0	S1 CL2	S1 CL8	S2 CL2	**S3 CL2**	S3 CL7	S3 CL0	**S3 CL1**		
SPARC	**S1 CL9**	**S1 CL12**	S2 CL6	S3 CL11	**S4 CL3**	S4 CL8		S1 CL0	**S1 CL11**	S1 CL2	S4 CL2	S4 CL4					
COL1A1	**S1 CL9**	**S1 CL3**	**S3 CL2**	S3 CL9	S3 CL11	**S4 CL3**	S4 CL2	S1 CL10	S1 CL0	S1 CL1	**S3 CL1**	S3 CL7	S4 CL2	S4 CL4			
IGKC	**S1 CL9**	**S1 CL3**	S3 CL0	S3 CL5	**S3 CL6**	S4 CL0		S1 CL1	S1 CL2	S1 CL4	**S3 CL1**	S3 CL3	S3 CL8	**S4 CL6**			
AEBP1	**S1 CL9**	**S3 CL2**	S3 CL11	**S4 CL3**				**S1 CL11**	S1 CL10	**S3 CL1**	S3 CL7	S3 CL0	S4 CL2	S4 CL4			
TIMP3	**S1 CL12**	**S1 CL9**	**S1 CL12**	**S4 CL3**				S1 CL2	S4 CL2	S4 CL4							
COL6A3	**S1 CL3**	**S1 CL9**	**S3 CL2**	S3 CL11	**S4 CL3**			S1 CL0	S3 CL7	S3 CL0	**S3 CL1**	S3 CL7	S4 CL4	S4 CL2			
TAGLN	**S1 CL9**	S3 CL10	**S4 CL3**	S4 CL8	**S4 CL3**			S1 CL0	S1 CL6	**S3 CL1**	S4 CL4	S4 CL2					
PRELP	**S1 CL9**	**S3 CL2**	S3 CL3	**S4 CL3**				**S3 CL1**									
CTSK	**S1 CL9**	**S4 CL3**	**S4 CL6**					S4 CL4	S4 CL2								
CDH11	**S1 CL9**	**S3 CL2**	**S4 CL3**					S3 CL7	S3 CL0	S4 CL2							
COL5A1	**S1 CL9**	**S3 CL2**	S3 CL11	**S4 CL3**				S3 CL7	S3 CL0	S4 CL2	S4 CL4						
VCAN	**S1 CL9**	**S3 CL2**	S3 CL3	S3 CL11	**S4 CL3**			**S3 CL1**	S3 CL7	S4 CL4	S4 CL2						
SULF1	**S1 CL9**	**S3 CL2**	**S4 CL3**					S4 CL2									

Table 20 Table 20 List of upregulated genes in at least three clusters associated with invasion. All clusters in which the gene

S = SAMPLE CL=CLUSTER

5.3 Real-Time PCR and digital PCR

Sample 3 has normal tissue adjacent to the tumor. Spatial transcriptomics showed the presence of tumor-associated cluster-related spots in the normal tissue (Figure 48). This suggests that tumor cells may have infiltrated the normal tissue. To verify this possible presence, we used real-time PCR (RT-PCR) and digital PCR (dPCR) to check whether the same *BRAF V600E* mutation found in the tumor tissue is present in the adjacent normal tissue.
As we can see in the figure (figure 49), RT-PCR detected the presence of *BRAF* V600E in normal tissue. The amplification curve in green represents the HEX fluorochrome that is associated with control, while the amplification curve in blue represents the FAM fluorochrome that is associated with *BRAF* V600E mutation. The HEX curve of the control rises at cycle 25, while the FAM curve of *BRAF* rises at cycle 32, thus indicating positivity to the mutation.



Posizione	Nome	Saggio	Canale	Cq	ΔR(Last)	ΔCq	Soglia
A1	21M094227	BRAF V600E/Ec	FAM	32.60	623	6.2	50
A1	21M094227	BRAF V600E/Ec	HEX	25.90	709	0.3	50

Figure 49 BRAF V600E and BRAF WT mutation amplification curve related to sample 3, the green curve associated with the HEX fluorochrome represents BRAF WT, while the blue curve associated with the FAM fluorochrome represents BRAF V600E

As for dPCR it too shows positivity for the presence of the *BRAF* V600E mutation. Two probes were used in this case: FAM to mark *BRAF* V600E and VIC to mark *BRAF* WT (table 21). Of the total of 20'480 microcameras, 20472 microcameras were included in the analysis, indicating a good quality of the analysis. 11 microchambers were positive for *BRAF* V600E, while 250 were positive for *BRAF* WT (table 21, figure 50). The concentration calculated as copy number per ul is 1.24 and 28.44(table 21), respectively, meaning that *BRAF* V600E would account for 4 percent of the total.

Dye	Target	Conc. cp/uL	95%CI	Positives
FAM	V600E	1.24	0.56	11
VIC	WT	28.44	3.32	250

Table 21 Short report of results related to digital PCR

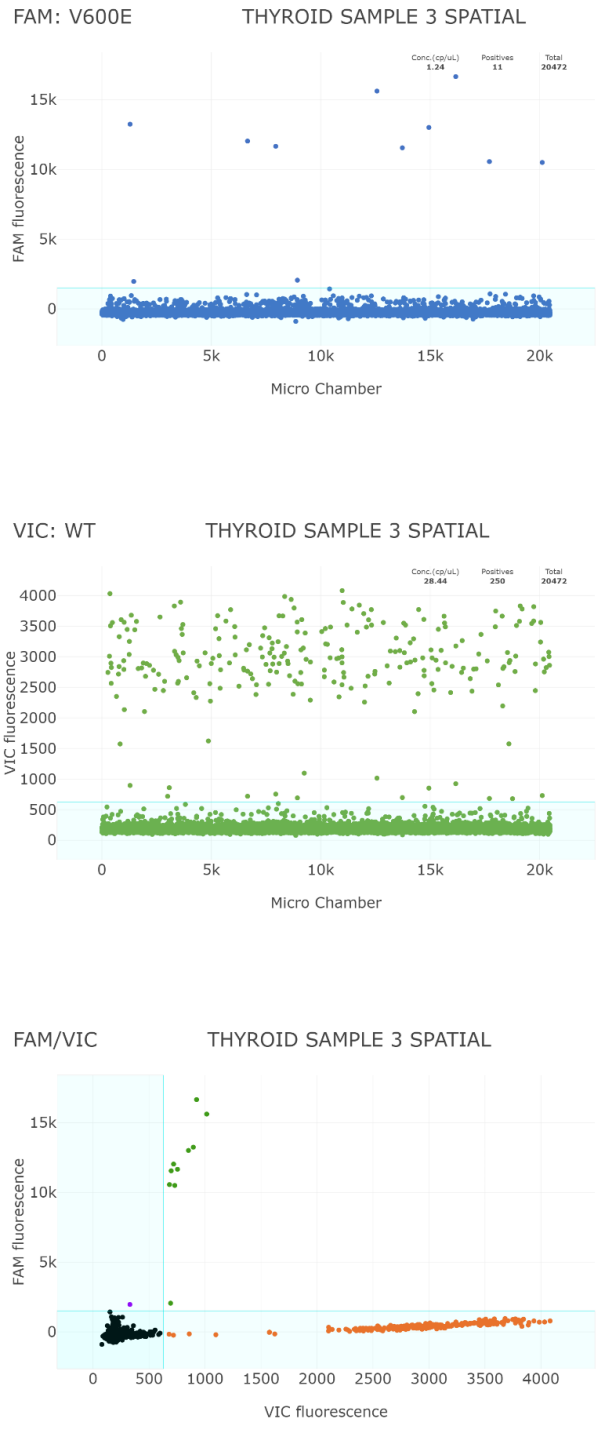


Figure 50 Graphs showing positive wells for BRAF V600E labeled FAM, BRAF WT labeled VIC, and in the last graph only positive wells are shown either individually for one fluorochrome or positive for both.

6. Discussion

Spatial transcriptomics is a state-of-the-art method for studying gene expression in space. In this project, we wanted to apply it to thyroid carcinomas with the aim of identifying specific expression patterns related to tumor progression.
This pilot study was based on the analysis of 4 thyroid carcinoma samples, one of which was subsequently excluded, and this represents one of the major limitations of the project. The small number of samples is due to the high costs of the method and the large amount of data being generated that are also difficult to manage and analyze.
Even at first glance, one can see that the gene expression in the various samples is heterogeneous, even considering their small size (6.5 x 6.5 mm maximum).
In sample 1, two clusters (cluster 5 and 9) are clearly visible close to the capsule, indicators that the cells near the capsule have a different gene transcript than the other cells in the sample.
Sample 2 was temporarily excluded from the study because the strong overlap of signals, due to the presence of different cell types in the same spot, does not allow immediate analysis of the sample. Future studies with deconvolution techniques could well separate the individual signals so that the data obtained from Sample 2 could be used.
Sample 1 and sample 2 are part of Case 1, which harbors the DICER1 mutation at 41%. Considering the percentage, this mutation could be present in the germ line and in the case would indicate the presence of DICER1 syndrome. In that case, confirmation by germline mutation analysis is required [116]. Cases 2 and 3, on the other hand, have the canonical *BRAF* V600E mutation characteristic of papillary carcinomas.
Sample 3 is particularly interesting because in the normal tissue adjacent to the tumor, the *BRAF* V600E mutation present in the neoplastic tissue was found at a lower percentage. Considering the mutation percentages, 20 percent for the tumor and about 4 percent for the adjacent normal, would rule out the presence of the mutation in the germline, which might instead account for the presence of the mutation in both tissues. This raises suspicions about the possible presence of infiltrating tumor cells in normal tissue. Further studies should be done to verify this. To the best of our knowledge no such studies have been conducted at present.
Regarding genes identified upregulated in at least three invasion-associated clusters, there is a group consisting of VCAN, COL6A3, COL5A1, COL1A2, LUM, DCN, COL3A1, SPARC and COL1A1 that are involved in extracellular matrix remodeling and have already been associated with tumor progression in a study by Tang et al[117].

On the other hand, as for the APOE gene, coding for Apolipoprotein E, it has already been found to be upregulated in papillary carcinoma compared with normal tissue [Lin Xu 2022] and in our study it is also upregulated in follicular carcinoma. This upregulation is concordant with the fact that APOE is involved in lipid metabolism, dysregulations of which have already been closely related to other cancers [117].
Another gene related to lipid metabolism is AEBP1. This gene has been identified as a marker in glioblastoma (GBM) because it interacts with PTEN and inhibits its function, so its overexpression promotes cell proliferation [118].
The CTSK gene, on the other hand, in several types of carcinomas (prostate, breast, lung, kidney, colon, ovary and others) can promote cancer cell proliferation, invasion and migration [119].
CDH11 was found upregulated in PTC and ATC, while in breast cancer it correlated with metastasis and invasion [120].
On the other hand, TIMP3, TAGLN, PRELP, and SULF1 appear to be oncosuppressors: TIMP3 in thyroid carcinoma [121], TAGLN in colorectal carcinoma [122], PRELP in hepatocarcinoma as its overexpression correlates with favorable prognosis [123], while the fourth inhibits proliferation and invasion of esophageal squamous cell carcinoma cells [124].
Regarding the last gene, IGKC is an immunoglobulin, thus a marker of Plasma cells and B cells.
Future studies could investigate in more detail what genes are associated with invasion, because in our case we found genes associated with cell proliferation and invasion in thyroid or other cancers (VCAN, COL6A3, COL5A1, COL1A2, LUM, DCN, COL3A1, SPARC, COL1A1, APOE, AEBP1, CTSK, CDH11) while others (TIMP3, TAGLN, PRELP, and SULF1) with which oncosuppressor activity is associated. Further studies could be conducted.
In the future, it would also be interesting to correlate the results of spatial transcriptomics with any mutations found in the samples under study, such as in our case.
An important aspect that could be investigated with a more in-depth study is to identify differentially expressed genes between papillary carcinoma and follicular carcinoma, considering the different patterns of invasion. This was among the aims of the project, but to date it was difficult with our data to come up with a clear result. It remains one of our future prospects.
Spatial transcriptomics is the new frontier in the study of cancer and may allow us to gain a deeper understanding of the mechanisms and interactions between cancer cells in order to identify new therapeutic targets.
Knowledge as a means to defeat.

Bibliography

1. *Basicmedical Key*. 2023.10.07]; Available from: https://basicmedicalkey.com/embryology-and-developmental-lesions]/.
2. *Etymology of the word ''Thyroid'' - etimo.it*.
3. Araque, D.V.P., A. Bleyer, and J.P. Brito, *Thyroid cancer in adolescents and young adults*. Future Oncol, 2017. 13(14): p. 1253-1261.
4. Ilahi A, M.E., Ilahi TB, *Anatomy, Head and Neck, Parathyroid.*, S. Publishing, Editor. 2022.
5. Fitzpatrick, T.H. and M.A. Siccardi, *Anatomy, Head and Neck, Adam's Apple*, in *StatPearls*. 2022, StatPearls Publishing Copyright © 2022, StatPearls Publishing LLC.: Treasure Island (FL).
6. Jacobsen, B., N. VanKampen, and J.V. Ashurst, *Anatomy, Head and Neck, Thyrohyoid Membrane*, in *StatPearls*. 2022, StatPearls Publishing Copyright © 2022, StatPearls Publishing LLC.: Treasure Island (FL).
7. Chaudhary, P., et al., *Levator glandulae thyroideae, a fibromusculoglandular band with absence of pyramidal lobe and its innervation: a case report.* J Clin Diagn Res, 2013. 7(7): p. 1421-4.
8. Jardaly, A.H. and B. Bordoni, *Anatomy, Thorax, Thyrocervical Arteries*, in *StatPearls*. 2022, StatPearls Publishing Copyright © 2022, StatPearls Publishing LLC.: Treasure Island (FL).
9. Mansberger AR, J., Wei JP, *Surgical embryology and anatomy of the thyroid and parathyroid glands*. Surg Clin North Am. Vol. 73. 1993.
10. Policeni, B.A., W.R. Smoker, and D.L. Reede, *Anatomy and embryology of the thyroid and parathyroid glands.* Semin Ultrasound CT MR, 2012. 33(2): p. 104-14.
11. Chintamani, *"Friend or Foe" of a Thyroid Surgeon?-the Tubercle of Zuckerkandl.* Indian J Surg, 2013. 75(5): p. 337-8.
12. Mullur, R., Y.Y. Liu, and G.A. Brent, *Thyroid hormone regulation of metabolism.* Physiol Rev, 2014. 94(2): p. 355-82.
13. G., G.D. and S.D. M., *Greenspan's Basic And Clinical Endocrinology*. 9 th edition ed.
14. Sharma., M.A.A.S., *Physiology, Thyroid Hormone*. 2023.
15. Chin, W.W., et al., *Thyroid hormone regulation of thyrotropin gene expression.* Recent Prog Horm Res, 1993. 48: p. 393-414.
16. Chaker, L., et al., *Hypothyroidism.* Lancet, 2017. 390(10101): p. 1550-1562.
17. Johnson, J.L. and J.V. Felicetta, *Hyperthyroidism: a comprehensive review.* J Am Acad Nurse Pract, 1992. 4(1): p. 8-14.
18. Hegedüs, L., *Clinical practice. The thyroid nodule.* N Engl J Med, 2004. 351(17): p. 1764-71.
19. Jemal, A., et al., *Cancer statistics, 2009.* CA Cancer J Clin, 2009. 59(4): p. 225-49.
20. Sung, H., et al., *Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries.* CA Cancer J Clin, 2021. 71(3): p. 209-249.
21. Zhai, M., et al., *The global burden of thyroid cancer and its attributable risk factor in 195 countries and territories: A systematic analysis for the Global Burden of Disease Study.* Cancer Med, 2021. 10(13): p. 4542-4554.
22. Rahbari, R., L. Zhang, and E. Kebebew, *Thyroid cancer gender disparity.* Future Oncol, 2010. 6(11): p. 1771-9.
23. National Cancer Institute. Surveillance, E.a.E.r.p. *SEER Surveillance Epidemiology and End-Results Cancer Registries Program (SEER) An interactive website for SEER cancer statistics*. [cited 2023 25/05/2023]; official website of the United States government]. Available from: https://seer.cancer.gov/.

24. Pizzato, M., et al., *The epidemiological landscape of thyroid cancer worldwide: GLOBOCAN estimates for incidence and mortality rates in 2020.* Lancet Diabetes Endocrinol, 2022. 10(4): p. 264-272.
25. Lortet-Tieulent, J., et al., *Thyroid cancer "epidemic" also occurs in low- and middle-income countries.* Int J Cancer, 2019. 144(9): p. 2082-2087.
26. Siegel, R.L., et al., *Cancer statistics, 2022.* CA Cancer J Clin, 2022. 72(1): p. 7-33.
27. EpiCentro, I.S.d.S., *Registro Tumori. I numeri del cancro in Italia nel 2021*. 2021.
28. Dal Maso, L., et al., *Incidence of thyroid cancer in Italy, 1991-2005: time trends and age-period-cohort effects.* Ann Oncol, 2011. 22(4): p. 957-963.
29. Luo, J., et al., *Hysterectomy, Oophorectomy, and Risk of Thyroid Cancer.* J Clin Endocrinol Metab, 2016. 101(10): p. 3812-3819.
30. Nikiforov, Y.E., *Is ionizing radiation responsible for the increasing incidence of thyroid cancer?* Cancer, 2010. 116(7): p. 1626-8.
31. Ron, E., et al., *A population-based case-control study of thyroid cancer.* J Natl Cancer Inst, 1987. 79(1): p. 1-12.
32. Richardson, D.B., *Exposure to Lonizing Radiation in Adulthood and Thyroid Cancer Incidence.* Epidemiology, 2009. 20: p. 7 pages.
33. Shakhtarin, V.V., et al., *Iodine deficiency, radiation dose, and the risk of thyroid cancer among children and adolescents in the Bryansk region of Russia following the Chernobyl power station accident.* Int J Epidemiol, 2003. 32(4): p. 584-91.
34. Aschebrook-Kilfoy, B., et al., *Thyroid cancer risk and dietary nitrate and nitrite intake in the Shanghai women's health study.* Int J Cancer, 2013. 132(4): p. 897-904.
35. Samani, A.A., et al., *The role of the IGF system in cancer growth and metastasis: overview and recent insights.* Endocr Rev, 2007. 28(1): p. 20-47.
36. Jorde, R. and J. Sundsfjord, *Serum TSH levels in smokers and non-smokers. The 5th Tromsø study.* Exp Clin Endocrinol Diabetes, 2006. 114(7): p. 343-7.
37. Derwahl, M. and D. Nicula, *Estrogen and its role in thyroid cancer.* Endocr Relat Cancer, 2014. 21(5): p. T273-83.
38. Zhao, S., et al., *Association of obesity with the clinicopathological features of thyroid cancer in a large, operative population: A retrospective case-control study.* Medicine (Baltimore), 2019. 98(50): p. e18213.
39. Büyükaşık, O., et al., *The association between thyroid malignancy and chronic lymphocytic thyroiditis: should it alter the surgical approach?* Endokrynol Pol, 2011. 62(4): p. 303-8.
40. Xu, J., et al., *Hashimoto's Thyroiditis: A "Double-Edged Sword" in Thyroid Carcinoma.* Front Endocrinol (Lausanne), 2022. 13: p. 801925.
41. Bogović Crnčić, T., et al., *Risk Factors for Thyroid Cancer: What Do We Know So Far?* Acta Clin Croat, 2020. 59(Suppl 1): p. 66-72.
42. Peiling Yang, S. and J. Ngeow, *Familial non-medullary thyroid cancer: unraveling the genetic maze.* Endocr Relat Cancer, 2016. 23(12): p. R577-R595.
43. Quayle, F.J. and J.F. Moley, *Medullary thyroid carcinoma: including MEN 2A and MEN 2B syndromes.* J Surg Oncol, 2005. 89(3): p. 122-9.
44. Cooper, D.S., et al., *Revised American Thyroid Association management guidelines for patients with thyroid nodules and differentiated thyroid cancer.* Thyroid, 2009. 19(11): p. 1167-214.
45. Durante, C., et al., *The Diagnosis and Management of Thyroid Nodules: A Review.* JAMA, 2018. 319(9): p. 914-924.
46. Grani, G., et al., *Contemporary Thyroid Nodule Evaluation and Management.* J Clin Endocrinol Metab, 2020. 105(9): p. 2869-83.

47. Bläker, H., et al., *Age-dependent performance of BRAF mutation testing in Lynch syndrome diagnostics.* Int J Cancer, 2020. 147(10): p. 2801-2810.
48. Nardi, F., et al., *Italian consensus for the classification and reporting of thyroid cytology.* J Endocrinol Invest, 2014. 37(6): p. 593-9.
49. Aster, J.C., V. Kumar, and A.K. Abbas, *ROBBINS & COTRAN PATHOLOGIC BASIS OF DISEASE*. 10th edition ed. 2020.
50. Oncology, A.S.o.C. *Cancer.net*. 05/06/2023]; Available from: https://www.cancer.net/cancer-types/thyroid-cancer/statistics#:~:text=For%20localized%20anaplastic%20thyroid%20cancer,papillary%20thyroid%20cancer%20is%2099%25.
51. Wei, X., et al., *Risk and Prognostic Factors for BRAF.* Biomed Res Int, 2022. 2022: p. 9959649.
52. de Biase, D., et al., *High-sensitivity BRAF mutation analysis: BRAF V600E is acquired early during tumor development but is heterogeneously distributed in a subset of papillary thyroid carcinomas.* J Clin Endocrinol Metab, 2014. 99(8): p. E1530-8.
53. De Leo, A., et al., *Expanding the Spectrum of BRAF Non-V600E Mutations in Thyroid Nodules: Evidence-Based Data from a Tertiary Referral Centre* Int J Mol Sci, 2023. 24(4).
54. Network, C.G.A.R., *Integrated genomic characterization of papillary thyroid carcinoma.* Cell, 2014. 159(3): p. 676-90.
55. Haroon Al Rasheed, M.R. and B. Xu, *Molecular Alterations in Thyroid Carcinoma.* Surg Pathol Clin, 2019. 12(4): p. 921-930.
56. Yang, J. and J.A. Barletta, *Anaplastic thyroid carcinoma.* Semin Diagn Pathol, 2020. 37(5): p. 248-256.
57. Lang, B.H., et al., *Classical and follicular variant of papillary thyroid carcinoma: a comparative study on clinicopathologic features and long-term outcome.* World J Surg, 2006. 30(5): p. 752-8.
58. *WHO Classification of tumours, Digestive System Tumours*. 5 ed. 2019.
59. Yu, X.M., et al., *Follicular variant of papillary thyroid carcinoma is a unique clinical entity: a population-based study of 10,740 cases.* Thyroid, 2013. 23(10): p. 1263-8.
60. Liu, J., et al., *Follicular variant of papillary thyroid carcinoma: a clinicopathologic study of a problematic entity.* Cancer, 2006. 107(6): p. 1255-64.
61. Nikiforov, Y.E., et al., *Nomenclature Revision for Encapsulated Follicular Variant of Papillary Thyroid Carcinoma: A Paradigm Shift to Reduce Overtreatment of Indolent Tumors.* JAMA Oncol, 2016. 2(8): p. 1023-9.
62. Redman, J.R. and S.M. Armstrong, *Reentrainment of rat circadian activity rhythms: effects of melatonin.* J Pineal Res, 1988. 5(2): p. 203-15.
63. Rivera, M., et al., *Molecular genotyping of papillary thyroid carcinoma follicular variant according to its histological subtypes (encapsulated vs infiltrative) reveals distinct BRAF and RAS mutation patterns.* Mod Pathol, 2010. 23(9): p. 1191-200.
64. Feldt-Rasmussen, U., *Iodine and cancer.* Thyroid, 2001. 11(5): p. 483-6.
65. Giuffrida, D. and H. Gharib, *Controversies in the management of cold, hot, and occult thyroid nodules.* Am J Med, 1995. 99(6): p. 642-50.
66. Badulescu, C.I., et al., *FOLLICULAR THYROID CARCINOMA - CLINICAL AND DIAGNOSTIC FINDINGS IN A 20-YEAR FOLLOW UP STUDY.* Acta Endocrinol (Buchar), 2020. 16(2): p. 170-177.
67. Tate, J.G., et al., *COSMIC: the Catalogue Of Somatic Mutations In Cancer.* Nucleic Acids Res, 2019. 47(D1): p. D941-D947.
68. Prete, A., et al., *Update on Fundamental Mechanisms of Thyroid Cancer.* Front Endocrinol (Lausanne), 2020. 11: p. 102.

69. Volante, M., et al., *Poorly differentiated thyroid carcinoma: the Turin proposal for the use of uniform diagnostic criteria and an algorithmic diagnostic approach.* Am J Surg Pathol, 2007. 31(8): p. 1256-64.
70. Ibrahimpasic, T., et al., *Outcomes in patients with poorly differentiated thyroid carcinoma.* J Clin Endocrinol Metab, 2014. 99(4): p. 1245-52.
71. Chao, T.C., J.D. Lin, and M.F. Chen, *Insular carcinoma: infrequent subtype of thyroid cancer with aggressive clinical course.* World J Surg, 2004. 28(4): p. 393-6.
72. Ibrahimpasic, T., et al., *Poorly Differentiated Carcinoma of the Thyroid Gland: Current Status and Future Prospects.* Thyroid, 2019. 29(3): p. 311-321.
73. Landa, I., et al., *Genomic and transcriptomic hallmarks of poorly differentiated and anaplastic thyroid cancers.* J Clin Invest, 2016. 126(3): p. 1052-66.
74. Rao, S.N., et al., *Patterns of Treatment Failure in Anaplastic Thyroid Carcinoma.* Thyroid, 2017. 27(5): p. 672-681.
75. Glaser, S.M., et al., *Anaplastic thyroid cancer: Prognostic factors, patterns of care, and overall survival.* Head Neck, 2016. 38 Suppl 1: p. E2083-90.
76. Khan, H.Y., et al., *Targeting XPO1 and PAK4 in 8505C Anaplastic Thyroid Cancer Cells: Putative Implications for Overcoming Lenvatinib Therapy Resistance.* Int J Mol Sci, 2019. 21(1).
77. Stamatakos, M., et al., *Medullary thyroid carcinoma: The third most common thyroid cancer reviewed.* Oncol Lett, 2011. 2(1): p. 49-53.
78. Cerrato, A., V. De Falco, and M. Santoro, *Molecular genetics of medullary thyroid carcinoma: the quest for novel therapeutic targets.* J Mol Endocrinol, 2009. 43(4): p. 143-55.
79. Moura, M.M., B.M. Cavaco, and V. Leite, *RAS proto-oncogene in medullary thyroid carcinoma.* Endocr Relat Cancer, 2015. 22(5): p. R235-52.
80. Filetti, S., et al., *Thyroid cancer: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-updagger.* Ann Oncol, 2019. 30(12): p. 1856-1883.
81. Turing, A.M., *Computing machinery and intelligence.* Mind, 1950. LIX.
82. Murdoch, T.B. and A.S. Detsky, *The inevitable application of big data to health care.* JAMA, 2013. 309(13): p. 1351-2.
83. Jiang, F., et al., *Artificial intelligence in healthcare: past, present and future.* Stroke Vasc Neurol, 2017. 2(4): p. 230-243.
84. Alrafiah, A.R., *Application and performance of artificial intelligence technology in cytopathology.* Acta Histochem, 2022. 4.
85. Baxi, V., et al., *Digital pathology and artificial intelligence in translational medicine and clinical practice.* Mod Pathol, 2022. 35(1): p. 23-32.
86. Haugen, B.R., et al., *2015 American Thyroid Association Management Guidelines for Adult Patients with Thyroid Nodules and Differentiated Thyroid Cancer: The American Thyroid Association Guidelines Task Force on Thyroid Nodules and Differentiated Thyroid Cancer.* Thyroid, 2016. 26(1): p. 1-133.
87. *The Pathologist*. 28/05/2923]; Available from: https://thepathologist.com/outside-the-lab/constant-demand-patchy-supply.
88. *The Bethesda System for Reporting Thyroid Cytopathol ogy. Definitions, Criteria, and Explanatory Notes*. 2018.
89. de Biase, D., et al., *Molecular Diagnostic of Solid Tumor Using a Next Generation Sequencing Custom-Designed Multi-Gene Panel.* Diagnostics (Basel), 2020. 10(4).
90. al., H.e. *Deep residual Learning fir Image recognition*. 23/09/2023]; Available from: https://arxiv.org/abs/1512.03385.
91. Weiss, K., T.M. Khoshgoftaar, and D. Wang, *A survey of transfer learning.* Journal of Big Data, 2016. 3.

92. Deng, J., et al., *ImageNet: a Large-Scale Hierarchical Image Database*, in *IEEE Conference on Computer Vision and Pattern Recognition*. 2009. p. 248-255.
93. Müller., A.C. and S. Guido., *Introduction to Machine Learning with Python: a guide for data scientist*. 2016: O'Reilly Media Inc.
94. Hastie, T., R. Tibshirani , and J. Friedman, *The Elements of Statistical Learning Data Mining, Inference, and Prediction*. 2013, Springer.
95. Abadi, M., et al., *TensorFlow: A system for large-scale machine learning.* Osdi., 2016. 16.
96. Karakitsos, P., et al., *Potential of the back propagation neural network in the morphologic examination of thyroid lesions.* Anal Quant Cytol Histol, 1996. 18(6): p. 494-500.
97. Ippolito, A.M., et al., *Neural network analysis for evaluating cancer risk in thyroid nodules with an indeterminate diagnosis at aspiration cytology: identification of a low-risk subgroup.* Thyroid, 2004. 14(12): p. 1065-71.
98. Guan, Q., et al., *Deep convolutional neural network VGG-16 model for differential diagnosing of papillary thyroid carcinomas in cytological images: a pilot study.* J Cancer, 2019. 10(20): p. 4876-4882.
99. Hirokawa, M., et al., *Application of deep learning as an ancillary diagnostic tool for thyroid FNA cytology.* Cancer Cytopathol, 2022.
100. Sanyal, P., et al., *Artificial Intelligence in Cytopathology: A Neural Network to Identify Papillary Carcinoma on Thyroid Fine-Needle Aspiration Cytology Smears.* J Pathol Inform, 2018. 9: p. 43.
101. Esteban, Á., et al., *A new optical density granulometry-based descriptor for the classification of prostate histological images using shallow and deep Gaussian processes.* Comput Methods Programs Biomed, 2019. 178: p. 303-317.
102. Teramoto, A., et al., *Automated classification of benign and malignant cells from lung cytological images using deep convolutional neural network.* Informatics in Medicine Unlocked, 2019. 16.
103. Awan, R., et al., *Deep learning based digital cell profiles for risk stratification of urine cytology images.* Cytometry A, 2021. 99(7): p. 732-742.
104. Moses, L. and L. Pachter, *Museum of spatial transcriptomics.* Nat Methods, 2022. 19(5): p. 534-546.
105. John, H.A., M.L. Birnstiel, and K.W. Jones, *RNA-DNA hybrids at the cytological level.* Nature, 1969. 223(5206): p. 582-7.
106. Gall, J.G. and M.L. Pardue, *Formation and detection of RNA-DNA hybrid molecules in cytological preparations.* Proc Natl Acad Sci U S A, 1969. 63(2): p. 378-83.
107. Harrison, P.R., et al., *Localisation of cellular globin messenger RNA by in situ hybridisation to complementary DNA.* FEBS Lett, 1973. 32(1): p. 109-12.
108. Langer-Safer, P.R., M. Levine, and D.C. Ward, *Immunological method for mapping genes on Drosophila polytene chromosomes.* Proc Natl Acad Sci U S A, 1982. 79(14): p. 4381-5.
109. Tautz, D. and C. Pfeifle, *A non-radioactive in situ hybridization method for the localization of specific RNAs in Drosophila embryos reveals translational control of the segmentation gene hunchback.* Chromosoma, 1989. 98(2): p. 81-5.
110. Rao, A., et al., *Exploring tissue architecture using spatial transcriptomics.* Nature, 2021. 596(7871): p. 211-220.
111. Nishino, M. and J. Jacob, *Invasion in thyroid cancer: Controversies and best practices.* Semin Diagn Pathol, 2020. 37(5): p. 219-227.
112. Sobrinho Simoes M, A.S., Kroll TG, Nikiforov Y, DeLellis R, Farid P, Kitamura Y, Noguchi SU, Eng C, Harach HR, et al., *World Health Organization Classification of Tumors: Pathology and Genetics of Tumors of Endocrine Organs.* . 2004: IARC Press.

113. Chiapponi, C., et al., *Lymphatic Vessel Invasion in Routine Pathology Reports of Papillary Thyroid Cancer.* Front Med (Lausanne), 2022. 9: p. 841550.
114. Baloch, Z.W., et al., *Overview of the 2022 WHO Classification of Thyroid Neoplasms.* Endocr Pathol, 2022. 33(1): p. 27-63.
115. Kopanos, C., et al., *VarSome: the human genomic variant search engine.* Bioinformatics, 2019. 35(11): p. 1978-1980.
116. Nosé, V., *DICER1 gene alterations in thyroid diseases.* Cancer Cytopathol, 2020. 128(10): p. 688-689.
117. Tang, X., et al., *Identifying gene modules of thyroid cancer associated with pathological stage by weighted gene co-expression network analysis.* Gene, 2019. 704: p. 142-148.
118. Reddy, S.P., et al., *Novel glioblastoma markers with diagnostic and prognostic value identified through transcriptome analysis.* Clin Cancer Res, 2008. 14(10): p. 2978-87.
119. Mikosch, P., et al., *High cathepsin K levels in men with differentiated thyroid cancer on suppressive L-thyroxine therapy.* Thyroid, 2008. 18(1): p. 27-33.
120. Chen, J., J.A. Gingold, and X. Su, *Immunomodulatory TGF-β Signaling in Hepatocellular Carcinoma.* Trends Mol Med, 2019. 25(11): p. 1010-1023.
121. Mazzoni, M., et al., *Transcriptomic landscape of TIMP3 oncosuppressor activity in thyroid carcinoma.* Cancer Cell Int, 2022. 22(1): p. 400.
122. Li, Q., et al., *TAGLN suppresses proliferation and invasion, and induces apoptosis of colorectal carcinoma cells.* Tumour Biol, 2013. 34(1): p. 505-13.
123. Hong, R., et al., *PRELP has prognostic value and regulates cell proliferation and migration in hepatocellular carcinoma.* J Cancer, 2020. 11(21): p. 6376-6389.
124. Liu, C.T., et al., *SULF1 inhibits proliferation and invasion of esophageal squamous cell carcinoma cells by decreasing heparin-binding growth factor signaling.* Dig Dis Sci, 2013. 58(5): p. 1256-63.