Immunotherapy in Recurrent/Metastatic Head and Neck Squamous Cell Carcinoma: biological landscape and prognostic biomarkers to improve patients' stratification

Thesis

Immunotherapy in Recurrent/Metastatic Head and Neck Squamous Cell Carcinoma: biological landscape and prognostic biomarkers to improve patients' stratification

Mara Serena Serafini

Thesis presented for the Degree of Doctor of Philosophy

The Open University, Milton Keynes (UK)
Faculty of Science Technology Engineering and Mathematics (STEM)
Discipline: School of Life, Health and Chemical Sciences

Date of submission: October 2022

Affiliated Research Centre:
Fondazione IRCCS Istituto Nazionale dei Tumori, Milan (Italy)

Director of study: Prof. Lisa Francesca Licitra
Internal supervisor: Dr. Loris De Cecco
Internal supervisor: Dr. Vera Cappelletti

TABLE OF CONTENT

ABSTRACT

Head and neck squamous cell carcinoma (HNSCC) is the eight most frequent cancer in the world, and approximately 2/3 of the patients are diagnosed at locally advanced stages (stage III or IV). Despite improvements in HNSCC management and the aggressiveness of first-line curative treatment, 65% of treated patients experience local recurrence or distant metastasis. Moreover, patients diagnosed with recurrence or distant metastasis have a poor prognosis (from 6 to 12 months) and 5-years survival less than 50%. From 2016 the clinical practice was revolutionized by the introduction of immuno-checkpoint inhibitors, approved for the treatment of recurrent/metastatic HNSCC patients. Nevertheless, only a small subset of patients respond to this therapy and currently predictive biomarkers are still under investigation. Herein, we investigated the tumor biology of R/M HNSCC patients, platinum-refractory, enrolled in the phase IIIb clinical trial Nivactor (EudraCT Number: 2017-000562-30), in which patients were treated with nivolumab. Across the study of single biomarkers and the extensive profiling through genomic and transcriptomic analyses, we aimed to characterize the tumor molecular peculiarity of patients that experienced response or those with the longer survival. While the prognostic/predictive role of Programmed Cell Death Ligand-1 (PD-L1; studied by IHC), Tumor mutational burden (TMB) and microsatellite instability (MSI) appeared to be relatively limited for R/M HNSCC patients, 8 expression signatures (retrieved from literature) showed up significant association with survival and contributed to highlight and extricate the extreme complexity of the tumor microenvironment of HNSCC, which appeared to be strongly immunosuppressive (suggesting and corroborating the activation of several mechanisms of immune evasion). Nevertheless, the testing of previously identified six HNSCC subtypes (*De Cecco et al.*) with specific biological and prognostic characteristics, indicated for two of them a strong prognostic role and a significant correlation with response. In conclusion, the current study demonstrated the strong relevance of gene expression signatures in HNSCC context (over the mere study of somatic mutations) to identify the biological features associated with benefits from immunotherapy. However, additional analyses for the validation of their significance are required.

DECLARATION OF AUTHORSHIP

I hereby certify that the thesis I am submitting is my own original work, except for:

a) Recruitment and clinical assessment of patients was performed by clinical partners in the 21 Italian centers;
b) PD-L1 IHC was performed by dr. Federica Perrone and dr. Andrea Vingiani;
c) Revision of clinical data collected was performed by dr. Stefano Cavalieri;
d) Bionformatic analyses for gene expression were performed by dr. Loris De Cecco and dr. Andrea Carenzo;
e) Bioinformatic analyses for mutational part were performed by dr. Andrea Carenzo;

All the data used in this thesis are publicly available and were properly acknowledged as references.

LIST OF ABBREVIATIONS

Acquired immunodeficiency syndrome (AIDS)

American joint committee on cancer stage (AJCC)

Areas under the curve (AUC)

Combined positive score (CPS)

Complete responder (CR)

Computed tomography (CT)

Damage response and repair (DRR)

DNA damage response (DDR)

Eastern cooperative oncology group (ECOG)

Epidermal growth factor receptor (EGFR)

False discovery rate (FDR)

Food and drug administration (FDA)

Formalin-fixed paraffin embedded (FFPE)

Gene-set enrichment analysis (GSEA)

Head and neck cancer (HNC)

Head and neck squamous cell carcinoma (HNSCC)

Hematoxiline & eosine (H&E)

Human immunodeficiency virus (HIV)

Human papilloma virus (HPV)

Immune-checkpoint inhibitors (ICI)

Immuno histochemistry (IHC)

Interferon (IFN)

Interferon-γ (IFN-γ)

Kyoto encyclopaedia of genes and genomes (KEGG)

Magnetic resonance imaging (MRI)

Major histocompatibility class I (MHC-I)

Microsatellite instability (MSI)

Mismatch repair system (MMR)

Natural killer cells (NK)

Next generation sequencing (NGS)

Normalized enrichment score (NES)

Objective response rate (ORR)

Oral cavity squamous cell carcinoma (OCSCC)

Oropharyngeal squamous cell carcinoma (OPSCC)

Overall response rate (ORR)

Overall survival (OS)

Partial responder (PR)

Performance status (PS)

Programmed death protein-1 (PD-1)

Programmed death-ligand 1 (PD-L1)

Progression free survival (PFS)

Progressive disease (PD)

Recurrence/metastasis (R/M)

Response Evaluation Criteria in Solid Tumor (RECIST)

Reverse transcriptase quantitative polymerase chain reaction (RT-qPCR)

Single nucleotide variants (SNVs)

Stable disease (SD)

Standard-of-care (SoC)

T-regulatory cells (Tregs)

Tumor infiltrating lymphocytes (TILs)

Tumor microenvironment (TME)

Tumor mutational burden (TMB)

Tumor proportion score (TPS)

Tumor, lymph node, metastasis (TNM)

Unique molecular identifiers (UMIs)

Whole exome sequencing (WES)

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

1.1 History of cancer

Evidence demonstrated that the history of cancer begins in ancient times, before the appearance of Homo sapiens: cancer was reported to be present in prehistoric animals by paleopathologists, studying hundred thousand years old fossils[1]. Nevertheless, cancer appeared to be closely related to Homo sapiens. The recent human history was described connected with oncologic disease in many written records during centuries, from 3000 BC by Egyptian and Chinese cultures, followed by Greek, Romans, till Contemporary Era. During the last 4000 years human begins are constantly trying to understand, characterize and eradicate the oncological disease[23]. Biologically speaking, cancer mostly represents the mirror of aging in all the living things; a reminder of oldness, an accumulation of tiredness, directly correlated to the rise in longevity. Nevertheless, it could be not only reduced to a mere representation of the time passing by. Cancer could mirror the results of an essential evolution condition, a consequence of the constant pressure at cellular and molecular level, inducing cells through several genomic changes, mostly influenced by the surrounding environment[4,5]. Nowadays, cancer is an urgent global issue. In 2018 cancer alone was responsible for nearly 10 million deaths, according to World Health Organization, and it is expected to be the leading cause of death in the world by the end of this century. The exponential growth worldwide could be partially explained by aging, social and economic development, limited access to oncology care, pollution, lifestyle choices (such as smoking and alcohol use). All these aspects result correlated with the increase in population and countries development. The biological contribution related to these factors is still under investigation[6,7,8]. The problems concerning this debilitating disease could not be discussed

examining only patient's health status, even if it's the imperative we must consider. As a matter of fact, cancer critically impacts the financial status of both patients and society. For instance, in 2015 in European Union more than €57 billion were spent for cancer-related healthcare costs[9].

1.2 Cancer origin

When we apply the definition of cancer we usually think of a specific disease. However, deepening insight, cancer cannot be described as a single, unique entity, and its heterogeneity is mirrored in several ways. At date, more than 100 cancer types have been described, reflecting the complexity of this disease. Major cancer classification is based on the type of primary tissue in which the oncogenic process originates; for this reason, cancer can be classified in two large categories: solid tumors (approximately 90%) and hematologic malignancies (10%)[10]. Solid tumors comprise:

i) Carcinomas (90% of solid tumors) are classified the malignancies that arise in epithelial cells, at different layers and sites. When the disease starts at the basal layer, it is called basal cell carcinoma; while when it happens at the surface it is called squamous cell carcinoma (from name of the thin, flat, squamous cells at the top of the skin).

ii) Sarcomas, it's the name given to those cancers that originate in the bone's cells and soft tissues, such as connective tissue, muscles, supportive tissue, etc. At date, more than 70 different types of sarcomas have been described.

Hematologic malignances include:

i) Leukemia is one of the three malignancies affecting the hematopoiesis process and blood forming organs, including bone marrow; leukemia impaired the formation of healthy blood cells, including myeloid and lymphoid cells, resulting in a large accumulation of unfunctional cells in the blood circle.
ii) Myeloma specifically involves the formation of plasma cells in bone marrow, consequently leading to anemia, or disease related to kidneys, and others.
iii) Lymphoma is the third malignancy, and it interests blood forming tissues (specifically the lymphatic system and lymphocytes), leading to the aggregation of immature white blood cells and the formation of masses in the lymphatic vessels; it can be classified in non-Hodgkin's lymphoma, the more prevalent form, or Hodgkin's lymphoma, a more aggressive form of the disease.

1.3 Types of carcinomas

Even if the tissue, in which cancer generates, maintains specific characteristics in different organs, we known that each organ of origin represents a different oncological disease. Thus, cancers are additionally classified based on the organ in which the disease originates. In 2020 more than 19 million cases worldwide were distributed, accordingly to the site of origin, in 36 more frequent different cancers. According to incidence, the five most frequently observed cancer (independently from patients' gender and country) were in order:
1) female breast cancer, with 2,261,419 (11,7%);
2) lung cancer, with 2,206,771 cases (11,4%);

3) prostate cancer, with 1,414,259 cases (7,3%);

4) non-melanoma skin cancer, with 1,198,073 cases (6,2%);

5) colon cancer, with 1,148,515 (6,0%)[11].

1.4 Head and nek cancer

Head and neck cancer (HNC) are a large spectrum of malignancies that affect the upper aerodigestive tract, although they aggregate in a single classification, they not to be considered as a single entity. Cancers arising from different anatomic sites are considered as HNC, including cancers deriving from salivary gland, soft tissues, skin, mucosal membrane and even bones. They are frequently discovery by primary care physician, dentists, in routine tests[12]. The most common symptoms are dysphagia, otalgia, weight loss, oral pain, mucosal irregularity, and ulceration. Moreover, head and neck cancers are associated with pain, disfiguration, dysfunction, and psychosocial distress [13,14]. In the complex heterogeneity of HNC, a distinct homogeneous characteristic is represented by histology: more than 90% of HNC are squamous cell carcinoma, cancers deriving from the epithelium. Head and neck squamous cell carcinomas (HNSCC) are the eight most common cancer in the world, accounting in 2020 for approximately 750,000 cases (of which 9.856 new cases in Italy, https://www.registri-tumori.it/) and over 400,000 deaths annually. The anatomical sites considered HNSCC are oral cavity (377,713 cases), larynx (184,615 cases), oropharynx (98,412 cases) and hypopharynx (84,254 cases, Figure 1)[15]. Therefore, even if the histology considered is univocal, the intricacy of HNSCC is strongly represented by the subsites that are considered[16].

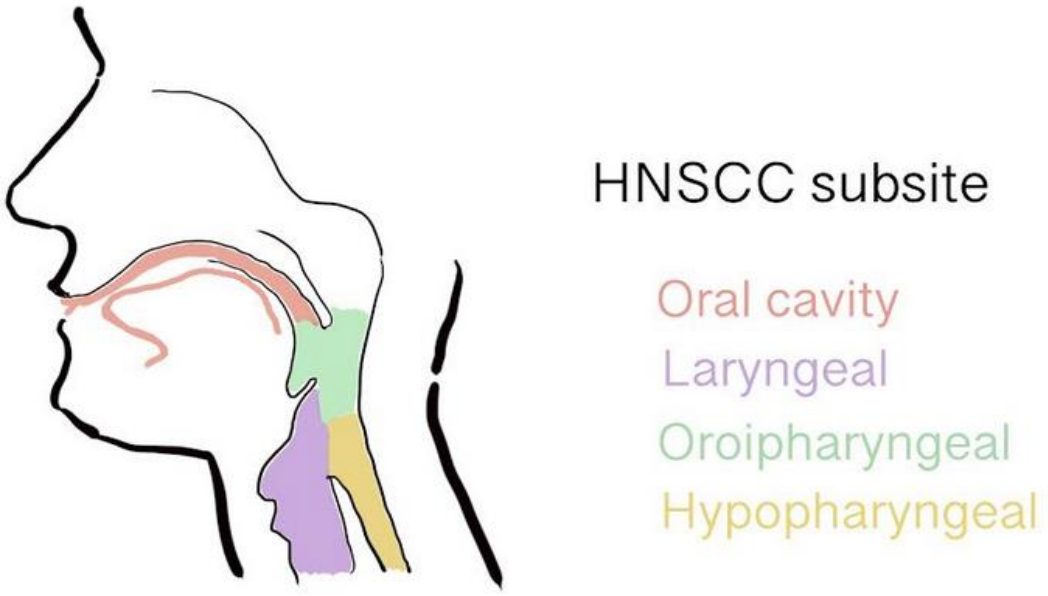


Figure 1. Head and neck squamous cell carcinoma subsites

1.4.1 Oral squamous cell carcinoma

Oral cavity cancer is the most frequent HNSCC, and it is defined as a malignant neoplasia of the oral cavity, including different subsites, such as floor of mouth, lips, upper alveolus and gingiva, hard palate, anterior tongue, and buccal mucosa. The features that can be present are exophytic (i.e., growing off the surface) versus ulcerative (i.e., invading below the surface)[17]. Mostly, the malignance originates from squamous tissues, and for this reason is known as oral cavity squamous cell carcinoma (OCSCC)[18,19]. Typically, the epithelial cells exhibit enlarged nuclei, an increased and abnormal mitosis, and abnormal cells. The number and the distribution of the atypical epithelial cells determine the degree of dysplasia, that can be mild, moderate, or severe[17]. OCSCC are mostly associated with the classical HNSCC risk factors, such as smoking tobacco and alcohol consumption, but also poor oral hygiene and diet lacking vegetables and fruit and rich in animal proteins and fats. In most the cases OSCC is diagnosed in advanced stages, leading to a shorter survival.

1.4.2 Laryngeal squamous cell carcinoma

Laryngeal squamous cell carcinoma is a malignant tumor of the respiratory tract, and it is the second common epithelial tumor of head and neck cancers. It can develop from different anatomical sites, such as supraglottis, glottis, subglottis. As for all HNSCC patients are prevalent male, age ranges between 50-70 years old and risk factors are smoking and alcohol abuse. Frequently, these habits lead to misunderstanding of cancer-related symptoms, inducing late diagnosis of cancer. The overall survival at 5-year highly depends on the staging and anatomical site involved[20].

1.4.3 Oropharyngeal squamous cell carcinoma

Oropharyngeal squamous cell carcinoma (OPSCC) indicates all the cancer that originates from epithelium of throat, tonsils, posterior tongue, soft palate, posterior and lateral pharyngeal walls. Major symptoms are sore throat, odynophagia, voice changes, weight loss and dysphagia. It can be observed ulcer or red/white patch when physical examination is performed.

1.4.3.1 OPSCC HPV negative

OPSCC can be divided in two groups depending on HPV-infection: OPSCC Human Papilloma Virus (HPV)-related, and OPSCC not HPV-related. For the last category the classical HNSCC risk factors (i.e., smoking tobacco and alcohol consumption) have been identified as major risk factors. Less common risk factors are poor nutrition, diet lacking vegetables and fruits, marijuana smoking.

1.4.3.2 OPSCC HPV positive

To date, more than 240 HPV types had been identified. HPV that can infect the human mucosal epithelia are divided in low risk (non-oncogenic) and

high risk (oncogenic). The most oncogenic types are HPV-16, HPV-18, HPV-31, and HPV-33, which are sexually transmitted and are involved in the malignant transformation of infected cells[21]. HPV-positive OPSCC patients differ from HPV-negative because of their age (they usually are younger), they do not have smoking and drinking dependency, they are mostly men and reported to have frequent oral sex with several partners. Notably, HPV-positive OPSCC patients have a better survival compared to HPV-negative OPSCC patients[22,23].

1.4.4 Hypopharyngeal squamous cell carcinoma

Hypopharyngeal cancer account for 3% of all head and neck cancers, and the most involved anatomical sites are posterior and lateral pharyngeal walls. Due to its anatomy, it usually involves lymphatic and vascular systems, frequently leading to metastasizing process. Risk factors are the common risk factors for HNSCC (smoking and alcohol abuse). However, symptoms are uncommon, and the involved sites are difficult to be inspected, making identification of cancer highly challenging and mostly diagnosed at advanced/metastatic state. For these reasons hypopharyngeal carcinoma has one of the worst prognoses of all HNSCC, with a 5-year overall survival rate of 30-35%[24]

1.5 HNSCC Diagnosis

When patients present suspicious lesion or symptoms, the first examination is performed by a surgeon or otolaryngologist and consists in an analysis of the upper digestive tract[17] . For the complete evaluation and diagnosis of the case, the biopsy or fine-needle aspiration of the suspicious lesion and the lymph node is needed. Additionally, imaging of head and neck and lymph nodes is performed, to understand the spread of the disease. The aggressiveness of the disease is evaluated looking at cell

differentiation, lymphatic invasion, lymph node metastasis, extranodal extension through the lymph node involvement[16]. Moreover, because of the distinct biology behind the HPV-positive and HPV-negative OPSCCs HPV testing should be performed for OPSCC cases. After the pathological diagnosis, radiological imaging (including positron emission tomography -PET, computer tomography scan -CT, and magnetic resonance imaging -MRI) is performed to determine the staging and the treatment approach. The stage of head and neck cancer (HNC) is based on the Tumor, Lymph node, Metastasis (TNM) system (Table 1), and for this reason HNSCC are included in head and neck chapter of the American Joint Committee on Cancer Staging (AJCC) manual, at the present at its eighth edition[25]. Early stages (stage I-II) generally include smaller tumors without the lymph nodes involvement. Meanwhile, later stages (stage III-IV) are characterized by invasion of surrounding structures, with the final stage characterized by distant metastasis spreading. Approximately 30-40% of HNSCC patients are diagnosed with stage I or stage II disease, while others are mostly (60-70%) diagnosed at locally advanced stages (stage III or IV) involving loco-regional lymph nodes. Moreover, 10% of patients with locally advanced disease already have distant metastases at initial presentation[26]. Despite improvements in HNSCC management and the aggressiveness of curative treatment, locally advanced disease carries a high risk of local recurrence and distant metastasis (developed in more than 65% of HNSCC patients). Patients diagnosed with recurrence or distant metastasis have a poor prognosis (6 to 9 months in absence of treatment), with a 5-year overall survival less than 50%[27].

T CATEGORY	T CRITERIA
Tis	Carcinoma in situ
T1	Tumor ≤ 2 cm
T2	2 cm < Tumor ≤ 4cm
T3	Tumor > 4 cm or extended to lingual surface of epiglottis
T4	Moderately advanced or very advanced local disease
T4a	Moderately advanced local disease; tumor invades the larynx, extrinsic muscle of tongue, medial pterygoid, hard palate, or mandible
T4b	Very advanced local disease; tumor invades lateral pterygoid muscle, pterygoid plates, lateral nasopharynx, or skull base or encases carotid artery
N CATEGORY	**N CRITERIA**
N0	No regional lymph node metastasis
N1	Metastasis in a single ipsilateral lymph node, < 3 cm and ENE-negative
N2	3 cm < Metastasis in a single ipsilateral lymph node < 6 cm and ENE-negative; or metastases in multiple ipsilateral lymph nodes < 6 cm and ENE-negative; or metastasis in bilateral or contralateral lymph nodes, none larger than 6 cm in greatest dimension and ENE-negative
N2a	3 cm < Metastasis in a single ipsilateral lymph node < 6 cm and ENE-negative
N2b	Metastasis in multiple ipsilateral lymph nodes < 6 cm and ENE-negative
N2c	Metastasis in bilateral or contralateral lymph nodes < 6 cm and ENE-negative
N3	Metastasis in a lymph node > 6 cm and ENE-negative; or metastasis in any lymph node(s) and clinically overt ENE-positive
N3a	Metastasis in a lymph node > 6 cm and ENE-negative
N3b	Metastasis in any node(s) and ENE-positive
M CATEGORY	**M CRITERIA**
cM0	No distant metastasis
cM1	Distant metastasis
pM1	Distant metastasis, microscopically confirmed

Table 1. Staging criteria according to Tumor, Lymph node, Metastasis (TNM) system

1.6 Cancer molecular profiling

During the past decades, genome-wide technologies enabled considerably progresses in the molecular biology of human cancers. Omics technologies are designed for the comprehensive detection, as an example, of DNA (genomics) and mRNA (transcriptomics). The acquired capability of investigating the whole genomes, transcriptomes or other different aspects in a single experiment has allowed an increased tumor profiling and tumor-related mechanism understanding[28,29]

1.6.1 Genomics

For years cancer have been described as a genomic disease. The classic model of carcinogenesis showed the accumulation of genomic variations in somatic cells. Indeed, all the cancer genomes carry mutations. A mutation, for its definition, must affect a gene and the gene alteration consequently must affect, in different ways, the protein function. Generally, it is known that mutations can be germline (occurring in gametes; this kind of mutation could be inherited by progeny) or somatic (occurring in the other cells of the body; this kind of mutation are not transferable through progeny)[30]. Moreover, mutations can affect the chromosome entirely, involving the number of chromosome (increasing the number in case of polyploidy, or decreasing the number in case of haploidy) or can affect the structure of chromosome (by deletion, duplication, inversion, insertion, and translocation events). Mutations can affect DNA sequence too with indels, which are insertion/deletion/duplication of nucleotides. Furthermore, mutations could be point mutations, involving a single nucleotide of DNA, and called single nucleotide variants (SNVs). Point mutations could be classified in three categories:

a) Missense mutation: it is a DNA alteration for a single base, and it has as result the substitution of one amino acid in the protein, the substitution in the protein sequence must affect its function;
b) Nonsense: it is a change in DNA that end prematurely the signal of the end of the protein, it results in impairing the protein function;
c) Frameshift: the addition or removal of a single DNA base alters the reading structure of the gene (for each base corresponds one codon that codification for an amino acid), and the shifts leads to the change of the amino acid encoded.

At genomic level, it is known that mostly two types of genes are responsible and must be involved in the malignant transformation, and they are called proto-oncogenes (then converted into oncogenes) and tumor-suppressor genes. However, complex organisms possess several biological mechanisms for protecting cell's health; for this reason, in order to start the oncogenic process, the accumulation of mutations in these two types of genes is required. In particular, somatic mutations in proto-oncogenes are effective when they make the gene constitutively active. A common analogy compares them to automobile's gas pedal stuck in the acceleration mode. Similarly, the overactivation of oncogenes leads to the exaggeration in encoding proteins which function increase the proliferative ability. On the other hand, tumor-suppressor genes are targeted by somatic mutations in the opposite way, leading to the inactivation of the gene. Remaining in the car analogy, tumor suppressor genes are like brakes for the car. However, the effective result is the same: the induction of a gain for the cell in terms of proliferation capability[30,31]. Advances in next generation sequencing (NGS) technologies gave the opportunity to investigate the whole genome/whole exome of tumor and normal tissue, and to generate an enormous catalogue of cancer-related somatic mutations, by subtracting the relative normal component. It was discovered that cancer genome was highly different between individuals, and that each cancer possessed a peculiar profile[32]. It has been reported that HNSCC develops by a multistep process through well-defined histopathologic phases. However, for HNSCC development and progression few genetic aberrations have been identified, and several candidates are still under investigation, such as mutations on TP53 and EGFR[33,34] . The difficulty to obtain a well-defined genomic profile of HNSCC could depends on its genomic heterogeneity, generated by the variety of anatomical sites involved, and by the different risk factors considered. Besides, the difficulty to establish a precise

genomic model of HNSCC, including mechanism of carcinogenesis and driven alteration, could be also explained by the difficulty of understand them based on genomic profile only. Therefore, to understand the HNSCC complexity additional omic areas need to be taken into consideration [35].

1.6.2 Transcriptomics

In multicellular organisms the same genome can be found in almost every cell. However, not all the genes are transcriptionally active in all the cells and the different gene expression patterns define the type of cell and the relative tissue. During the last twenty years, the introduction of concepts such as alternative splicing, RNA editing, alternative transcription initiation and termination sites and the study of their effects have revolutionized the concept of cancer biology. In contrast to the genome, which appears to be more static, transcriptome changes in response to various cellular stimuli. As for example, the organisms' transcriptome dynamically depends on epigenetic factors and even from environmental conditions. During 1990s early methods to assess the gene expression were Northern blotting and reverse transcriptase quantitative polymerase chain reaction (RT-qPCR). However, these methods evaluated a small and limited number of transcripts. With the introduction of high-throughput technologies, such as microarrays and RNA sequencing, it was quickly possible to investigate at the same time the expression of thousands of genes and their transcript variations. Moreover, in contrast to genomic study, the transcriptome allows the inference of biological activities. Indeed, transcriptomics hold the promise to be an accurate tool for largely detecting the tumor complexity. Studies based on both genomic and transcriptomic provide a better understanding of the structure of cancer genome and the possible mechanisms behind. The introduction of transcriptomic in cancer profiling has the potentiality that allows the characterization and comprehension of different molecular subtypes and

the enabling patients' stratification, crucial aspect for reaching the final goal of personalized medicine. Indeed, in HNSCC several gene expression subtypes have been identified, overtaking the rigid definitions of genomic only study[36–39]. Moreover, several studies have identified and evaluated the presence of biomarkers using gene expression signatures. A gene signature can be defined as a single or a combination of specific expression profiles, evaluating the association of the gene expression with cancer diagnosis, disease prognosis or prediction of the response to specific therapies[40]. A prognostic signature is described as a biologic tool that once measured provides information about outcome and course of the disease. Instead, a predictive signature is a biological tool that once measured provides information about the patient's benefit from a therapy, independently from the prognosis. The accuracy of the gene signature should then be validated in independent cohorts, with different techniques, and by several teams. Gene expression signatures have been included even in clinical trials and preliminary results suggest that transcriptomic analysis could better define responding patients to specific therapies than genomic data alone. However, at date, only few gene expression signatures have been translated in clinical practice for patient management[35].

1.7 Tumor microenvironment

For several years cancer has been described as a disease of proliferating and invasive cells, a mere genetic disease, caused by a graduate accumulation of proliferative advantageous mutations, with the final result of a malignant cell transformation. However, recent evidence showed that cancer can not be considered as autonomous cells with a hyper proliferative profile; rather cancer must be described as a complex disease involving several biological components. The first evidence of the presence of biological elements (not historically considered as "malignant"

or “cancer-related”) in tumor microenvironment (TME) was proposed during the 19th century, when Rudolf Virchow reported the detection of leucocytes in tumor tissue, opening the concept of the balance and link between the inflammation and the cancer development[41]. At date, we know that tumors are complex ecosystems, greatly shaped by TME, which is composed by several non-transformed cell types, such as immune cells (i.e., neutrophils, macrophages, lymphocytes), stromal cells (i.e., endothelial cells, pericytes, fibroblast), and non-cellular components (i.e., extracellular matrix as collagen, fibronectin, and others). The dynamic interaction and the crosstalk between cancer cells and the other cells present in the TME influences growth, tumor progression and invasion, shaping the cancer architecture. As an example, the chronic inflammation, induced in the context of TME, is strictly related to cancer progression and drug resistance and stromal cells promote the cancer invasion by inducing formation of new blood vessels[5]. For these reasons, TME investigation is essential to understand the mechanisms behind cancer formation, response to therapy, drug resistance and to develop new therapeutic strategies[42–44].

1.7.1 Immune components

1.7.1.1 Macrophages

These cells are considered as part of innate immune system and differentiate from monocytes. After their activation, macrophages could be referred as type M1 and type M2. Type M1 macrophages are considered as proinflammatory, and their activation is driven by INF-gamma. They are considered as “anti-tumor” and have been observed in the first stages of tumor progression. Nevertheless, when tumor progresses, TME influences macrophages differentiation in type M2, defined as “anti-inflammatory”[45]. Type M2 macrophages stimulate tumor progression by increasing angiogenesis, proliferation, and epithelial-mesenchymal transition. Moreover, type M2 macrophages (in concomitance with Tregs)

secrete cytokines such as IL-10 and TGF-beta, both inducing potent immunosuppression, impairing the activities of NK, B-cells, and T-cells. A high level of type M2 macrophages is associated with poor prognosis in many oncological diseases. Macrophages, when associated to tumorigenesis are called tumor-associated macrophages (TAM). TAM can secrete chemokines and cytokines (e.g., TGF-β, IL-6, IL-10, and TNF-α), enhancing stemness and promoting epithelial mesenchymal transition[46].

1.7.1.2 Cytokines

Cytokines are small, secreted proteins that, based on their function and structure, can be divided into different superfamilies, including interferons (INFs), interleukins (ILs), tumor necrosis factors (TNFs), transforming growth factors (TGFs), and chemotactic cytokines (chemokines). One of the cytokines roles is to alert of an infection and tissue damage the immune system; however, the signalling pathways of inflammation response are shared with carcinogenesis, and a persistent signalling in the TME can lead to chronic inflammation state, a tumor-supportive immune microenvironment, in which anti-tumor immunity is suppressed. In this state, cytokines are usually overexpressed, and they are regulated by transcriptional and post-transcriptional mechanisms, modified by oncogenic transformation, hypoxia, and other metabolic alterations. Moreover, cytokines impact on anti-tumor immune response, promote proliferation, and influence drug response[43,47]. As an example, interleukin-2 has been shown to be one of the major proinflammatory cytokines produced by T-cells, enhancing their proliferation and cytotoxic response[48,49].

1.7.1.3 Neutrophils

Neutrophils are among the first immune cells recruited during inflammation, and they are involved in different processes, such as

phagocytosis and exocytosis of protease. High levels of neutrophils and high ratio between neutrophil/lymphocytes in the TME of cancer patients have been associated with poor prognosis. Their recruitment in the TME and reprogramming in protumor neutrophils is mediated by TGF-beta, and chemokines, such as CXCL1, CXCL2, CXCL5 [45].

1.7.1.4 NK cells

Natural killer (NK) cells play their role in the innate immune response, having a cytolytic activity in response to transformed cells. Their immunosurveillance is mediated by different stimuli, such as the inhibitory effects that has the major histocompatibility class I (MHC-I), a target receptor presents on normal cells, but a usually a deficiency for cancer cells. When the NK binds the MHC-I present on healthy cells, the NK is inhibited in its function. On contrary, when the MHC-I is missing, NK cells mark the cancer cell as unhealthy and induced them to programmed cell death[45].

1.7.1.5 T-cells

They are part of adaptative immune system, with an inflammatory or anti-inflammatory role. In the primary steps of cancer proliferation, naïve T-cells migrate in TME where they eliminate cancer cells. High levels of T-cell infiltration are associated with a favourable prognosis in several cancer types. Among the several T-cell types, CD8+T-cells are the most abundant against cancer cells, while CD4+T-cells mediate the anti-tumoral response through the secretion of high number of cytokines, such as IL-2, TNF-alpha and IFN-gamma, involving a cascade that recruits NK and macrophages[50,51].

1.7.1.6 B-cells

B lymphocytes are the main components of adaptive immunity response, and although the presence of B cells in the TME has been described in

different carcinomas, the role of B cells in cancer progression is far less investigated and understood than the role of T cells. It has been observed that B cells promote and as well as inhibit the anti-tumor immunity. However, for several different cancer types, the presence of B cells in tumor microenvironment have been primarily associated to a negative outcome, and B cells to immunosuppressive effects[52–56].

1.7.2 Stromal components

1.7.2.1 Endothelial cells

These cells are involved in the building of blood vessels, essential for tumor formation and growth. When they are involved in tumorigenesis, they are usually called tumor endothelial cells; their shape and phenotype change, becoming similar to the tumor itself. Tumor endothelial cells could derive directly from differentiation of cancer cells, or they can be recruited. The most known process in which they are involved is angiogenesis (strictly connected to tumor hypoxia), nevertheless they also promote distant metastasis and drug resistance, impairing even drug delivery. For all their processes they take advantage from various chemokine receptors (CXCR), such as atypical chemokine receptor 1 (ACKR1), ACKR2, CXCR7 and others[44].

1.7.2.2 Pericytes

They are multifunctional cells, and in the context of TME, along with endothelial cells they are involved in angiogenesis process and tumorigenesis. Moreover, pericytes are strictly related to function of immune system, including recruitment of leucocytes from blood vessels and lymphocyte activation. It has been shown that greater amount of pericytes correspond to a better prognosis, while in some tumors it has been seen that pericytes production promotes the growth of cancer. To

date, addressing a better knowledge of pericytes subpopulations is necessary to understand their role in promoting tumor formation[41,44].

1.8 History of cancer therapy

During the last century cancer therapy has evolved. Several progresses have been obtained from 19th century, when X-ray and radiation were discovered, and when in the 20th century surgery started to be studied. From 1930s treating cancer became an object of research, and 90 years later several improvements have been introduced. Nevertheless, the history of cancer treatment showed several ups and downs, and still nowadays more than one approach involves high treatment doses to kill cancer cells (treatment dose that contemporary damages healthy cells). In 1950s the word "chemotherapy" was coined, since 1958 when the first cancer patient was cured with the use of a single chemotherapeutic agent. In 1960s surgery and radiotherapy were introduced as therapy for solid tumors, and promising results for curative intent were obtained in breast cancer patients. Moreover, in 1960s the concept of adjuvant chemotherapy was proposed. In 1978 the combination of cisplatin, and other agents for the cure of metastatic cancer was an innovation, underlying that the possibility to combine drugs and its importance in fighting cancer were related to the different phases of cellular cycle in which cancer cells were passing through. The acquired wisdom enlightened the concept of the essentiality of tumor biological characterization to treat patients as best as it could. In 1980s drugs, specifically targeting cancer biological mechanisms, were tested and in 1990s targeted therapy was introduced. Since 1990, cancer incidence and mortality have been decreasing despite the increasing in population's age. However, what highly changed the paradigm and contributed to cancer therapies were the genetic and molecular biological studies introduced and rising from 1990s. At date, the acquired knowledge

has allowed to understand that effectiveness of treatments highly depends on many individual factors, such as tumor stage, location, as well as patients' overall health, and age. Several personal factors should be considered for selecting cancer treatment, and above all biological profile of the tumors. The last and the future decade of study regarding cancer treatment have been and are going to be dedicated to gain more and more knowledge towards improvements for personalized treatment[57,58].

1.9 HNSCC therapy

For HNSCC treatment decision is based on primary site, stage, tumor histology, human papilloma virus status, comorbid health condition and patient's performance status and a multidisciplinary team is required for the choice evaluation.

1.9.1 Standard-of-Care (SoC) treatment for early stages

For these patients, the treatment modality depends on the accessibility of anatomical site, with the primary aim to minimize the morbidity and preserve the function. Innovative techniques, including robotic surgery, minimally invasive microsurgery, and image-guided radiotherapy are utilized. Surgery and radiotherapy are preferred over other treatments, allowing the total resection of the disease, and reducing the function impairing [59].

1.9.2 SoC treatment for locally advanced stages

For patients with a locally advanced disease a multimodality approach is recommended. Treatment decision for these patients strongly depends on the tumor size and stage, anatomical location, patient's age, and performance status. If possible, surgical resection is preferred, followed by radiotherapy or chemoradiotherapy. When surgery is not applicable, chemoradiotherapy is considered as curative standard. No matter of

previous curative treatment was applied, more than half of patients experience a relapse or a distant metastasis[60,61].

1.9.3 SoC treatment for recurrent/metastatic disease

Patients with local recurrence or metastasis (R/M) can not be treated with surgery, because of the size of the disease or with radiotherapy, since re-irradiating the normal tissue is limited by the toxicity and tolerability. Until recently, for years the SoC was based on platinum with or without cetuximab (anti-epidermal growth factor receptor -EGFR- targeted drug) or targeted therapies alone (cetuximab, methotrexate, and taxanes). These agents were used alone or in combination, depending on patient's age, performance status, symptoms and co-existing condition caused by the prior therapy[62]. Anyhow, no matter which chemotherapy regiment in combination with cetuximab is applied, prognosis for R/M HNSCC remains poor (with a median overall survival after diagnosis < 1 year). Moreover, all these agents are associated with side effects and the response rate is low (from 10 to 13%). Unfortunately, the choice between these systemic therapies reduced the treatment as palliative therapy because regression (when present) was transient and failed to significantly increase patients' survival. The heterogeneous phenotypes of HNSCC made the response to targeted therapies limited, creating an urgent demand for effective new therapies[6364].

1.10 Immunotherapy

From 2016 the oncology practice was drastically transformed by the introduction of immune-checkpoint inhibitors (ICI), particularly regarding anti-programmed death protein-1 (PD-1) antibodies (nivolumab, pembrolizumab) for the treatment of patients with recurrent or metastatic HNSCC[26,65,66] . These agents are now used for both first- e or second-line settings and promising prospects have been shown for the treatment

algorithm for HNSCC. The fascinating aspect of immunotherapy relays on the immune system potentiality to fight cancer alone, enhancing immune system response by blocking suppressive signals through PD-1/PD-L1 (Programmed death-ligand 1) pathway. Indeed, these agents do not target cancer cells directly, but bind receptors/ligands on immune cells, modulating their activity to eliminate cancer cells. Immunotherapy demonstrated the existence of anti-cancer immunity even in HNSCC patients. Undoubtedly, the curative effects obtained by immune-checkpoint inhibitors showed a greater survival benefit than traditional targeted chemotherapy drugs, and ICI has transformed the lives of HNSCC patients. At the present, anti-PD-1 agents have become the current standard of care for management of HNSCC R/M[67,68].

1.11 Major clinical trials using anti-PD-1 agents for R/M HNSCC

Recently, two PD-1 inhibitors (pembrolizumab and nivolumab) have been approved for the treatment of R/M HNSCC, both in first line and second line. Two trials were conducted contemporaneously, such CheckMate-141 and KEYNOTE-012. The first was a phase III randomized trial, comparing nivolumab with chemotherapy (docetaxel, cetuximab or methotrexate)[69]. The study demonstrated an improved survival in patients treated with nivolumab compared to those treated with chemotherapy (7.5 vs 5.1 months, Table 2). Overall, it was observed that patients responded to nivolumab regardless the therapy received before, but those who did not received cetuximab responded better to immunotherapy. Moreover, despite the clear benefit for some patients, the PD-L1 expression made no differences on survival endpoints in CheckMate-141 and the same was observed in KEYNOTE-012. KEYNOTE-012 was a phase Ib nonrandomized trial, in which patients were treated with pembrolizumab[70].

Some of R/M HNSCC patients experienced a durable response, lasting more than 2 years. The study assessed the drug's safety and clinical activity in treatment of R/M HNSCC. As a result, pembrolizumab and nivolumab were approved in 2016 for the treatment of R/M HNSCC following progression on platinum chemotherapy, despite the PD-L1 expression. KEYNOTE-040 was a trial that investigated pembrolizumab to confirm the previous observation of KEYNOTE-012. Also in this trial, it was observed that patients tended to respond better if previously they did not receive cetuximab. Moreover, in the trial it was observed an association between PD-L1 (TPS $\geq$ 50%) and overall survival[71]. More recently, in KEYNOTE-048 was investigate the role of pembrolizumab in first line therapy versus SoC in patients with platinum-sensitive, newly diagnosed R/M HNSCC[72]. KEYNOTE-048 was the only trial to date that investigated the treatment of platinum-sensitive R/M HNSCC: It was a phase III trial that randomly allocated patients with R/M HNSCC to pembrolizumab monotherapy (n=301), pembrolizumab plus platinum and 5-fluorrouracil (n=281) or cetuximab plus a platinum and 5-fluorouracil (n=300). In this trial patients were predominantly male (>80%) and age varies from 20 to 94 years old. Scientists investigated the role of PD-L1 expression (CPS), p16 status (for HPV status), and performance status. They observed a correlation in the population of patients treated with pembrolizumab between PD-L1 (CPS $\geq$ 1 and CPS $\geq$ 20) and Overall survival (OS), but they did not observe a correlation of PD-L1 with Progression free survival (PFS). Chemoimmunotherapy OS resulted superior to standard chemotherapy in all the cohorts regardless CPS score, and response rate were higher in the chemotherapy containing arms compared to pembrolizumab alone. The investigators concluded that pembrolizumab containing arms should be considered as first-line treatments in R/M HNSCC. Pembrolizumab was approved as single-agent for CPS $\geq$ 1 or

chemoimmunotherapy for platinum-sensitive R/M HNSCC patients, not amenable to surgical salvage or radiotherapy[73–75]. Meanwhile, first-line combination of nivolumab plus ipilimumab did not significantly improved OS compared to EXTREME regimen for patients with R/M HNSCC, according to results from the phase 3 Checkmate-651 trial[76].

Clinical trial	**Setting**	**Trial phase**	**Treatment**	**Enrolled patients**	**Median PFS (months)**	**Median OS (months)**	**PD-L1 expression (cut-off considered)**
Checkmate-141	Second line (platinum refractory)	Phase III	Nivolumab vs SoC	240 and 121 R/M HNSCC	2.0 vs 2.3	7.5 vs 5.1	TPS ≥ 1%
Keynote-012	Second line (platinum refractory)	Phase Ib	Pembrolizumab	60 R/M HNSCC	2.0	13.0	TPS ≥ 1%
Keynote-040	Second line (platinum refractory)	Phase III	Pembrolizumab vs SoC	247 and 248 R/M HNSCC	2.1 vs 2.3	8.7 vs 7.1	CPS ≥ 1% TPS ≥ 50%
Keynote-048	Frontline (platinum sensitive)	Phase III	Pembrolizumab vs Pembro + PFE vs Cetuximab + PFE	301 and 281 and 300 R/M HNSCC	2.3 vs 4.9 vs 5.0	11.6 vs 13.0 vs 10.7	CPS ≥ 1% CPS ≥ 20% TPS ≥ 50%

Table 2. Comparison between the major clinical trials using anti-PD-1 agents for R/M HNSCC.

Legend. PD-L1, programmed cell death ligand-1; R/M, recurrent and/or metastatic; PFE, platinum/5-fluorouracil/cetuximab regimen according to the EXTREME trial; SoC, standard of care (methotrexate or docetaxel or cetuximab*)*

Nevertheless, despite the clear progresses obtained introducing immunotherapy in the treatment of HNSCC patients, limitations should be considered. The first and major limitation is that only a small proportion of patients have benefited from this therapy and a low response rate is obtained. An estimated 82-87% of R/M HNSCC patients, platinum-refractory treated in second line, do not have any benefit from these agents, or if any response is assessed, successively it is followed by disease progression or/and death[77]. A first limitation of immunotherapy is linked to the several side effects, comprising also autoimmune adverse events, that could become even life-threatening and highly challenging in clinical practice[78]. The second limitation is the absence of evident predictive tools to assess patients' response, required for stem the immune-related toxicities and the high cost of these antibodies made. Accounting the HNSCC heterogeneity, identify patients that will respond to immunotherapy still remain a challenge. Therefore, the necessity of clear predictive markers is undeniable[79,80].

1.12 Association of clinical variables with ICI response

In the context of HNSCC some clinicopathological variables are under investigation and are here detailed.

1.12.1 HPV status for OPSCC

Two trials, KEYNOTE-012 and Checkmate-141 investigated the possible role of HPV status for assessing Objective Response Rate (ORR) and OS. In the first trial it was observed that patients with HPV infection had an increase ORR compared with those without HPV infection; meanwhile, in

the second trial no differences in terms of response and survival were observed in the two categories. It is known that HPV-positive OPSCCs have a less suppressive tumor microenvironment compared to HPV-negative cases. The predictive role of HPV is still debated.

1.12.2 Smoking status

As described in the literature, smoking is associated with specific genetic signatures and an immunosuppressive tumor microenvironment, and the efficacy of immunotherapy could be compromised by smoking habits[81]. For instance, in Checkmate-141 it was observed that patients with smoking habits had a decreased survival compared to non-smokers, while in other oncological disease, such as lung cancer, smokers were associated with improved response to immunotherapy[82].

1.12.3 Age

Another aspect to consider when we investigate the response to immunotherapy is age, that has been underlying to be essential in anticancer immunity regulation[83]. Specifically, age has been already investigated in melanoma patients and it has been observed that elderly patients had major benefits from immunotherapy compared to younger patients[82]. However, a metanalysis published in 2020 reported the opposite result[84]. No differences in terms of effectiveness of ICIs were observed in Checkmate-141 between the two age groups (cut-off 65 years).

1.13 Biomarkers of immunotherapy

Immune-checkpoint inhibitors have shown a significant and consistent benefit in survival when compared to standard therapies, however the Overall Response Rates (ORRs) ranged from 13-18%. Currently, a solid immune predictive biomarker has not been identified. The finding of predictive biomarker is still an unmet need[85–87].

1.13.1 Tumor-related biomarkers – Programmed death-ligand 1 (PD-L1)

PD-L1, a ligand expressed by various cell types (including tumor and immune cells), was the first biomarker examined and the most widely investigated. Although its expression is currently used as a guide for treatment decision, the expression could vary over time and across multiple tumor types. In some specific oncological diseases, it has been associated with improved response. However, ICIs have demonstrated efficiacy regardless PD-L1 expression. PD-L1 is commonly assessed by Immuno histochemistry (IHC), and various cut-off are under investigation, such as tumor proportion score (TPS), defined as the fraction of viable tumor cells showing membrane staining at any intensity, and combined positive score (CPS) defined as the percentage of tumor and inflammatory cells within and near the tumor expressing PD-L1. PD-L1 is considered as a surrogate biomarker of T-cell infiltration, resulted from an upregulation of INF- γ. In three different HNSCC trial, investigating ICI treatment, it has been seen that PD-L1 expressing tumors tend to have an improved response rates to ICI therapies in KEYNOTE-040 and KEYNOTE-048. However, CHECKMATE-141 failed to show a significant correlation between PD-L1 expression and response. This last trial suggested that PD-L1 expression may help predict the clinical benefit of the ICI treatment, however patients that do not express PD-L1 should not be precluded from the therapy. Results discordance on the relevance of PD-L1 could be explained by the variability of the antibody assays and cut-off levels used, timing of the testing, by the variability of cells that express PD-L1 (such as tumor or/and tumor-infiltrating immune cells) and by which scoring is considered for PD-L1 assessment[85]. Therefore, to understand the real contribution of PD-L1 in predicting the response, it must be aggregated

with additional factors that may contribute to treatment response, such as tumor immune infiltration and Tumor mutational burden (TMB)[87,88].

1.13.2 Tumor-related biomarkers – Tumor mutation burden

TMB is the total number of non-synonymous somatic point mutations per coding area of a tumor genome (somatic mut per megabase - Mut/Mb). The idea behind the possible correlation between TMB and immunotherapy is based on the hypothesis that TMB should reflects the cancer mutation-derived neoantigens[89,90]. Consequently, neoantigens, if expressed and processed, may increase the formation of novel peptides, resulting in higher tumor immunogenicity, and activating immune cells, such as CD8+ T-cells. The higher the number of somatic nonsynonymous mutations, the higher should be the number of neoantigens generated that can be recognized by the immune system. Mostly, TMB have been evaluated by whole exome sequencing (WES) and various cut-off have been taken in consideration. Nowadays, several sequencing panels (>300 genes detected) have been developed to measure TMB, limiting costs, required DNA input, and reducing the time consuming[91,92]. Thank to next generation sequencing techniques it has been possible to investigate TMB in hundreds of patients, and a significant relationship between high TMB and improved responsiveness to ICI has been observed in various cancers, such as HNSCC, melanoma and lung cancer[66,93]. In 2017, the US Food and Drug Administration (FDA) approved the use of ICI for cancer patients with a TMB > 10 (mutations/megabase: mut/Mb), defined as TMB-high[94,95]. As for PD-L1 high TMB alone does not guarantee response to immunotherapy and patients with a low TMB could benefit from the therapy too[96]. Studies have shown that patients with TMB-high displayed also a microsatellite instability. These correlation between these two parameters seems to be a possible predictive factor for response to checkpoint inhibitors in cancer[97,98].

1.13.3 Tumor-related biomarkers – Microsatellites instability

Microsatellites are short DNA regions (that usually are 1-6 nucleotides long), which are tandemly repeated through all the genome, in both intronic and exonic regions. Microsatellites are frequently present also in promoter regions, and terminal regions. Microsatellite instability (MSI) is an event that usually occurs when a microsatellite site gains or loses 1 or more repeats. It is usually connected with issue related to the functionality of DNA repair gene systems. Above all, DNA mismatch repair (MMR) system is the most involved in repairing DNA replication errors, including microsatellite loci. The genes coding proteins for MMR functionality are MLH1, MSH2, PMS2, and MSH6[99]. When these genes present mutations the direct consequence is the impairment of MMR system and, defects in MMR system could be observed and associated with the increased number of instable microsatellite regions. Tumors that present a high microsatellite instability are named MSI-High (MSI-H), while those that do not present microsatellite instability are label as MSI-Stable (MSI-S). The association between possible benefit from immunotherapy and DNA repair system and microsatellites instability is given by the observation in colorectal cancer, in which the deficiency in MMR system were correlated to an increased number of neoantigens. The increased number of neoantigens could be related to more immunogenic tumors, thus, to be more likely to respond to immunotherapy[100]. In routine MSI is traditionally analyzed with PCR (MSIPCR) and immunohistochemistry[101]. limited to five microsatellite markers recommended by National Cancer Institute[102]. Nevertheless, a greater number of microsatellite loci through DNA sequencing could allow a better identification of MSI profile in different cancer types.

1.13.4 Tumor-related biomarkers – Damage response and repair

Damage response and repair (DRR) gene alterations in DNA are associated with higher genomic instability, that could be also inferred studying the MSI and TMB. Mutations in DDR genes have been proven to be associated with higher immunity in some cancers and with treatment benefits in patients treated with immune-checkpoint inhibitors[103]. For example, DNA mismatch repair (MMR) is a biological system that is used by cells for identifying and repairing DNA variations that occurs during replication. Uncorrected DNA replication can produce numerous aberrant neoantigen and deficient tumor cells in MMR could carry from 10 to 100 times more mutations than a tumor cell with a proficient MMR. Particularly, mutations in MMR-related genes, such as MLH1, MSH2, MSH6, and PMS2 impair efficacy of the system and are argument of study[96,104,105].

1.13.5 Biomarkers of the overall status of tumor microenvironment

Tumor infiltrating immune cells seems to have an important role on treatment response to immunotherapy. For instance, T-cells significantly correlated with a better outcome in different cancer types and could predict the efficacy of immunotherapy. On the contrary, other immune cells such as T regulatory cells (Tregs) and tumor-associated macrophages can be an indicator of immunosuppressive environment and consequently be linked to a poor response to ICI[106]. Likewise, Interferon-γ (IFN-γ), a cytokine with antiviral, antitumor and immunomodulatory function, able to trigger the activation of the immune response and in the meanwhile to prevents the overactivation of the immune system is now considered object of study for biomarkers evaluation. For instance, IFN-γ plays a vital role in enhancing CD8+ T-cells activity against tumor cells and otherwise

upregulate PD-L1 expression on tumor cells, that efficiently protect them from immune surveillance. Several interferon-γ gene expression signatures (built up through the study of genes associated with T-cell activation) have been proposed in different cancer types[107–109].

1.13.6 Immune-related and TME-related gene expression signatures

Gene expression profiling allows the assessment of various inflammatory-related markers contemporaneously. Several studies investigated the role of immune-gene signatures to predict prognosis or response to immunotherapy in different cancer types, including for example Interferon-γ and T-cell gene expression signatures in HNSCC[110]. Moreover, the complexity of tumor microenvironment, which is known to play a critical role, was considered. However, the intricacy of biological mechanisms results difficult to manage by scientists. A possible way to reduce such methodological complexity could be offered by gene expression analysis. For instance, gene expression of tumor cells and tumor microenvironment infiltrating cells is under investigation as biomarker for the efficacy of immunotherapy. TME could be classified into three different distinct phenotypes based on immune infiltration, which are "inflamed", "immune-exhausted" or "immune-desert"[111]. The large availability of multi-omics data generated by Pan-cancer studies has facilitated the dissection of molecular principles that made possible the inferring of biological characteristics. Gene expression signatures has been proven to be a relevant surrogate tool to interrogate the complex interactions and mechanisms behind the oncogenic processes, allowing the possibility to reduce the methodologies used for the tumor biology evaluation (such as IHC, FACS and other techniques). On the other hand, gene expression signatures gave the opportunity to interrogate an

additional manageable tool, easily translatable to clinical practice. At date, several gene expression signatures are under investigation, comprising de-regulation in immune and stromal infiltrated cells and DNA repair systems[112,113].

2. AIMS

Prognosis for R/M HNSCC patients is dismal, and second line treatment options are limited. Since the approval in 2016 of the use of immunotherapy (single agents, i.e., nivolumab, pembrolizumab) as SoC for these patients, the paradigm has changed. However, response rate is low (13-18%) and association of the outcome with clinical, demographical variables and predictive biomarkers are still under investigation. The present PhD project took advantage of the clinical trial "Nivactor", in which R/M HNSCC platinum-resistant patients were treated in second line with nivolumab. Among the "Exploratory objectives" of the trial there was the identification of pre-existing biological markers associated to nivolumab response potentially useful as predictive surrogate markers.

The tumor tissue of these patients was profiled, and by the analysis of gene expression and mutations in specific cancer-related genes, we planned to investigate:

a) The differences between responding and non-responding patients;
b) The role of specific markers (such as PD-L1, TMB and MSI) highlighted in other clinical trials;
c) The association of specific gene expression signatures with response and survival;
d) The contribution of specific mutation to response and survival.

Through this extensive tumor profiling, we aimed to define specific molecular features able to guide the clinical practice through a more personalized approach.

3. MATERIALS AND METHODS

3.1 Clinical trial

The present Ph.D. project took advantage of the clinical trial "Nivactor": a single-arm, open-label, multicentre Phase IIIb clinical trial, with NIVolumab in subjects with recurrent or metastatic (R/M) platinum-refrACTORy squamous cell carcinoma of head and neck cancer (EudraCT Number: 2017-000562-30). The study was performed in compliance with the requirements of the AIFA. The study was full approved from Comitato Etico Fondazione IRCCS Istituto Nazionale dei Tumori Milano on 4 Oct 2017 (INT 128_17). Patients were enrolled from 21 different Italian centres, following the inclusion and exclusion criteria. All patients provided written informed consent to participate in the study.

3.2 Inclusion and exclusion criteria

Patients were included in the study only with confirmed recurrent or metastatic HNSCC (oral cavity, oropharynx, hypopharynx, larynx), not amenable to local therapy with curative intent (surgery or radiation therapy with or without chemotherapy), male and female, with an age ≥ 18 years and an Eastern Cooperative Oncology Group (ECOG) performance status ≤ 2. Moreover, the presence of tumor progression or recurrence must have occurred within 6 months of last dose of platinum therapy in the adjuvant (i.e., with radiation after surgery), primary (i.e., with radiation or prior to it or to surgery as induction chemotherapy), recurrent, or metastatic setting. Patients were included only if the disease could be measurable by Computed tomography (CT) or Magnetic resonance imaging (MRI) per Response Evaluation Criteria in Solid Tumor (RECIST) 1.1 criteria[114] . On the contrary, patients were not considered for the study with histologically confirmed recurrent or metastatic carcinoma the nasopharynx, and salivary gland or non-squamous histology (e.g., mucosal melanoma) and patients

prior treated with anti-PD-1, anti-PD-L1, anti-PD-L2, anti-CTLA-4 antibody, or any other antibody or drug specifically targeting T-cell co-stimulation or immune checkpoint pathways. Moreover, patients were not considered for the trial if they had an active, known, or suspected autoimmune disease, known history of testing positive for human immunodeficiency virus (HIV) or known acquired immunodeficiency syndrome (AIDS).

3.3 Nivolumab treatment

Subjects received treatment with nivolumab monotherapy at 240mg flat dose on Day 1 of treatment cycle and every 14 days, until confirmed progression of disease, unacceptable toxicity, death, or withdrawal of consent. To monitor the disease, CT scans and/or MRI were performed. The treatment efficacy was assessed following RECIST 1.1, a set of published criteria used in the definition of tumor changes through imaging, allowing the definition into the category of "responder" or "partial responder" (when the disease reduces or disappears), "stable disease" (when the disease stays the same), "progress disease" (when the disease worsen). Clinical information and demographic data (such as age, gender, tumor subsite, tumor stage, treatment history, and others) were collected and update till the end of 2020.

3.4 Tumor specimen collection and PD-L1 evaluation

Tumor biopsy specimens were obtained from patients prior to treatment with nivolumab from: a) metastatic or recurrence setting; b) an archived biopsy of the primary tumor. All the tumor specimens were collected at IRCCS Istituto Nazionale dei Tumori (Milan, Italy). A minimum of 1 formalin-fixed paraffin embedded (FFPE) tumor tissue block or a minimum of 10 FFPE unstained sections were required for assessment of PD-L1 status and other biomarker evaluation. PD-L1 IHC was performed

utilizing PD-L1 IHC 22C3 pharmDx kit (Dako, Carpinteria, CA) on the Dako ASL48 platform, according to manufacturer recommendations. The activity was centralized at IRCCS Istituto Nazionale dei Tumori (Milan, Italy). PD-L1 expression was evaluated both in tumor cells and in inflammatory cells. TPS was defined as the percentage of tumor cells presenting PD-L1 membranous immunoreactivity at any intensity. The CPS was defined as the number of PD-L1 staining cells (tumor cells, lymphocytes, macrophages) divided by the total number of viable tumor cells, multiplied by 100. A minimum of 100 viable tumor cells had to be present in the PD-L1 stained slide for the specimen to be considered adequate for PD-L1 evaluation. Tumor infiltrating lymphocytes (TILs) were assessed on hematoxiline & eosine (H&E) slides, as described in Salgado et al[115] . Mononuclear inflammatory cells (i.e., lymphocytes and plasma cells) present in the stromal compartment were assessed. The percentage of stromal TILs was calculated as the area of stromal tissue (within and at the invasive edge of tumor area) occupied by inflammatory cells over the total stromal area.

3.5 Nucleic acid extraction

Pathological revision was performed to obtain adequate non- necrotic tumor areas free of contamination from normal tissue, and after, macrodissection was done on sections to obtain at least a percentage > 70% of tumor cells. Nucleic acid extraction on FFPE pre-immunotherapy tumor samples (primary or recurrence/metastasis specimens) and the following related activities were all centralized at IRCCS Istituto Nazionale dei Tumori (Milan, Italy). RNA was isolated using Qiagen RNeasy FFPE kit (Qiagen), in accordance with the manufacturer's instructions. To avoid possible bias of lack of uniformity of nucleic acid material, RNA quality and quantity were assessed by 4200 TapeStation system (Agilent

Technologies), and Qubit 4.0 Fluorimetric RNA Broad Range or High Sensitivity Assay (Thermo Fisher Scientific), respectively. Samples that did not meet the standard criteria for quality (DV200 < 15; i.e., DV200 is the percentage of RNA fragments that are >200 nucleotides in size) and quantity (concentration < 10 ng/µL) were not considered adequate for gene expression profiling. For mutational analysis, DNA were extracted using GeneRead FFPE kit (Qiagen), following the manufacturer's instructions. After nucleic acid quality check and quantification by 4200 TapeStation (Agilent) and Qubit 4.0 Fluorometer (Thermo Fisher Scientific) samples. Samples with a poor quality (DIN < 2) nucleic acid, or a low quantity (< 7 ng/µL) were not considered for DNA sequencing experiments.

3.6 Gene expression experiments

Gene expression experiments were performed accordingly to GeneChip WT Pico standard protocols (Affymetrix, Thermo Fisher Scientific). For the protocol execution, 100 ng of RNA was used as starting input; RNA was retrotranscribed with "First-Strand Master Mix" and incubated in a thermal cycler for 1 hour at 25°C, then for 1 hour at 42°C, then for at least 2 minutes at 4°C. To clean up the reagents from the previous reaction, the samples were incubated for 30 min at 37°C, for 10 minutes at 80°C and then for at least 2 min at 4°C with "WT Pico Cleanup Reagent" in the thermal cycler. To the single-stranded cDNA 3' adaptors were added, utilizing the "Adaptor Master Mix" (thermal cycler program: 2 min at 95°C, 9 cycle of 30 sec at 94°C and 5 min at 70°C and then for at last 2 min at 4°C). Second strand cDNA was synthetized using "Second-Strand Master Mix", following the incubation for 1 hour at 16°C, then for 10 minutes at 65°C, then for at least 2 minutes at 4°C. Antisense RNA (complimentary RNA or cRNA) was synthesized and amplified by in vitro transcription (IVT) of the second-stranded cDNA template using T7

RNA polymerase (thermal cycler incubation at 40°C for 14 hours). The purification of cRNA was performed utilizing “Purification Beads” and washed three times with EtOH 80%. Samples were eluted in 27 µL of nuclease free water, previously heated at 65°C. For the synthesis of the sense-strand cDNA 833 ng/µL of cRNA (in a final volume of 24 µL) were utilized, following the incubation reaction in the thermal cycler (10 min at 25°C, 90 min at 42°C, 10 min at 70°C and then for at least 2 min at 4°C). RNase H was then utilized to remove the residual cRNA template incubating the samples for 45 min at 37°C, 5 min at 95°C and then for at least 2 min at 4°C. The RNase H activity was stopped adding 11 µL of nuclease free water. The purification of cDNA was performed utilizing “Purification Beads” and washed three times with EtOH 80%. Pre-heated water (65°C) was utilized for elution. Samples with a concentration > 120 ng/µL were used for the fragmentation step. It was utilized 120 ng/µL of ss-cNDA, equal to 5.5 ug in 46 µL. Sense-strand cDNA was fragmented by uracil-DNA glycosylase and apurinic/apyramidinic endonuclease 1 at the unnatural dUTP residues and breaks the DNA strands. Moreover, the fragmented cDNA is labeled by terminal deoxynucleotidyl transferase using “DNA labeling reagent” that was covalently linked to biotin. Probes were hybridized on human Clariom D chips for 16 hours at 45°C; after chips washing and staining, the chips were scanned with Affymetrix Gene Chip Scanner 3000 7G. The Affymetrix system was designed to detect genes, exons, and alternative splicing events from >540,000 transcripts. Primary data were acquired using the Affymetrix GeneChip Command Scan Control version 4.0 (developed by Thermo Fisher Scientific). The generated CEL files were analyzed for an additional quality check using Affymetrix Expression Console Software (version 1.4), which normalized array signals using Signal Space Transformation (SST) and a robust multiarray averaging (RMA) algorithm.

3.7 TSO500 NGS library preparation

The experiments for the target sequencing were performed using Illumina TruSight Oncology 500 protocol (Illumina, UK). The Illumina TruSight Oncology 500 panel (TSO500) allows the detection of somatic variants in 523 cancer related genes; moreover, it generates a score for TMB and MSI calculation. Before library preparation, DNA was evaluated and restored using Infinium FFPE QC and DNA restoration kit (Illumina), following the manufacturer's instructions. FFPE samples are known to generally yield highly degraded DNA, they typically perform poorly in whole-genome genotyping. For this reason, after DNA extraction, we used the Illumina FFPE QC Kit, that through real-time PCR assessed the quality and the integrity of DNA for each samples. Each FFPE samples that pass the QC test was eligible for restoration, using the Infinium HD FFPE restore kit, following the manufacturer' instructions, which is able to repair the degraded DNA samples.

For library preparation, 80 ng in 12 µL of DNA were used as starting material from each sample. The library preparation was performed manually according to manufacturer's instructions. The very first step of the protocol is based on mechanical fragmentation (Covaris *E220evolution*) of the genomicDNA (gDNA). After the fragmentation the generated dsDNA were checked in size (90-250 bp) using the Tapestation 2200 (Agilent Technologies, UK). Then, the ends are repaired using A-tailing master mix. Unique molecular identifiers (UMIs) adapters are ligated to the fragments, and then a clean-up is performed. UMI are complex indexes added to the fragments before the amplification to reduce the number of false positive calls, to have a more accurate result during the sequencing process (Figure 2).

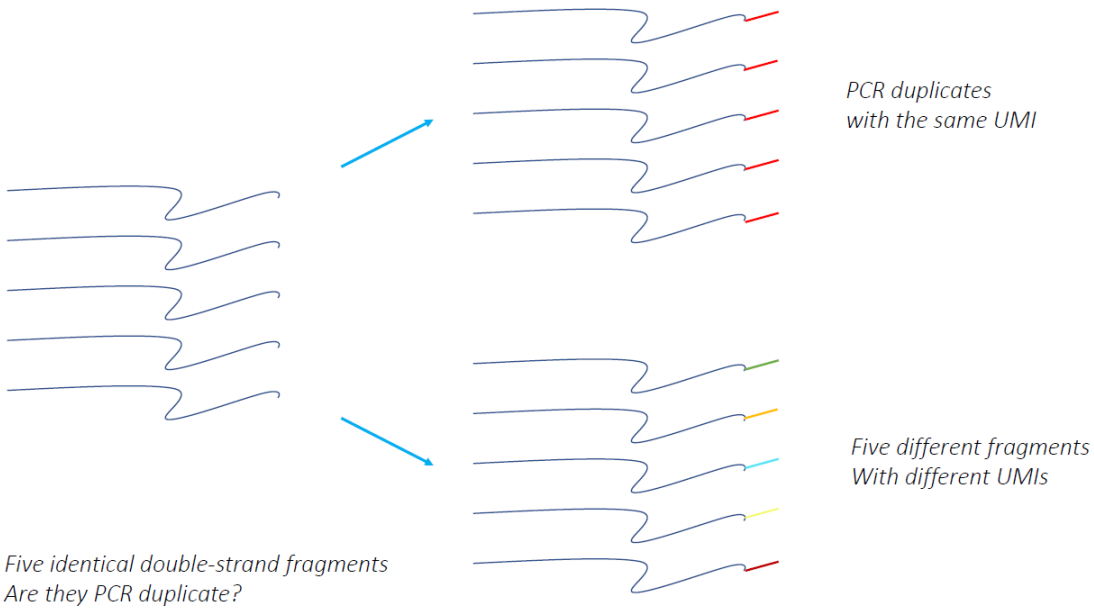


Figure 2. Functional mechanism of UMIs

Library fragments were amplified, and index sequences are added, to allow the multiplexing of samples. This step was fundamental for cluster generation during sequencing. Libraries were then hybridized overnight with a pool of oligo specific to the 523 targeted genes. Probes were captures and cleaned-up with streptavidin magnetic beads. Library quality and quantity were checked before the creation of the pool, using qubit Fluorometer (Thermo Fisher Scientific) and Tapestation 2200 (Agilent Technologies). Only library with the specific peak were considered [240bp-290bp], and library with a concentration < 1 ng/µL were excluded. Library nano molarity was calculated using the following formula:

$$\mathbf{nM = ([qubit] / 660*size\ bp) *10^6}$$

At the end of the process the pools are created and sequenced. Pools were run on NextSeq 550 instrument (Illumina). Libraries were multiplexed for sequencing with up to 8 DNA libraries for each run. The Sequencing

results were analyzed using TSO500 Docker pipeline. Possible sequencing PCR bias were removed thanks to the use of UMIs.

3.8 TSO500 NGS library analysis

The sequence alignment to the human genome (hg19) was completed using SOPHiA DDM™ for TSO500 pipeline. Moreover, for the analysis the following threshold were considered:

i) "Variant reads" is the number of independent sequence reads supporting the presence of a variant. Due to the high error rate of NGS at the per-base call level, calls supported by less than 5 variant reads are typically considered to be likely false positive calls. We considered as minimum variant reads = 10;

ii) "Variant allele frequency (VAF)" is the percentage of sequence reads observed carrying a specific DNA variation divided by the overall coverage obtained at that precise locus. Thus, VAF is a surrogate measure of the proportion of DNA molecules in the original specimen. The number ranges between 0.0 to 100.0. VAF could be interpreted as measure of diploid zygosity in germline DNA sequencing, in which heterozygous loci have a VAF = ~50% VAF and homozygous loci a VAF = ~100%. However, for somatic testing in clinical setting, in which even variants with low VAF could have an important role (and high level of sensitivity is required) the combination of normal and tumoral DNA in the sample causes heterogeneous VAFs. In cancer context, VAF analysis could not be considered accurate because intratumoral heterogeneity and impurity of tumor DNA cause confusing deviations from expected VAFs. Moreover, polymerase and sequencing errors make it difficult to robustly detect low-frequency mutations <5% VAF. For this reason,

UMIs were developed to suppress the errors to detect mutations even below 1% VAF. Nonetheless, accurate detection using UMIs with a low VAF (< 1%) is possible only with large amount of total DNA (> 250 ng) and they were not considered for the present analysis[116]. On the other hand, high VAF percentages (> 50%) were included, because even if it could indicate that the variant is germline or in a region of loss of heterozygosity, the pairing with normal tissue was not performed to confirm the assessment. Also, even if the tumor percentage in our tissue sample was selected > 70%, it is important to underly that since percent tumor cell is a prediction performed by pathological/histological analysis, and it is never intended to be an accurate measurement. Additional, even the VAF in the "grey" zone were included in the analysis, because it is common observing true cancer mutation in this range[117]. To conclude, for all these reasons, with considered all the VAF > 2 %.

iii) Coverage, usually indicated with a number followed by "×", is the number of independent reads with overlapping alignment at a locus of interest. This is often expressed as an average or percentage exceeding a cutoff over a set of intervals (such as exons, genes, or panels). To assess proper parameters the SOPHiA DDM™ for TSO500 pipeline generated a specific output, allowing the elimination of all those variants with low coverage or high background noise.

Moreover, only INDELS and SNPs were included, filtering all the variations that were nonsynonymous, and so including only the "missense", "nonsense", "frameshift" annotations. The allele frequencies in control populations could be useful for assessing clinical significances of somatic variants[118]. For this reason, we excluded all the variants with a

prediction of frequency > 3% in the population using the control population of 1000 Genome Project[119]. The distinction between somatic and germline alterations with the somatic only next-generation sequencing panel was not addressable[120]. For the assessment of somatic prediction for specific considered mutations, the Catalog of Somatic Mutations in Cancer (http://cancer.sanger.ac.uk/cosmic) was utilized, where are annotated millions of somatic alterations across different tumor types. For the germline prediction of the DNA variations instead ClinVar database (http://www.ncbi.nlm.nih.gov/clinvar) was utilized. ClinVar database contains only germline variants, both pathogenic and benign, providing in addition clinical evidence when available[121]. Data were visualized by Maftools, a R package developed to analyze and visualize NGS data[122].

3.9 Tumor mutational burden

Tumor mutational burden (TMB) is defined as the number of non-synonymous (missense, nonsense, frameshift) mutations within 1 Mb (mut/Mb) of coding region on tumor genome. TMB score was generated using SOPHiA DDM™ for TSO500 pipeline. For the evaluation of TMB in the present analysis, either single-nucleotide variants and small insertion/deletions in coding regions with a minimum coverage of 50X and 5% ≤ variant allele frequency ≤ 90% were considered. Contrary, multi-nucleotide variants (MNVs) were excluded. Only eligible somatic mutations per Megabase (Mb) were considered, after filtering germline variants (in-house SOPHiA GENETICS database). Even if TMB is historically assessed in whole exome sequencing studies, Illumina demonstrated that TMB can be effectively estimated using targeted sequencing panels covering at least 1.1 Mb of genomic content. The TruSight Oncology 500 panel covers 1.94 Mb genomic content and take advantages of UMIs, which reduce the sequencing noise during the initial

steps. This addition of UMIs is a fundamental step, which strengthens the measurement of TMB. Tumor samples were classified as TMB-high or TMB-low using 10 mut/Mb as cut-off value, as suggested by Illumina Pipeline.

3.10 Microsatellite instability

Next generation sequencing (NGS) of 175 noncoding homopolymer regions was performed for assessing the microsatellite instability. MSI calculation was performed by SOPHiA DDM™ for TSO500 pipeline. Regions with low sequencing coverage (minimum coverage > 50X) were not considered in the MSI status calculation, and at least 20 well-sequenced sites were required to be assessed to determine the MSI score. Samples were divided in: a) microsatellite stable (MSS) if the score is less than 0.013; b) For scores above 0.018, the sample is considered as ”unstable” with high confidence and marked as ”MSI-H” (MSI-High); c) “reject” if the percentage of homopolymer loci used is less than 70% of the loci sequenced.

3.11 Differentially expressed genes (DEGs)

The investigation of differential expressed genes between the responders versus non responders, clinical benefit versus non-clinical benefits categories was performed using limma R package[123] . Only genes with the absolute value of log2FoldChange > 1and adjusted $p < 0.05$ were considered as significantly deregulated transcripts. The p values were adjusted on the false discovery rate (FDR) according to the Benjamin-Hochberg method. Volcano plot was utilized for visualizing the results, using R software ggplot2[124].

3.12 Functional enrichment analysis

To explore the de-regulated biological pathways, the gene-set enrichment analysis (GSEA) Hallmark pathways[125] was applied. The hallmark gene sets (n=50) represent coherently expressed signatures derived by aggregating many MSigDB (i.e. Molecular Signature Data Base, https://www.gsea-msigdb.org/gsea/msigdb/) gene sets to represent well-defined biological states or processes. Genes were ranked using the t-statistic, the number of permutations was set to 10,000 and gene sets with fewer than 15 or more than 500 genes were filtered out. The comparison was performed using two categories or using scores in continuum. We distinguished significantly enriched biological pathways setting as screening criteria a false discovery rate (FDR, correction for multiple comparisons) q-value < 0.05 and normalized enrichment score (NES) > |1.5|. GSEA was run using the Java desktop application and GSEA v4.2.2 for Windows. Text files with the following specific formats were uploaded for the analyses: i) GCT format for the gene-expression matrix; ii) CLS file format to define phenotype labels (classes or continuous traits). The differential pathways analysis was calculated by limma package[123].

3.13 Tumor microenvironment composition

For the evaluation tumor microenvironment components, we used on gene expression normalized data the xCell method[126], a deconvolution algorithm that is based on gene expression, and allows through an extensive in silico analysis, the inference of 64 immune and stromal cell types and 3 summary scores (i.e., immune, stroma and microenvironment scores). The 64 cell types were categorized into: lymphoid, myeloid, stromal, and others. The comparison was performed using two categories or, when available, using scores in continuum.

3.14 Statistical analysis

Clinical endpoints were Overall Survival (OS) and Progression Free Survival (PFS). OS was defined as the time interval between the date of the first diagnosis and date of death from any cause or the date of the last follow-up. PFS was defined as defined as the time between date of randomization and the date of radiological evidence of progression or significant clinical symptomatic progression or of death without evidence of progression, whichever occurred first, or the date of the last follow-up. Objective response rate (ORR) was assessed according to Response Evaluation Criteria In Solid Tumors (RECIST criteria, Version 1.1). Last follow-up data were updated in 2020. Survival curves were estimated using Kaplan-Meier method and statistical differences between categories were assessed by log-rank test, using R packages "survival" [127], "survminer"[128] or by the use of the online tool "Statistic Kingdom" (https://www.statskingdom.com/kaplan-meier.html).

Patients were divided in two groups by specific parameters, explained in the reference or calculated (see M&M chapter 3.16). The Fisher's test or chi-square statistic test was used for comparing categorical data of two or more groups, respectively. T-test or one-way analysis of variance (ANOVA) for quantitative (numerical or continuous) data comparing two or more groups (online tool "Aatbio" https://www.aatbio.com/tools/). The statistical significance level was set for all the comparisons at adjusted $p < 0.05$.

3.15 Gene expression and mutational signatures

For the catalogue of gene expression and mutational signatures literature research was conducted on Embase and PubMed, using the key words ("head and neck cancer" OR "head and neck squamous cell carcinoma" OR "HNSCC") AND ("gene signature" OR "gene list" OR "mutational

signature" OR "gene expression signature" OR "signature" OR "mutational profile" OR "gene expression profile") AND ("immunotherapy" OR "nivolumab" OR "pembrolizumab" OR "ICI" OR "ICIs" OR "immune checkpoint inhibitor"). Publications unavailable in English were excluded. The signatures were tested following the methods of each specific publication (see Results for the list of signatures). When gene weight was not available, we used the ssGSEA score for each gene in the gene list. The gene lists related to signatures/pathways were retrieved from primary literature sources (i.e. tab delimited text files, microsoft word tables or pdf files). Gene ID were reannotated to include the official HUGO Gene Nomenclature and EntrezID/RefSeq from NCBI Gene (https://www.ncbi.nlm.nih.gov/gene). When the method for testing the gene expression signatures was not available the publication was excluded from the analysis. For dichotomizing the gene expression score in order to perform the survival analyses, we utilized for each signature the Cutoff Finder R package method[129], which allow the biomarker cut-off determination. The present method allows to optimize the cut-off point taking into account the presence of a complex distributions of data. The biology behind the gene expression was instead performed using the score in continuum, correlating the biological marker evaluated. For the evaluation of mutations in repair systems related genes we followed the mutated gene lists published by Hsiehchen et al.[130].

4. RESULTS

4.1 Nivactor study population

The first evaluation was made considering the complete picture of Nivactor, in order to corroborate the clinical validity of the trial and then proceeding with translational analyses. From November 2017 to July 2018 a total of 127 R/M HNSCC patients with platinum-refractory unresectable R/M HNSCC were enrolled in Nivactor phase IIIb trial. The enrolment closed at the end of March 2019 and the follow-up phase was concluded in March 2020. However, out of the 127 patients enrolled, 124 R/M HNSCC patients were treated with anti-PD-1 antibody (nivolumab) and 3 were excluded for screening failure. Of the 124 treated patients we observed that patients were balanced for age groups at the first nivolumab dose (patients with age ≥ 65 years were 61/124 patients, corresponding to 49%; patients with age < 65 years were 63/124 patients, corresponding to 51%. The median age was 64 years [31-84] Table 3). A prevalence of male (N=101; 81%) than female (N=23; 19%) was found in the total cohort, prevalence already exhaustively reported in literature for HNSCC. The subsite of the primary disease in the total cohort was oral cavity for 43/124 patients; oropharynx for 28/124; larynx for 30/124; hypopharynx for 18/124; not annotated for 5/124. For oropharyngeal cases, HPV status was tested in 18/32 cases. The response was assessed by RECIST v1.1 and it was available for 122/124 the patients: 2 complete responders (CR, 1.6%); 17 partial responders (PR, 13.78%); 22 patients with stable disease (SD, 17.7%); 81 patients with progressive disease (PD, 65.3%). For 2 patients the response was not available (N.A., 1,6%). Objective response rate resulted in line with the literature.

Successively we compared the characteristics of responders versus non-responders. First, we did not observe differences in terms of median age

(Table 3). Smoking status was annotated for all the patients excepting 6 (for 1/6 patient response was not available). Notably, a significant prevalence of present smokers was observed in non-responders, while past smokers were significantly more abundant in responders (p=0.012, Chi-square test). Moreover, in responders we observed a significant increase in oropharyngeal cancers (specifically HPV-positive) as primary site of disease, and a decrease in the other subsites (except hypopharynx). In non-responding patients an abundance of oral cavity cancer was observed. The other clinical-pathological characteristics resulted balanced between the categories of responders and non-responders. The median overall survival (months) observed for the total cohort of 124 patients was 5.77 [0.03 – 25.72], while for responders was 19.97 [8.78 – 25.72] and for non-responders was 4.30 [0.03 – 25.36]. The median progression free survival (months) for the total cohort was 2.23 [0.01 – 25.63], for responders was 15.56 [2.73 – 25.62], while for non-responders was 2.03 [0.01 – 25.36].

Clinical-pathological characteristics		**Treated patients (N=124)**	**Responders (CR+PR) (N=19)**	**Non-responders (SD+PD) (N=103)**	**p value**
Age, years	**median [range]**	64 [31-84]	63 [47-78]	64 [31-84]	NA
Gender	**male**	101	14 (73%)	85 (82%)	.1602
	female	23	5 (26%)	18 (18%)	
Smoking status	**present**	22	1 (5%)	21 (21%)	**.0012**
	past	76	15 (79%)	60 (58%)	
	never	20	3 (16%)	17 (16%)	
	N.A.	6*	0 (0%)	5 (5%)	
Site of primary disease	**Oral cavity**	43	4 (21%)	39 (38%)	**.0424**
	Oropharynx	10	4 (21%)	6 (6%)	
	HPV positive **Oropharynx**	8*	0 (0%)	7 (7%)	

	HPV negative **Oropharynx** N.A.	10	3 (16%)	7 (7%)	
	Larynx	30	3 (16%)	26 (25%)	
	Hypopharynx	18	3 (16%)	15 (14%)	
	N.A.	5	2 (10%)	3 (3%)	
Performance status	**0** **1** **2**	42 78 4	9 (47,5%) 9 (47,5%) 1 (5%)	33 (32%) 67 (65%) 3 (3%)	.0521

Table 3. Clinical and pathological characteristics of 124 patients enrolled in trial Nivactor. Differences were evaluated between responders and non-responders; *for 2 patients the response was not assessed. P-value were calculated with Chi-square test. Significance was set at $p < .05$

Following the data published in literature, we evaluated differences in response comparing patients with age lower or higher than 65 years old, instead the median age (Figure 3). However, even evaluating the age with a specific cut-off, no significant differences were recorded between the two response categories (p=0.589268, Chi-square test).

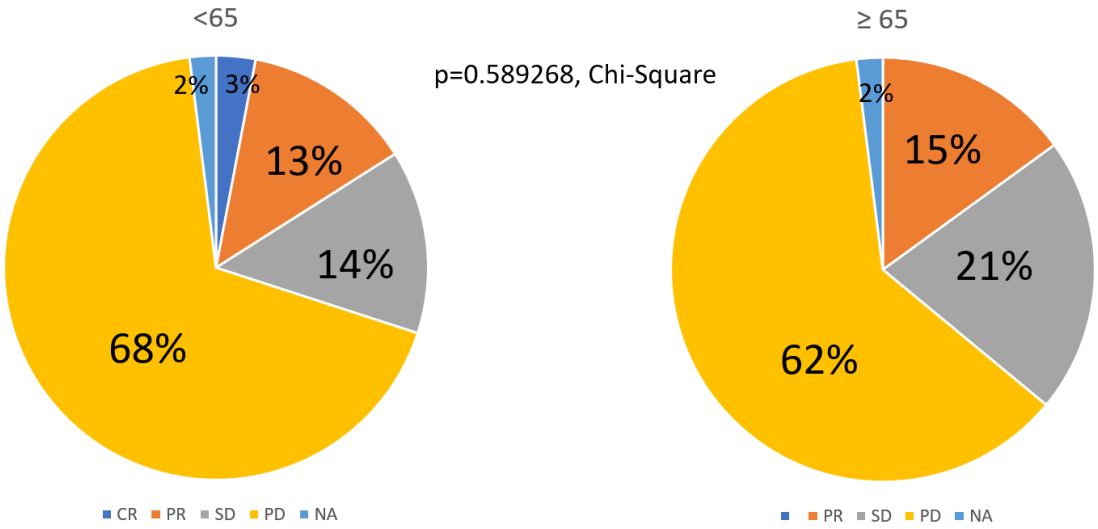


Figure 3. Pie charts of patients with age higher or lower than 65 years old associated with response. 63 patients with age < 65: 2 CR (3%), 8 PR (13%), 9 SD (14%), 43 PD (68%), 1 NA (2%); 61 patients with age ≥ 65: 0 CR (0%), 9 PR (15%), 13 SD (21%), 38 PD (62%), 1 NA (2%)

Deepening the result of differences in primary disease, we decided to compare patients that experienced an oropharyngeal cancer versus patients with cancer in other subsites. Notably, we observed a significant increase in oropharyngeal cancer (N=7, 37%) in patients that experienced a response compared to those who did not (N=20, 20%; Figure 4, p=0.007747, Chi-Square test).

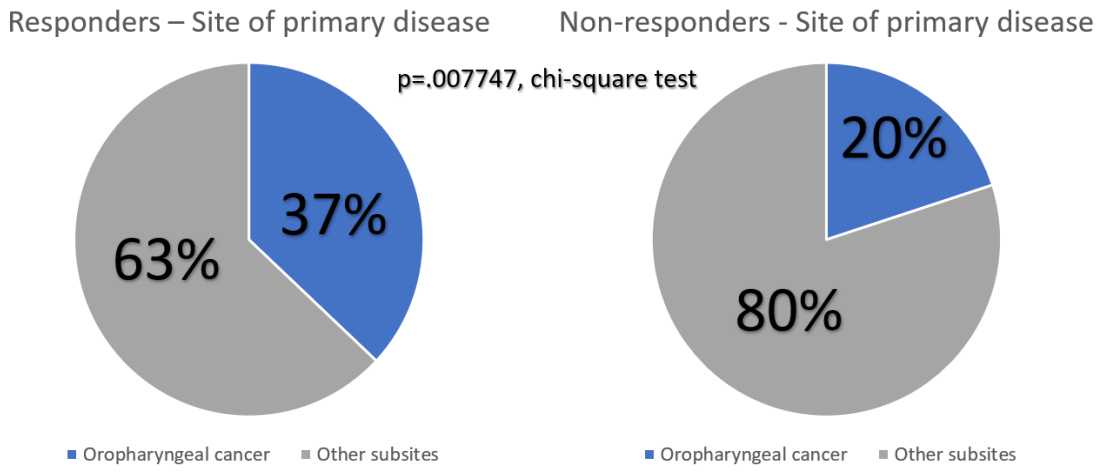


Figure 4. Pie chart of site of primary disease of responders vs non-responders

Furthermore, we interrogated the role of RECIST in assessing the effective response and the correlation with the related survival for patients treated with immunotherapy. Patients with an objective response (CR+PR) had a longer OS and PFS than non-responders (respectively p=5.5452e-8; p=0.00000016, Figure 5).

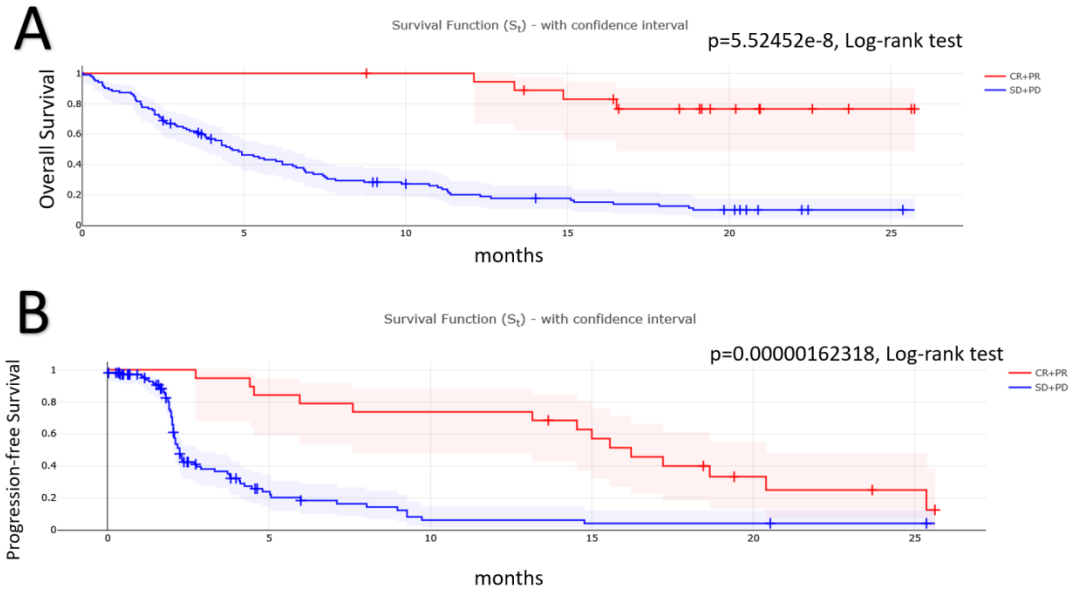


Figure 5. Kaplan-Meier curves (response).
A) OS comparing responders vs non-responders (red and blue curves, respectively); B) PFS comparing responders vs non responders (red and blue curves, respectively); the analyses were performed on 122/124 patients; for 2 patients the response status was not annotated. P-values were calculated using Log-rank test

Similarly, patients achieving a disease control rate (CR+PR+SD) had a longer OS and PFS than those with PD as best response (respectively, p=3.35369e-9; p=2.03848e-12, Log-rank test, Figure 6).

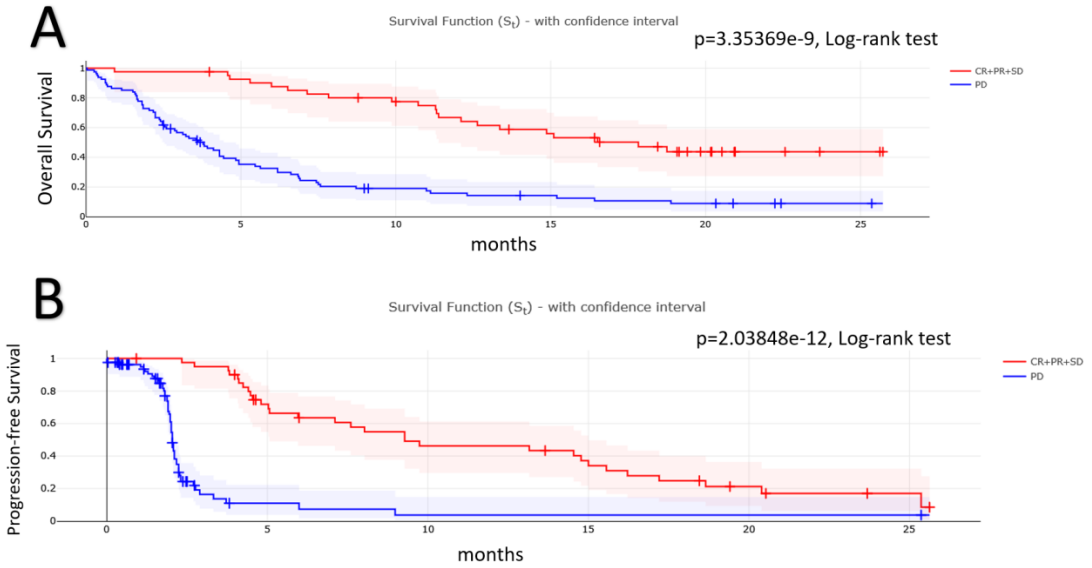


Figure 6. Kaplan-Meier curves (disease control rate) A) OS comparing DCR patients vs non-DCR patients (red and blue curves, respectively); B) PFS comparing DCR vs non-DCR patients (red and blue curves, respectively); the analyses were performed on 122/124 patients; for 2 patients the response status was not annotated. P-values were calculated using Log-rank test

In the Nivactor trial two patients experienced the complete response: both both male (age 55-63 years old). One patient was current smoker and the other past smokers. The primary disease for one patient was oral cavity and for the other HPV-positive oropharyngeal cancer. The performance status was PS=1 in both.

4.2 Tumor specimens

In the trial Nivactor the translational analyses focused on biomarker investigation, profiling the patients' tumor tissue. Specifically, tumor specimens from all the institutions were collected during 2019 and 2020, and at the end of the collection, we obtained the FFPE samples of 110 patients, while for 14 patients the tumor tissue was not available (Figure 7). The 110 collected tumor tissues were classified based on the type of specimen: 71 were from the primary tumors, 36 from recurrences or metastases, and 17 were unknowns. Notably, all the tumor samples were collected before the immunotherapy. For omics analyses, the pathologists performed the specific revision (details in Material and Methods) and after the revision 94 FFPE samples (56 primary diseases; 4 recurrences; 27 metastases; 7 unknown) were considered adequate for the extraction of both nucleic acids. However, no one of those 94 tumor specimens (available for the experiments of gene expression and DNA sequencing) derived from a complete responder.

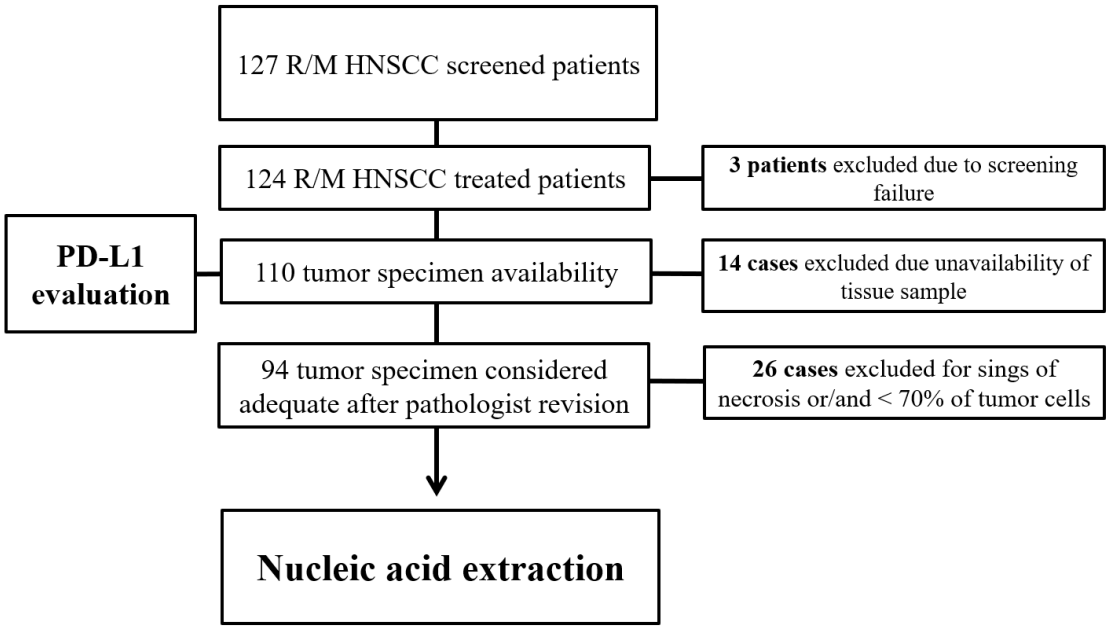


Figure 7. Tumor specimens collected in Nivactor trial

4.3 PD-L1 status

The first biomarker investigated was the expression of the PD-L1 protein by IHC. As described in Figure 7, the tumor samples available for PD-L1 examination were 110. However, for 6 samples the material was scant, while for additional 4 samples (considering TPS analysis) and for 5 samples (considering CPS analysis), the scores were not methodologically evaluable (Figure 8). The measurement of PD-L1 was possible for 94 samples analyzing TPS and for 93 samples analyzing CPS score (Figure 8). The two scores were considered in a separate manner, and R/M HNSCC ICI trials cut-offs were investigated (see Introduction, 1.11 Major clinical trial using anti-PD-1 agents for R/M HNSCC for details). Considering TPS, we observed 6 samples which resulted ≥ 50 (2 PR; 1 SD; 3 PD), while 33 samples were CPS ≥ 20 (1 CR; 6 PR; 8 SD; 18 PD). The two patients with a complete response resulted with TPS < 50 and CP ≥ 1.

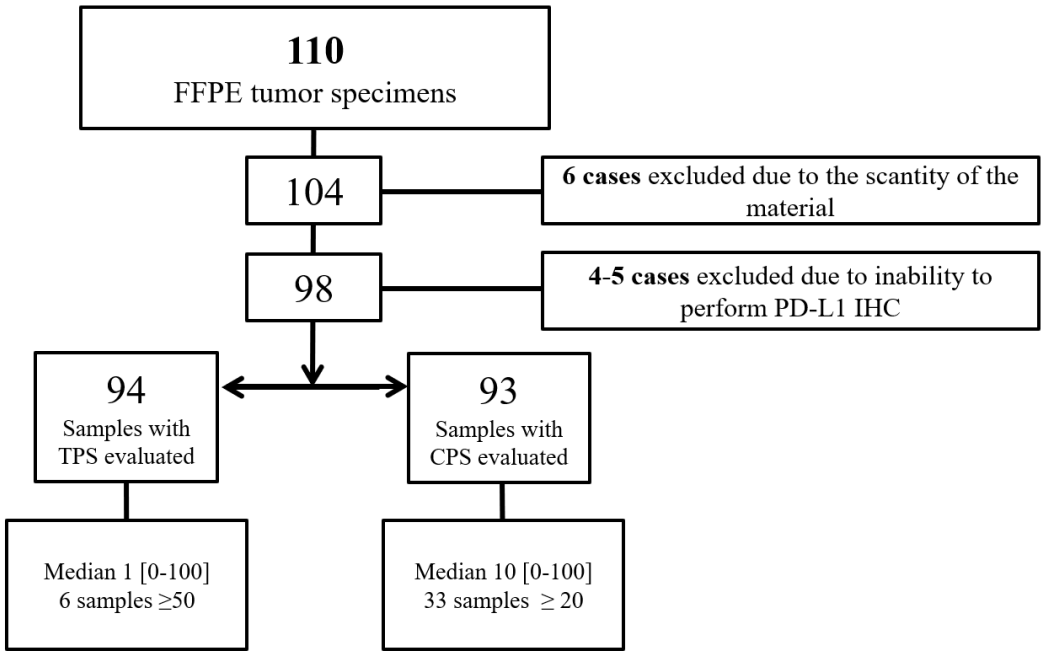


Figure 8. PD-L1 IHC evaluation

To understand if any difference in PD-L1 expression could be observed between primary tumor and R/M, we compare TPS and CPS score in the samples. Evaluating the TPS in only primary tissues we observed that for 55 primary tissue cases the median value was 1 (ranging from 0 to 100) and the same TPS median value was observed in the 31 cases of recurrence/metastasis (ranging from 0 to 80). While, for CPS the median value in primary tissue was 9 (ranging from 0 to 100), while in R/M the CPS score was 5 (ranging from 0.5 to 100). No significant differences in terms of PD-L1 expression were observed comparing primary and R/M tumor tissue (considering both scores, Table 4).

	Primary tumor tissue (N=55 for TPS) (N=54 for CPS)	**Recurrence/Metastasis Tumor tissue** (N=31)	**p-value**
TPS % score, median [range]	1 [0-100]	1 [0-80]	0.69627
CPS score, median [range]	9 [0-100]	5 [0.5-100]	

Table 4. Comparison of PD-L1 expression in primary and recurrence/metastasis samples. For 8 samples details about tumor tissue were not available. Significance was set at $p \leq 0.05$ and it was calculated by Chi-Square test

After the comparison between samples, we aimed to explore the correlation between TPS score and response (CR+PR vs SD+PS) or disease response rate (CR+PR+SD vs PS, Figure 9). Indeed, we correlated TPS and CPS with RECIST, and we observed a significant association between TPS and response (Figure 9A, p=0.0121, Wilcoxon test), but not between TPS and disease response rate (Figure 9B, p=0.0527, Wilcoxon test).

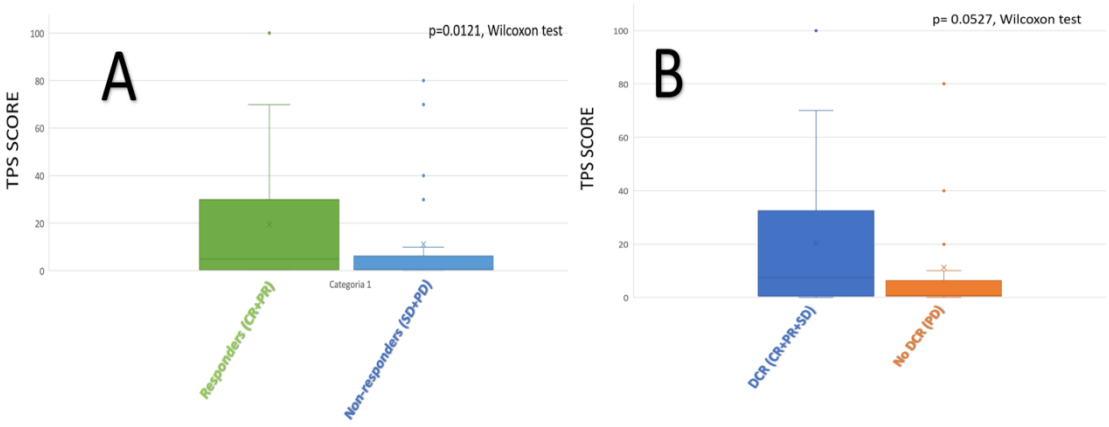


Figure 9. Response and TPS % score A) Boxplots showing the correlation between TPS score and response (15 patients CR+PR vs 79 patients SD+PD). B) Boxplots showing the correlation between TPS score and disease control rate (30 CR+PR+SD vs 64 PD). P-values were calculated using Wilcoxon test, significance was set at $p \leq 0.05$

The same was applied for CPS score, and we did not observe a significant correlation between response and CPS score (Figure 10A). However, contrary on what observed for TPS score, we observed a significant difference in CPS score between patients that experienced a disease control rate versus those who did not (Figure 10B, p=0.0287, Wilcoxon test).

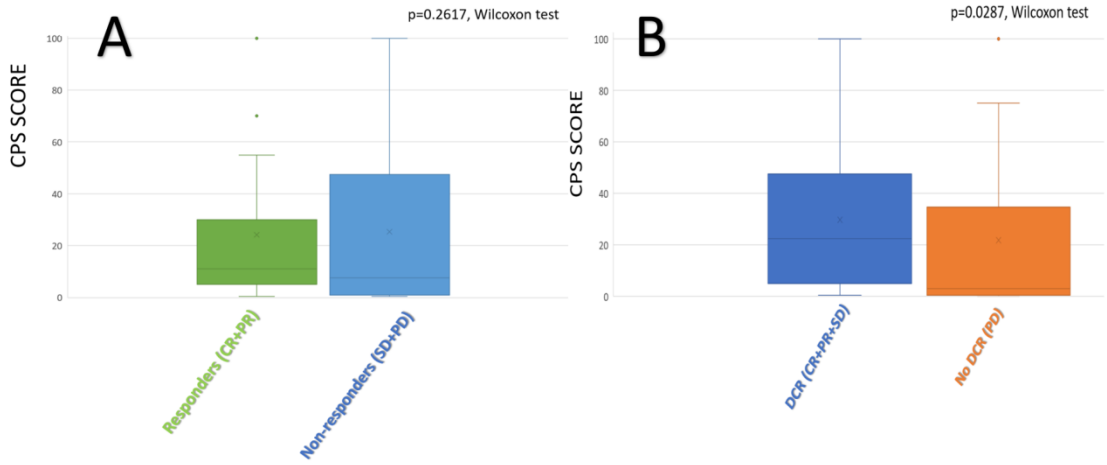


Figure 10. Response and CPS score. A) Boxplots showing the correlation between CPS score and response (15 patients CR+PR vs 78 patients SD+PD). B) Boxplots showing the correlation between CPS score and disease control rate (30 CR+PR+SD vs 63 PD). P-values were calculated using Wilcoxon test, significance was set at $p \leq 0.05$

Moreover, we considered TPS and CPS scores for the survival analyses, evaluating both OS and PFS. These explorative analyses considered the cut-off already utilized in different trials in which R/M HNSCC patients were treated with anti PD-1 agents. The first score investigated was TPS, and we interrogated about the possible difference in terms of survival of patients with TPS ≥ 1% than those with TPS < 1% (Figure 11).

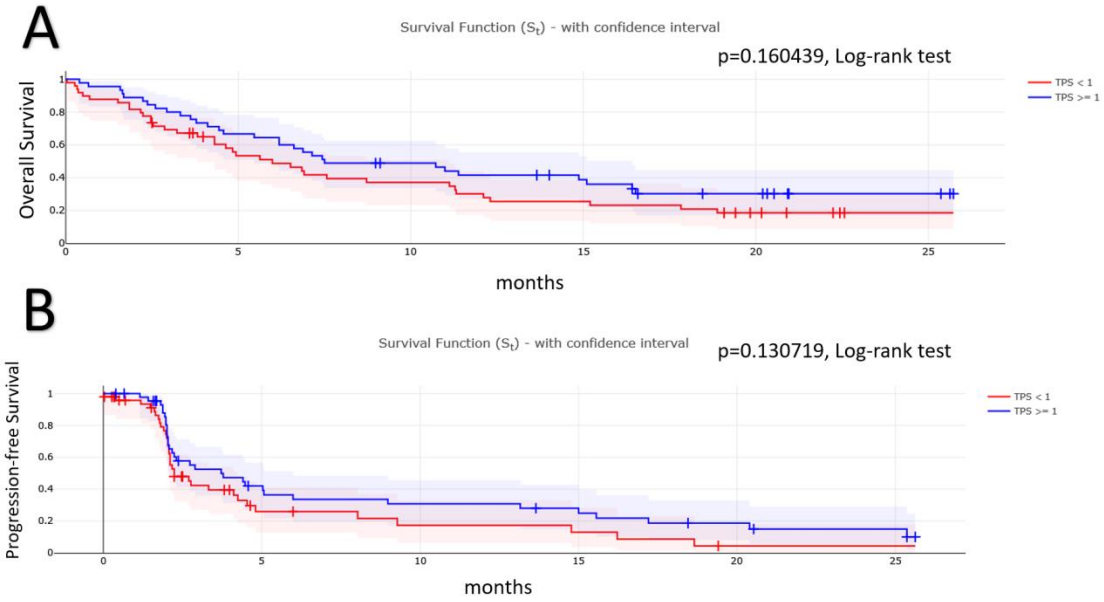


Figure 11. Kaplan Meier curves (N=94) TPS % score≥1 A) OS analysis considering the score TPS using the cut-off=1; B) PFS analysis considering the score TPS using the cut-off=1%. TPS<1% curve was coloured in red while TPS≥1% was in blue. P-values were calculated using Log-rank test, significance was set at $p \leq 0.05$.

However, observing the curves no differences could be seen in both OS and PFS analyses (Figure 11 A and Figure 11 B, respectively). Proceeding with TPS score, we utilized also 50% as cut-off.

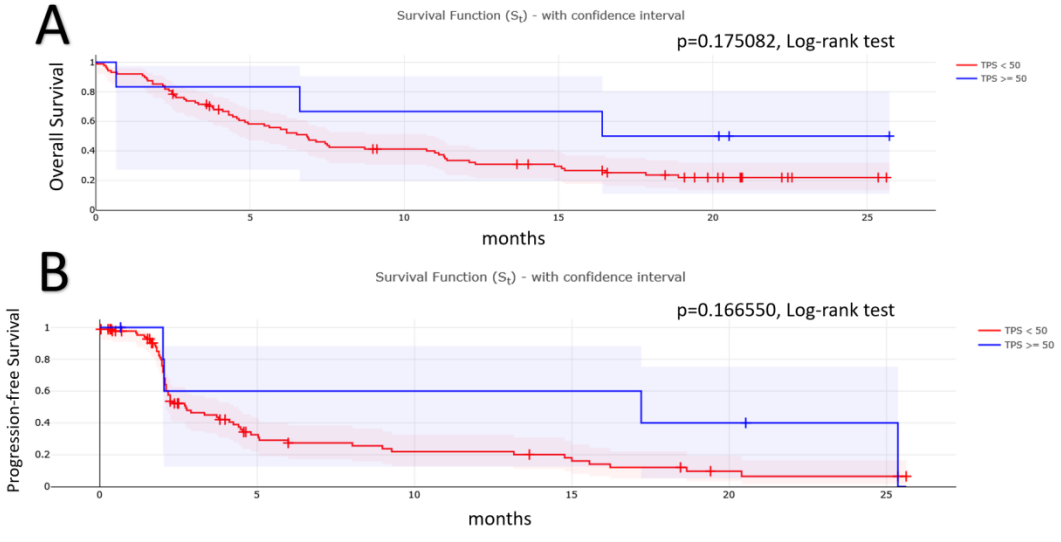


Figure 12. Kaplan Meier curves (N=94) TPS score≥50%. A) OS analysis considering the score TPS using the cut-off=50; B) PFS analysis considering the score TPS using the cut-off=50. TPS<50 curve was coloured in red (N=88) while TPS≥50 was in blue (N=6). P-values were calculated using Log-rank test, significance was set at $p \leq 0.05$.

Nevertheless, interrogating TPS score and its prognostic role with another cut-off we did not observed differences in survival for both OS and PFS using 50 as cut-off (Figure 12). Moreover, we investigated the prognostic role of CPS score.

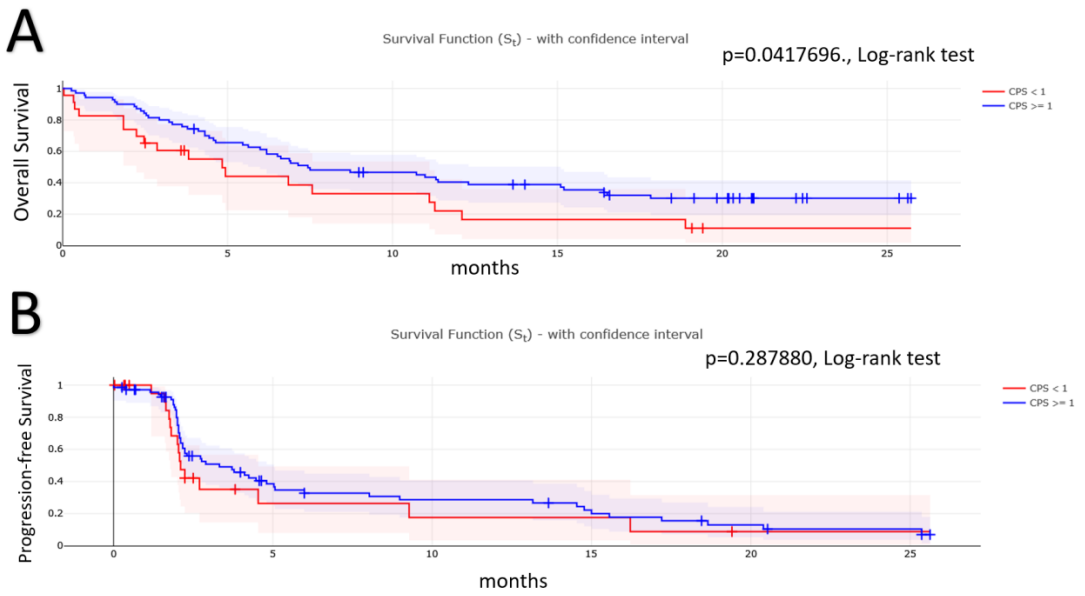


Figure 13. Kaplan Meier curves (N=93) CPS score≥1 A) OS analysis considering the score CPS using the cut-off=1; B) PFS analysis considering the score CPS using the cut-off=1. CPS<1 curve was coloured in red while CPS≥1 was in blue. P-values were calculated using Log-rank test, significance was set at $p \leq 0.05$.

Studying the differences in survival utilizing different CPS cut-off, what we observed was that patients with CPS≥1 had benefits in terms of OS, but not studying PFS (Figure 13). Considering CPS≥20 as cut-off we did not observed any survival benefit (Figure 14).

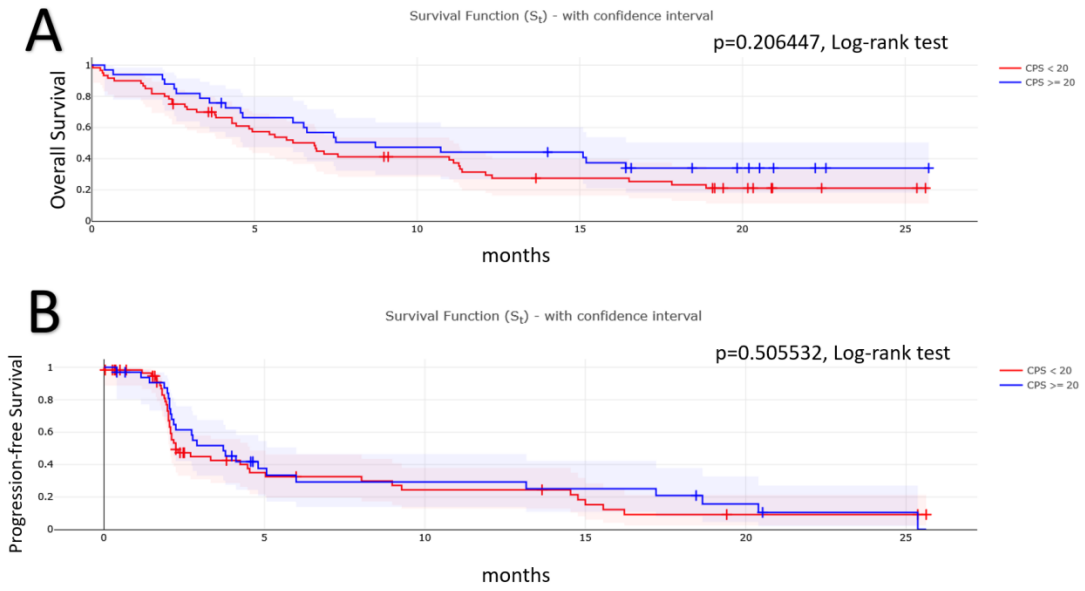


Figure 14. Kaplan Meier curves (N=93) CPS score≥20. A) OS analysis considering the score CPS using the cut-off=20; B) PFS analysis considering the score CPS using the cut-off=20. CPS<20 curve was coloured in red while CPS≥20 was in blue. P-values were calculated using Log-rank test, significance was set at $p \leq 0.05$.

4.4 Gene expression experiments

After pathologists' revision, 94 samples were considered adequate for omics analyses (Figure 7). Therefore, the RNA extraction was carried out, and consequently nucleic acids quality and quantity check were performed. For quality and quantity 5 and 4 samples were excluded, respectively. Following the gene expression experiments, 2 samples did not well-performed the cDNA reverse transcription and they were not considered for the protocol next steps. After the chip hybridization, the primary images were examined, and 3 samples were excluded for poorly hybridization. The Clariom-D assay contains probes designed on intron (i.e. negative controls) and exon (i.e. positive controls) regions of housekeeping genes. Negative controls are a measure of false positive signals and, in theory, no signal should be detected in the negative controls; to evaluate how well the probe set summary for a specific gene separates the positive controls from the negative controls the area under the curve

(AUC) for a receiver operator curve (ROC) comparing the intron controls to the exon controls was assessed. An AUC of 1 reflects perfect separation whereas as an AUC value of 0.5 would reflect no separation (Figure 15A). The Positive vs Negative AUC controls is visualized as a bar graph to monitor assay quality across samples.

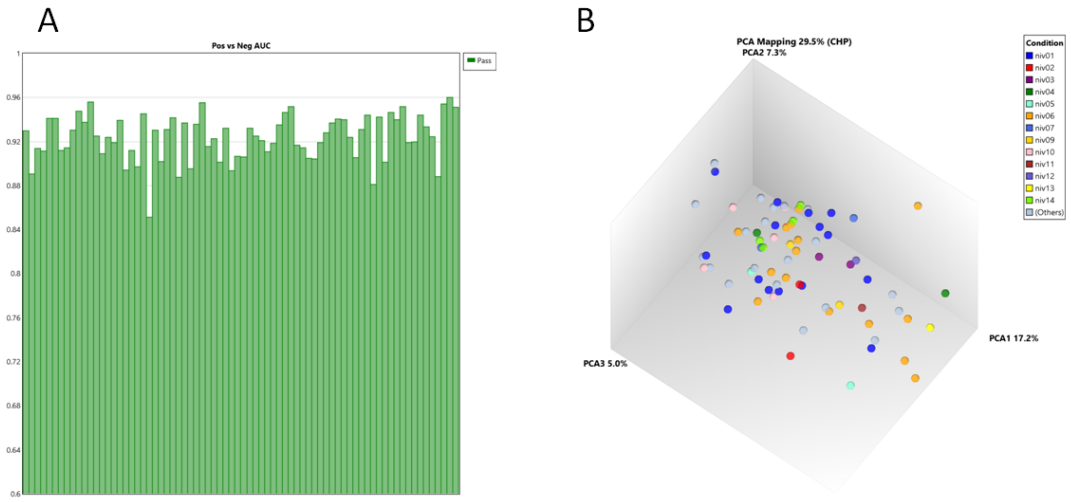


Figure 15 Quality control of the gene expression profiling. A) AUC of the positive and negative controls; B) PCA of the NIVACTOR samples based on samples centre of origin

To check the quality of the gene-expression profiling, we performed an exploratory grouping analysis. This approach highlights the presence and relationships among groups of samples, proving to be a helpful way to discover the underlying clusters within an experiment. Moreover, the analysis is usually applied to discover possible batch effects (i.e, patterns due to technical issues) that can be removed in later analysis. To visualize the results, the large number of probe set signals per sample was reduced to three for viewing, using Principal Component Analysis (PCA). The plot shows the data distribution by PCA of the NIVACTOR samples and the different colors show the different centers where the patients were enrolled. No significant association can be detected in the PCA (Figure 15B). The final dataset cohort for gene expression was composed by 80

cases (Figure 16). Of the 80 samples, 52 profiled samples derived from the primary tissue, while 19 from metastasis, 3 from recurrence, and 6 from unknown source. The 80 patients were divided for response status, according to the RECIST in: 12 partial response (PR, 15%); 14 stable disease (SD, 18%); 53 with progressive disease (PD, 66%); 1 without assessed response (N.A., 1%). However, we did not dispose of the tumor material of the two complete responding patients in order to obtain their biological profile.

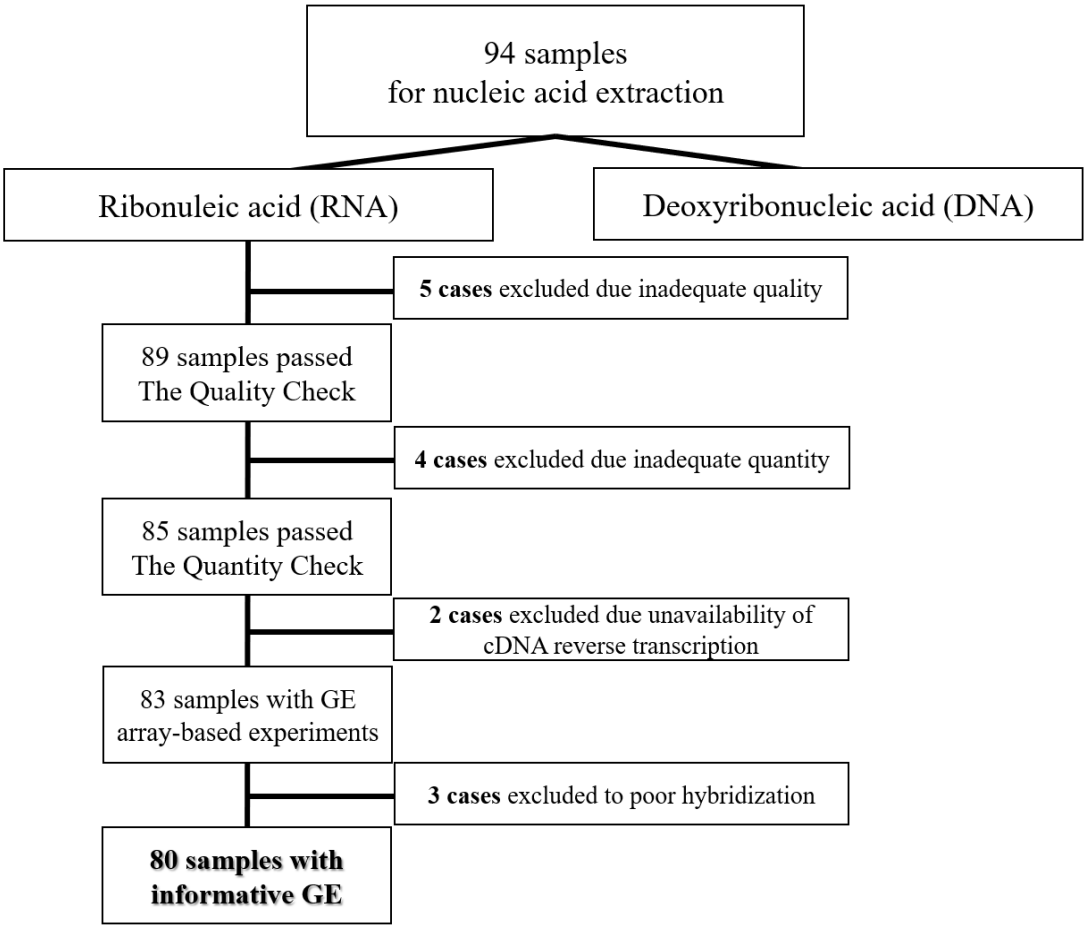


Figure 16 Consort diagram of gene expression dataset

Comparing the treated patients dataset (N=124) and the gene expression dataset (N=80) we observed no differences between all the clinical-pathological characteristics (Table 5). The median overall survival for the gene expression cohort was 6.72 months [0.03 – 25.72], while the median

progression free survival was 2.20 months [0.01 – 25.62]. For this reason, we considered the gene expression dataset (N=80) representative of the entire cohort (N=124). We proceed with the evaluation of biological characteristics of the 80 patients.

Clinical-pathological characteristics		**Treated patients (N=124)**	**GE dataset (N=80)**	**p value**
Age, years	**median [range]**	64 [31-84]	65.5 [33-84]	NA
Gender	**male**	101 (81%)	65 (81%)	1
	female	23 (19%)	15 (19%)	
Smoking status	**present**	22 (18%)	16 (20%)	.86335
	past	76 (61%)	50 (62%)	
	never	20 (16%)	11 (14%)	
	N.A.	6 (5%)	3 (4%)	
Site of primary disease	**Oral cavity**	43 (35%)	30 (37%)	.934331
	Oropharynx HPV positive	10 (8%)	7 (9%)	
	Oropharynx HPV negative	8 (6%)	5 (6%)	
	Oropharynx N.A.	10 (8%)	7 (9%)	
	Larynx	30 (24%)	18 (22%)	
	Hypopharynx	18 (15%)	11 (14%)	
	N.A.	5 (4%)	2 (3%)	
Performance status	**0**	42 (34%)	27 (34%)	.927346
	1	78 (63%)	50 (62%)	
	2	4 (3%)	3 (4%)	
Response	**CR**	2 (1.5%)	0 (0%)	.984821
	PR	17 (14%)	12 (15%)	
	SD	22 (18%)	14 (18%)	
	PD	81 (65%)	53 (66%)	
	NA	2 (1.5%)	1 (1%)	

Table 5. Clinical and pathological characteristics of gene expression dataset (N=80) compared to the total cohort (N=124). For 3/124 and 1/80 patients the response was not assessed. P-values were calculated with Chi-square test. Tests significance was set at $p < .05$

Notably, in the cohort of gene expression database (N=80) two patients experienced an OS and PFS > 25 months. The two patients were both male and they were 62 and 63 years old, respectively. One was a past smoker, while the other never smoked. The primary site of oncologic disease was oropharynx and they both resulted HPV-positive. TPS was equal to 10 for the first patient and 100 for the second, while CPS was 11 and 100. They both experienced a partial response to nivolumab. From this paragraph to the end, we will refer to analyses on the gene expression database cohort or DNA sequencing cohort, and not to the total cohort of 124 patients.

4.5 Immunotherapy response

4.5.1 Responders vs non-responders (RECIST v1.1)

To confirm the results observed in the total cohort (N=124), we evaluated the correlation between response status and survival by the Kaplan-Meier analyses. Specifically, we investigated the OS and PFS of partial responders (PR, N=12) versus not responders (SD+PD, N=67). For 1/80 patient the response was not annotated.

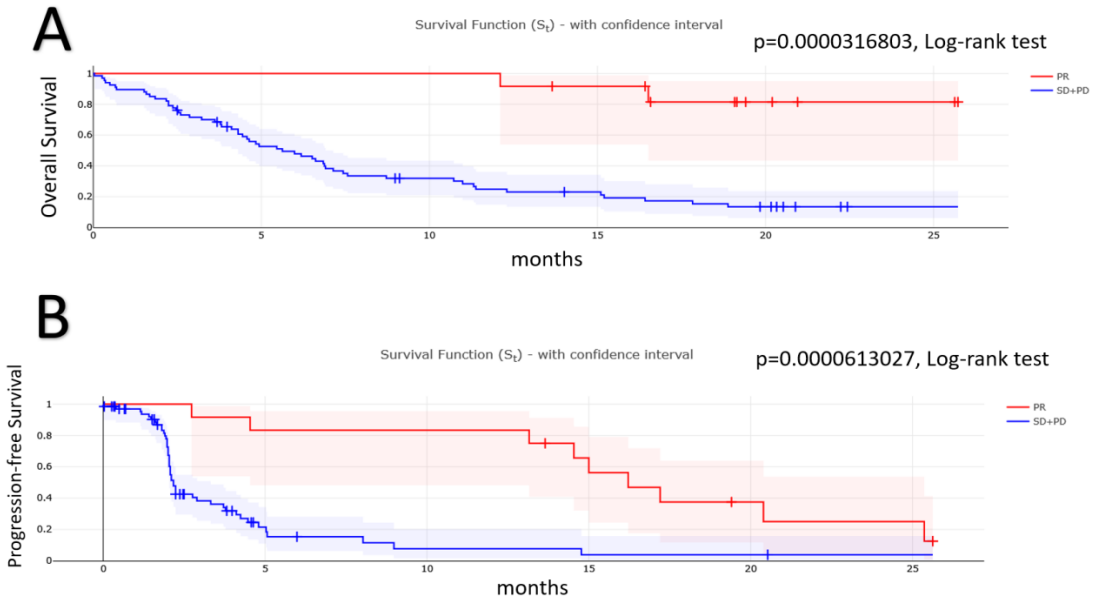


Figure 17 Kaplan Meier curves (response; N=79). A) OS analysis of partial responders (PR, red curve) vs non-responders (Stable Disease and Progressive Disease, SD+PD; blue curve); B) PFS analysis of partial responders vs non-responders. P-values were calculated using Log-rank test, significance was set at $p \leq 0.05$.

The survival analyses, for both OS and PFS probabilities, revealed a strong significance (Figure 17), confirming the results observed in the total cohort. Proving the differences in survival between the two categories of patients that experienced a response or did not (assessed by RECIST), we explored the biological differences of the two groups. Analyzing the gene expression of the 79 patients with response annotated, we did not observe any de-regulated gene (significance set at adjusted p-value ≤ 0.05). Investigating differences in terms of microenvironment composition, no one of the cell population resulted de-regulated in one of the two categories, performing both xCell (64 cell types + 3 summary scores) and TIMER methods (six cell types: B-cells, Dendridic cells, Macrophages, Neutrophils, CD4 T-cells, CD8 T-cells). Furthermore, no differences were observed comparing the de-regulated pathways in the two categories, evaluating both Hallmark GSEA and KEGG pathways.

4.5.2 Disease control rate (RECIST v1.1)

Then, we evaluated the disease control rate, exploring the differences between patients who experienced a benefit from immunotherapy (PR+SD) who does not (PD). In the survival analyses, we observed a significant difference considering both in OS and PFS curves (Figure 18), confirming what observed in the total cohort.

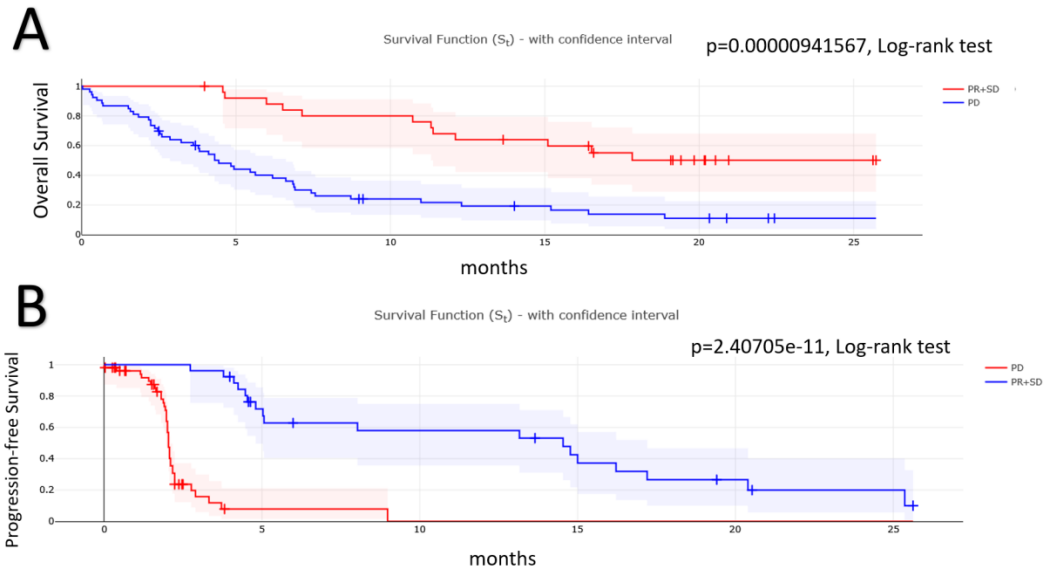


Figure 18 Kaplan-Meier curves (disease control rate; N=79). Survival of disease control rate patients' categories (Partial Response and Stable Disease, PR+SD vs Progressive Disease, PD)

Analyzing the differentially expressed genes no gene resulted significantly de-regulated comparing the two categories (significance set at adjusted p-value ≤0.05). For the assessment of tumor microenvironment components, xCell and TIMER were interrogated. In both xCell and TIMER analyses no one of the cell resulted differently enriched in one the two categories. Moreover, investigated the de-regulated biological pathways by Hallmark GSEA and KEGG methods, we did not observe any significant difference between patients with a disease control rate and patients that experienced progression. Through the gene expression we were not able to discern biological differences between responders vs not responders or patients with a treatment benefit vs patients without. At this point, we decided to

assess if any biological difference was present in the clinical-epidemiological groups that resulted significantly different between responders and non-responders in the cohort of 124 patients (Table 3). Investigating the smoking status, we did not observe any differences in terms of gene expression, evaluating de-regulated genes, microenvironment composition and de-regulated pathways. On the contrary, the anatomical site was not investigated for the biological characteristics, being the biological differences between the four classical HNSCC subsites extensively narrated in literature. Specific molecular features of patients that experienced a response or a benefit from ICI therapy were not addressed. For this reason, we proceed in evaluating additional biological markers.

4.5.3 PD-L1

After observing significant association between PD-L1 expression and response and differences in terms of survival for using TPS and CPS score with different cut-off, we aimed to understand if a biological background could resonate with these results. Indeed, biological differences between the TPS and CPS categories were investigated in the 80 cases having both gene expression and PD-L1 analyses. We characterized the tumor microenvironment composition (xCell and TIMER) and the de-regulated pathways (GSEA Hallmark and KEGG). However, no significant differences were observed between the different cut-off categories. Neither using the continuous scores of TPS and CPS a significant correlation was observed between specific biological characteristic and PD-L1 scores.

4.6 Gene expression signatures

4.6.1 Literature research

To investigate TME and its potential implication as biomarker, immune-related gene expression signatures were retrieved from literature, and specifically 27 gene signatures had been selected, published from 2015 to 2021 (Table 6). However, for 14/27 signature the methods were lacking or partially explained, and it was impossible to test the gene signatures. As final result, it was possible to test 13/27 signatures.

Name	Type fo tumor: HNSCC ?	HNSCC site specific?	Dataset	Validation	Reference	Year	DOI	Possible to be tested?
Cytolitic activity	Yes, with other solid cancer types	No	TCGA	No	Rooney, Cell	2015	https://doi.org/10.1016/j.cell.2014.12.033	**Applicable, methods available**
IFN-γ gene signature	Yes, with other 8 cancer types	No	Not deposited	Yes, KEYNOTE-001 (n=62 additional metastitc melanoma pts), KEYNOTE-012 (n=40 HNSCC pts; n=33 gastric cancer pts) and all the patients from KEYNOTE-012 and KEYNOTE-028 (not deposited)	Ayers, J Clin Invest	2017	DOI: 10.1172/JCI91190	**Applicable, methods available**
Expanded immune gene signature	Yes, with other 8 cancer types	No	Not deposited	Yes, KEYNOTE-001 (n=62 additional metastitc melanoma pts), KEYNOTE-012 (n=40 HNSCC pts; n=33 gastric cancer pts) and all the patients from KEYNOTE-012 and KEYNOTE-028 (not deposited)	Ayers, J Clin Invest	2017	DOI: 10.1172/JCI91190	**Applicable, methods available**
Immunophenoscore (IPS)	Yes, with 20 solid tumors	No	TCGA	37 microarray deposited dataset	Charoentong, Cell Rep	2017	doi: 10.1016/j.celrep.2016.12.019.	**Applicable, methods available**

Immunogenomics pipeline	Pancancer	No	TCGA	No	Thorsson, Immunity	2018	doi: 10.1016/j.immuni.2018.03.023	**Not applicable, methods not explained**
Immune-related clusters	Yes, only	No	TCGA	No	Chen, Annals of Oncology	2019	https://doi.org/10.1093/annonc/mdy470	**Not applicable (immune classes are not defined)**
six-gene prognostic signature for OSCC	Yes, only	Yes, OSCC	GSE85195, GSE23558, and GSE10121	TCGA (331 patients) + 28 OSCC/normal patients tissues to further validate the signature	Jiaying Wang	2019	DOI: 10.1002/jcp.29210	**Not applicable, methods not explained**
Immune/Inflammatory-Related Risk Signature	Yes, only	Yes, OSCC	A cohort of 314 OC-SCC samples possessing whole genome expression data that were sourced from The Cancer Genome Atlas (TCGA) database	GSE41613	Shuang Bai, Journal of Oncology	2019	DOI: 10.1155/2019/3865279	**Applicable, methods available**
Hypoxia and Immune Prognostic Classifier	Yes, only	No	TCGA	TCGA	Jill M Brooks	2019	https://doi.org/10.1158/1078-0432.CCR-18-3314	**Not applicable, methods not explained**
TMEscore	Yes, only	No	TCGA	Yes, GSE65858	Huo, Sci Rep	2020	https://doi.org/10.1038/s41598-020-68074-3	**Not applicable (genes list and weithgs are missing)**
Immune-related signature	Yes, only	No	TCGA	Yes, not deposited (n=115)	Yao, Journal for Immunotherapy for cancer	2020	http://dx.doi.org/10.1136/jitc-2019-000444	**Not applicable (genes list and weithgs are missing)**
EMT gene signature	Yes, only	No	TCGA	GSE65858, GSE41613, GSE42743	Jung, Scientific report	2020	https://doi.org/10.1038/s41598-020-60707-x	**Not applicable (genes weithgs are missing)**
Immune Risk Model	Yes, only	No	TCGA	No	Liu, Cancer genetics	2020	https://doi.org/10.3389/fgene.2020.576566	**Applicable, methods available**
tumor immune cell infiltration (ICI)	Yes, only	No	GSE41613, GSE42743, E-MTAB-1328 and GSE65858; TCGA	No	Zhang, Molecular therapy acid nucleici	2020	https://doi.org/10.1016/j.omtn.2020.08.030	**Not applicable, methods not explained**

27 prognostic IRGs genes signature	Yes, only	No	TCGA	GSE65858	Yangyang She	2020	DOI: 10.1186/s12935-020-1104-7	**Applicable, methods available**
Immune-related gene prognostic index	Yes, only	No	TCGA	GSE65858	Chen, Clinical Cancer Research	2020	DOI: 10.1158/1078-0432.CCR-20-2166	**Not applicable, methods not explained**
Immune related signature	Yes, only	No	TCGA	GSE41613	Fang, Aging	2021	DOI: 10.18632/aging.202842	**Applicable, methods available**
Immune signature	Yes, only	No	TCGA	GSE65858	Qiang, Computational and Structural Biotechnology Journal	2021	https://doi.org/10.1016/j.csbj.2021.01.046	**Applicable, methods available**
Microenvironment subtypes	Pancancer	No	TCGA, ICGC, GTEx	No	Bagaev A, Cancer Cell	2021	https://doi.org/10.1016/j.ccell.2021.04.014	**Not applicable, methods not explained**
Tumor immunological phenotypes	Pancancer	No	TCGA	No	Wang	2021	DOI: 10.1126/sciadv.abd7851	**Not applicable, methods not explained**
Subtyping based on immune signatures	Yes, only	No	TCGA	GSE65858, GSE30784, GSE39366	Song, Int Immunopharmacol	2021	https://doi.org/10.1016/j.intimp.2021.108007	**Not applicable, methods not explained**
Invasiveness score	Pancancer	No	TCGA	7 GEO database	Bi Ji, J Transl Med	2021	https://doi.org/10.1186/s12967-021-02773-x	**Applicable, methods available**
quanTIseq deconvolution algorithm	Pancancer	No	TCGA	No	Finotello, Genome Med	2021	https://doi.org/10.1186/s13073-019-0638-6	**Not applicable, methods not explained**
Immune exhaustion gene signature	No, Thyroid cancer	No	Not deposited	TCGA, GSE29265, GSE33630	Li, The Journal of Clinical Endocrinology & Metabolism	2021	https://doi.org/10.1210/clinem/dgab132	**Applicable, methods available**
myeloid signature	Yes, only	No	TCGA	GSE65858	Liu, Frontiers in Immunology	2021	https://doi.org/10.3389/fimmu.2021.659184	**Applicable, methods available**
RNA-binding protein (RBPs) trascription level	Yes, only	No	TCGA	GSE65858	Hu, Frontiers in Genetics	2021	https://doi.org/10.3389/fgene.2020.571403	**Applicable, methods available**
IRGPs signature	Yes, only	No	TCGA	GSE41613	Jiang, Translational Oncology	2021	https://doi.org/10.1016/j.tranon.2020.100924	**Not applicable, methods not explained**

Table 6. Gene expression signatures retrieved from literature

4.6.2 Gene expression signatures from literature

4.6.2.1 Gene expression signature non significantly associated with survival

For the 13 gene expression signatures, which methods were available, the score was calculated, and their association with survival was evaluated by Kaplan-Meier curves. Five of the thirteen gene expression signature resulted non-significantly associated with both to OS and PFS.

Specifically, statistical results for *Bi et al* were p=0.400404 for OS and p=0.114788 for PFS (Log-rank test); for *She et al* p=0.195569 and p=0.225746 (OS and PFS, respectively); for *Qiang et al* were p=0.189808 and p=0.0846285 (OS and PFS, respectively); for *Ayers et al (immune expanded)* p=0.392714 for OS and p=0.175512 for PFS. More details of the gene signature were descripted in Table 7.

GE name	Genes	Details	doi
Invasiveness-related molecular features (Bi et al.)	4-gene expression signature *(COL11A1, POSTN, EPYC, ASPN, COL10A1, THBS2, FAP, LOX, SFRP4, INHBA, MFAP5, GREM1, COMP, VCAN, COL5A2, COL5A1, TIMP3, GAS1, TNFAIP6, ADAM12, FBN1, SULF1, COL1A1, and DCN)*	Positive correlation between invasiveness score and immune-associated biomarkers across multiple cancer types; potential application of the score in the identification of tumor patients who are more likely to benefit from immunotherapy.	https://doi.org/10.1186/s12967-021-02773-x ([131])
Prognostic immune-related gene signature (IRGS) (She et al.)	27-gene expression signature *(RFXAP, ULBP1, TMSB4Y, RBP4, LCNL1, CCR6, KLRK1, PTX3, MASP1, HRG, CCL22, OLR1, ROBO1, BTC, CHGB, DKK1, HBEGF, INHBB, PDGFA, AVPR2, IL20RA, RORB, TNFRSF18, TNFRSF25, TNFRSF4, SH3BP2, ICOS)*	IRGS signatures revealed that a higher score of immune cell infiltration was present in the low-risk group, and immunosuppression is an indispensable factor of carcinogenic progression in HNSCC.	https://doi.org/10.1186/s12935-020-1104-7 ([132])
Prognostic immune-related genes (IRGs) signature (Qiang et al.)	*13-gene expression signature (PLAU, IRF9, CCL26, BLNK, SEMA3G, FPR2, GAST, IL34, SLURP1, STC1, STC2, TNFRSF12A* and *TNFRSF25)*	IRG-based prognostic signature for HNSC and proved its predictive capability in multiple datasets; the signature could potentially provide a foundation for individualized cancer immunotherapy.	10.1016/j.csbj.2021.01.046 ([133])
Expanded immune gene signature (Ayers et al.)	18-gene expression signature *(CD3D, IDO1, CIITA, CD3E, CCL5, GZMK, CD2, HLA-DRA, CXCL13, IL2RG, NKG7, HLA-E, CXCR6, LAG3, TAGAP, STAT1, GZMB)*	T cell–inflamed phenotype necessary for the clinical activity of PD-1–/PD-L1–directed monoclonal antibodies; the signature was prospectively validated in a large, independent cohort of PD-L1–unselected patients with HNSCC.	https://doi.org/10.1172/JCI91190 ([107])

Table 7. Detail of the 4 gene expression signature not significantly associated with survival

In addition, the 4 signatures did not correlate with response or disease response rate neither. We decided to additionally test the Immunophenoscore (IPS), a pancancer-algorithm[134]. Authors characterized the intratumoral immune landascape of 20 solid cancers, and they observed a correlation between the response to immunotherapy. However, in Nivactor cohort we did not observe a correlation between the score and the response (Figure 19). Moreover, it was not possible to divide patients in two group based on immunophenoscore and consequently, it was not possible to investigate the IPS prognostic role.

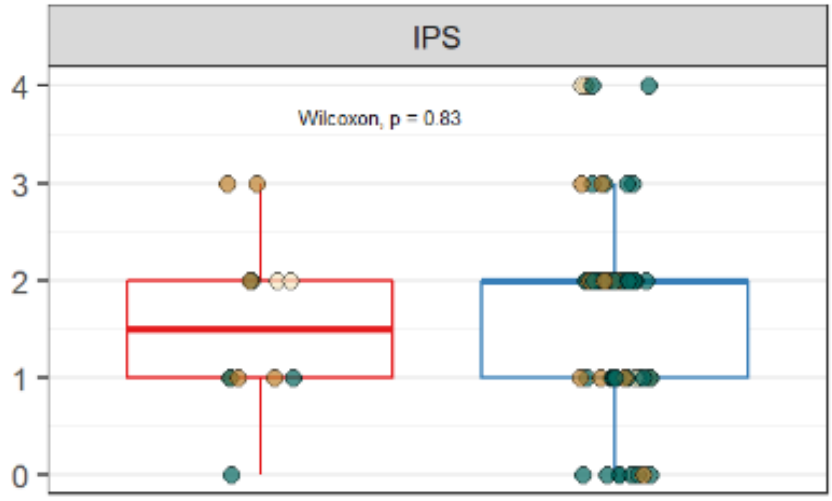


Figure 19 Boxplots of Immunophenoscore (IPS). Red: responding patients (PR); blue: non-responding patients (SD+PD). Significance was calculated by Wilcoxon test

4.6.2.2 Gene expression signature significantly associated with survival

On the other hand, 8/13 literature signatures resulted significantly associated with survival (OS or PFS).

1.Bai et al. The first was the prognostic immune/inflammatory signature of Bai et al.[135], comprising 18 genes (CD27, CD79B, CMA1, CCR4, CCR7, CNR2, CTLA4, CTSG, GZMM, IL16, MASP1, SAA1, CCL11, TNFAIP3, BATF, IL19, PGLYRP4, and TREML1) and specifically designed for OSCC. In their work, based on the median cut-off value of the risk, they divided patients into high (worse prognosis) and low (better

prognosis) risk group. Their results showed that the lower risk score was correlated with inflammatory response and the higher risk score with cell-cycle related process. Testing the signature in the Nivactor cohort (N=80) we observed that the signature resulted associated with PFS (p=0.0455087, Log-rank test, Figure 20), but not with OS, however confirming the high-risk score group (red curves) with a poorer prognosis.

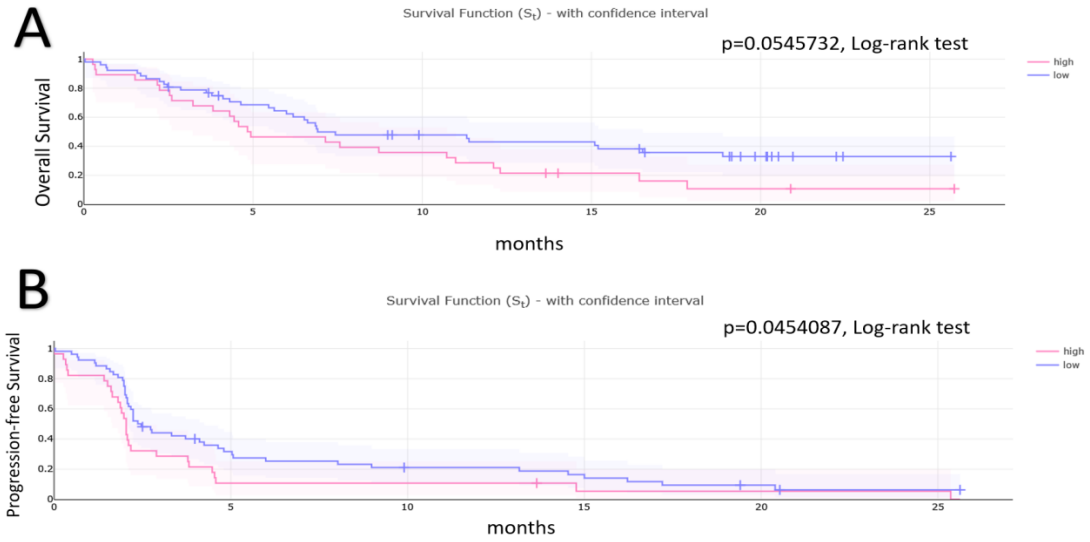


Figure 20 Kaplan-Meier analyses for OS (A) and PFS (B) testing Bai et al. immune-inflammatory signature on Nivactor cohort (N=80)

Investigating the biological differences between the two risk groups, we observed different de-regulated Hallmark GSEA pathways in the high group (worse prognosis), showed in Table 8. As observed in the original work of Bai et al. we evidenced an up-regulation of pathways related to signaling and proliferation. Moreover, we observed an enrichment in pathway related to DNA repair and those correlated with a more aggressive disease and worse prognosis, such as Epithelial mesenchymal transition and hypoxia pathways. Moreover, the high-risk group resulted similar in 4 Hallmark GSEA pathways comparing to the profile of "Immune exhausted HNSCC" described by Chen et al[136] .

Group enriched	Hallmark GSEA pathway	p-value	FDR	Process category ([31])	Common with "Immune Exhausted" (Chen et al.)?	Common with "Immune Active" (Chen et al.)?
High group (Bai et al.)	MYC_TARGETS_V1	0.00000368	0.000184	Proliferation	No	No
High group (Bai et al.)	TGF_BETA_SIGNALING	0.0000115	0.000565	Signaling	Yes	No
High group (Bai et al.)	G2M_CHECKPOINT	0.0000255	0.00122	Proliferation	No	No
High group (Bai et al.)	PROTEIN_SECRETION	0.0000309	0.00145	Pathway	No	No
High group (Bai et al.)	EPITHELIAL_ MESENCHYMAL_ TRANSITION	0.000035	0.00161	Development	Yes	No
High group (Bai et al.)	MTORC1_ SIGNALING	0.0000361	0.00163	Signaling	No	No
High group (Bai et al.)	OXIDATIVE_ PHOSPHORYLATION	0.0000407	0.00179	Metabolic	No	No
High group (Bai et al.)	ANDROGEN_ RESPONSE	0.0000619	0.00266	Signaling	No	No
High group (Bai et al.)	HYPOXIA	0.0000835	0.00351	Pathway	Yes	No
High group (Bai et al.)	UV_RESPONSE_DN	0.0001	0.00412	DNA damage	Yes	No

Table 8. Hallmark GSEA pathways de-regulated in the high group of Bai et al.

Using xCell for the investigation of the microenvironment composition, we observed that the high group was enriched in lymphoid (CD8+ naive T-cells) and stromal cells (Table 9), while low group resulted enriched in Immune score (p=0.00000183, FDR= 0.000115), and specifically in different lymphoid cell population such as CD4+ Tem (p=5.82e-9, FDR=3.9e7), CD4+ Tcm (p=0.00000115, FDR= 0.0000746), Immunoscore (p=0.00000183, FDR= 0.000115), and myeloid cells, such

as Basophils (p=0.0000202, FDR= 0.00119) and Eosinophils (p=0.0000318; FDR= 0.00181).

Group enriched	Cell type	pvalue	FDR	xCELL CATEGORY
Low group (Bai et al.)	CD4+ Tem	5.82e-9	3.9e-7	lymphoids
Low group (Bai et al.)	Platelets	1.75e-8	0.00000115	Stem cells
Low group (Bai et al.)	CD4+ Tcm	0.00000115	0.0000746	lymphoids
High group (Bai et al.)	Smooth muscle	0.00000159	0.000102	Stromal
Low group (Bai et al.)	ImmuneScore	0.00000183	0.000115	Not applicable
Low group (Bai et al.)	Hepatocytes	0.00000229	0.000142	Others
Low group (Bai et al.)	Neurons	0.00000434	0.000265	Others
Low group (Bai et al.)	Melanocytes	0.00000445	0.000267	Others
Low group (Bai et al.)	Basophils	0.0000202	0.00119	Myeloids
High group (Bai et al.)	CD8+ naive T-cells	0.0000241	0.0014	Lymphoids
Low group (Bai et al.)	Eosinophils	0.0000318	0.00181	Myeloids
Low group (Bai et al.)	MicroenvironmentScore	0.0000845	0.00473	Not applicable

Table 9. Microenviroment composition (xCell) of high and low groups of Bai et al. signature

Investigating KEGG pathways, the already observed characteristics in Hallmark GSEA were confirmed, or rather that the high-risk group resulted associated with cell cycle, proliferation, and DNA repair pathways (Table 10). Nevertheless, no one of the KEGG pathways resulted associated with the immune profiles described by Chen et al.

Group enriched	Geneset KEGG	p-value	FDR	Common with "Immune Exhausted" (Chen et al.)?	Common with "Immune Active" (Chen et al.)?
Low group (Bai et al.)	RETINOL_METABOLISM	9.23e-9	0.00000114	No	No
High group (Bai et al.)	RIBOSOME	1.72e-7	0.000021	No	No
Low group (Bai et al.)	LINOLEIC_ACID_METABOLISM	7.45e-7	0.0000901	No	No
High group (Bai et al.)	GLYOXYLATE_AND_ DICARBOXYLATE_METABOLISM	0.00000281	0.000337	No	No
Low group (Bai et al.)	DRUG_METABOLISM_ CYTOCHROME_P450	0.0000119	0.00141	No	No
High group (Bai et al.)	PROTEIN_EXPORT	0.0000252	0.00297	No	No
High group (Bai et al.)	NUCLEOTIDE_EXCISION_REPAIR	0.0000328	0.00384	No	No
High group (Bai et al.)	ADHERENS_JUNCTION	0.0000755	0.00876	No	No
High group (Bai et al.)	OXIDATIVE_PHOSPHORYLATION	0.000079	0.00909	No	No
High group (Bai et al.)	SPLICEOSOME	0.0000821	0.00936	No	No
High group (Bai et al.)	UBIQUITIN_MEDIATED_ PROTEOLYSIS	0.000108	0.0122	No	No
High group (Bai et al.)	MISMATCH_REPAIR	0.000136	0.0152	No	No
High group (Bai et al.)	CELL_CYCLE	0.00014	0.0156	No	No
Low group (Bai et al.)	STEROID_HORMONE_BIOSYNTHESIS	0.000174	0.0191	No	No

Table 10. KEGG pathways de-regulated in high and low risk groups of Bai et al.

2. Fang et al. Testing the literature immune signature in Nivactor cohort, we found another signature correlated with PFS. It was the prognostic immune-related HNSCC gene signature of Fang et al. (Figure 21), composed by 10 immune-related genes (DEFB1, EDNRB, ADM, BTC, DKK1, FAM3D, GNRH1, STC2, TNFRSF12A, CTLA4)[137]. Authors claimed that the prognostic value of the signatures was superior to TNM and, the prognostic model divided patients in two groups based on risk scores. Moreover, they observed that the low-risk group showed a higher immune checkpoint expression (observing CTLA-4 and PD-1). In Nivactor cohort the prognostic value was not confirmed in OS analyses

and investigating the biological composition no one of the Hallmark GSEA and KEGG pathways and xCell population resulted significantly de-regulated between the two categories.

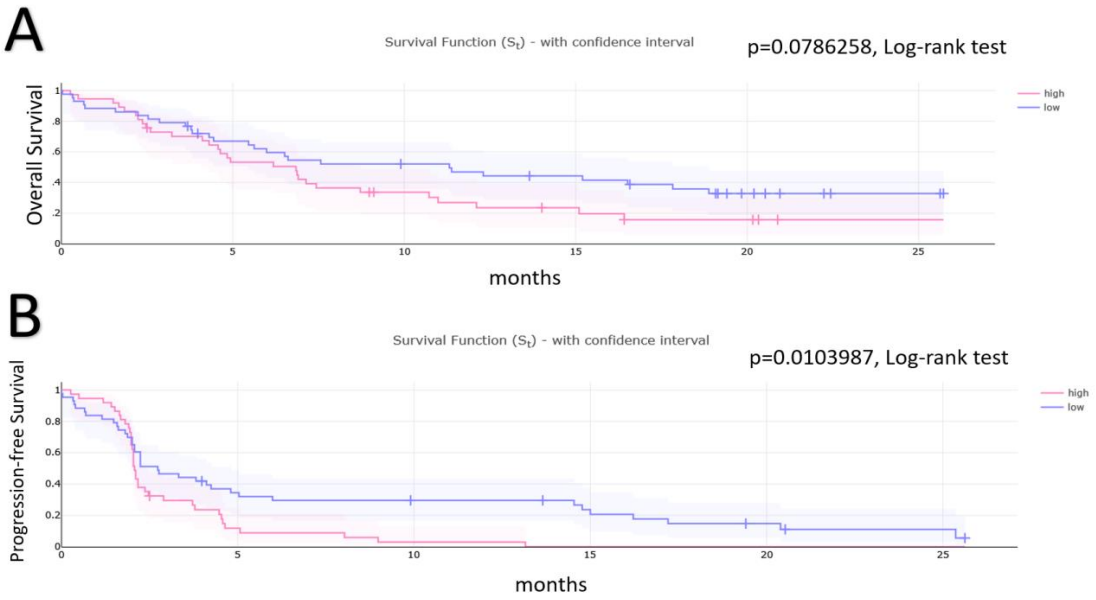


Figure 21 Kaplan-Meier analyses for OS (A) and PFS (B) testing Fang et al. immune-related gene signature on Nivactor cohort (N=80)

3. Hu et al. The third signature tested was the prognostic HNSCC RNA-binding proteins (RBPs) gene expression signature by Hu et al[138]. (Figure 22), composed by six RBPs-related genes, including NCBP2, MKRN3, MRPL47, AZGP1, IGF2BP2, and EZH2. In their paper patients were stratified into two risk group, in which high-risk had the poorer prognosis. Authors individuated a correlation between the RBPs gene expression signatures and the function regulation of immune cells. They observed that the low-risk group was enriched in immune, inflammatory response, fatty acid metabolism, B cell receptor signalling pathway and T cell receptor signalling pathway. Meanwhile high-risk score resulted associated with different signals that reconducted to immunosuppression. In the Nivactor cohort the signatures resulted significantly associated with survival considering PFS analysis (p=0.0378015, Log-rank test), but the same

result was not observed for OS (Figure 21). Moreover, even in the Nivactor cohort, the high-risk group was confirmed to have the poorer prognosis.

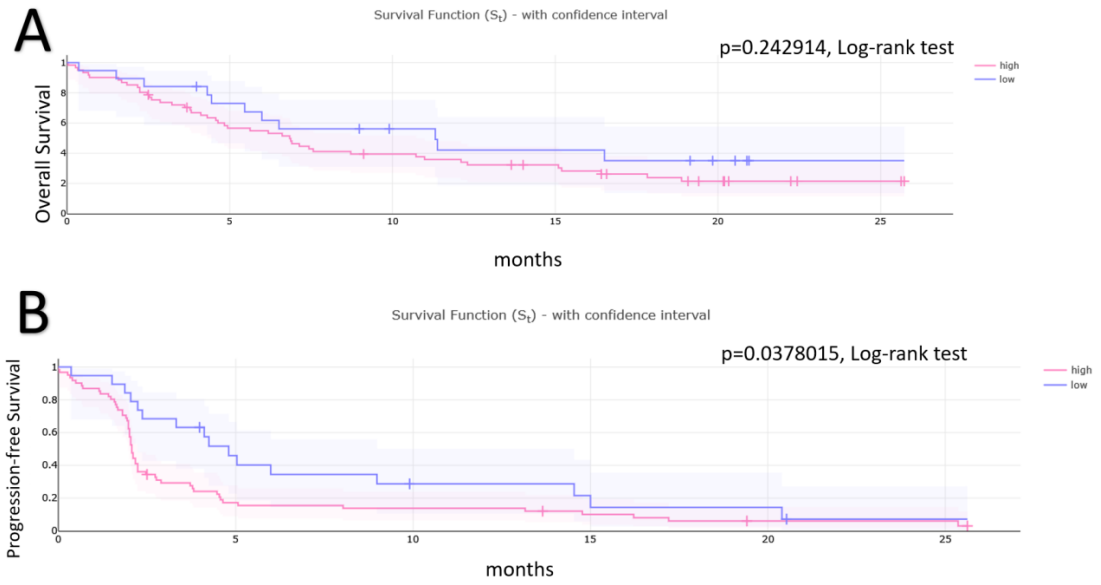


Figure 22. Kaplan-Meier analyses for OS (A) and PFS (B) testing Hu et al. immune-related gene signature on Nivactor cohort (N=80)

In addition, investigating the biological differences of the two-risk group we observed that different Hallmark GSEA pathways resulted enriched in low-risk group of Hu et al, but no one of the pathways resulted enriched in the high-risk group (Table 11). The low risk enriched pathways were related to proliferation, and development. Moreover, 4/10 enriched pathways resulted similar to the "Immune exhausted" HNSCC profile detailed by Chen et al.

Group enriched	Hallmark GSEA pathway	p-value	FDR	Process category ([31])	Common with "Immune Exhausted" (Chen et al.)?	Common with "Immune Active" (Chen et al.)?
Low group (Hu et al.)	G2M_CHECKPOINT	0.00000265	0.000132	proliferation	No	No
Low group (Hu et al.)	REACTIVE_OXYGEN _SPECIES_PATHWAY	0.0000114	0.000558	pathway	No	Yes
Low group (Hu et al.)	E2F_TARGETS	0.0000156	0.000749	proliferation	No	No
Low group (Hu et al.)	MYC_TARGETS_V1	0.0000197	0.000926	proliferation	No	No
Low group (Hu et al.)	OXIDATIVE_PHOSPHORYLATION	0.0000202	0.000928	metabolic	No	No
Low group (Hu et al.)	MITOTIC_SPINDLE	0.0000735	0.00331	proliferation	No	No
Low group (Hu et al.)	CHOLESTEROL _HOMEOSTASIS	0.0000754	0.00332	metabolic	No	No
Low group (Hu et al.)	MTORC1_SIGNALING	0.000108	0.00465	signaling	No	No
Low group (Hu et al.)	DNA_REPAIR	0.000171	0.00717	DNA damage	No	No
Low group (Hu et al.)	UNFOLDED_PROTEIN _RESPONSE	0.000176	0.00721	pathway	No	No
Low group (Hu et al.)	ANDROGEN_RESPONSE	0.000225	0.00902	signaling	No	No
Low group (Hu et al.)	ESTROGEN_RESPONSE _LATE	0.000367	0.0143	signaling	No	No
Low group (Hu et al.)	MYC_TARGETS_V2	0.000407	0.0155	proliferation	No	No
Low group (Hu et al.)	PROTEIN_SECRETION	0.000508	0.0188	pathway	No	No
Low group (Hu et al.)	PI3K_AKT_MTOR_SIGNALING	0.000579	0.0208	signaling	No	Yes
Low group (Hu et al.)	HEME_METABOLISM	0.000759	0.0266	metabolic	No	No
Low group (Hu et al.)	NOTCH_SIGNALING	0.000829	0.0282	signaling	No	No

Table 11. Hallmark GSEA pathways de-regulated in the low group of Hu et al.

The microenvironment between the two groups resulted differently de-regulated, the high-risk group (worse PFS) resulted enriched in eosinophils, endothelial cells, neurons, hepatocytes, NKT and platelets, while the low-risk group (better PFS) was enriched in CLP, Tgd cells, smooth cells (Table 12). Therefore, stromal and lymphoid components were present in both risk categories.

Group enriched	Cell type (xCell)	p-value	FDR	xCell category
High group (Hu et al.)	Eosinophils	0.00000104	0.0000695	Myeloids
High group (Hu et al.)	Neurons	0.00000221	0.000146	Others
High group (Hu et al.)	Endothelial cells	0.00000313	0.000204	Stromal cells
Low group (Hu et al.)	Common Lymphoid progenitor (CLP)	0.00000389	0.000249	Lymphoids
Low group (Hu et al.)	T gamma delta (gd) cells	0.0000141	0.000888	Lymphoids
High group (Hu et al.)	Hepatocytes	0.0000264	0.00164	Others
High group (Hu et al.)	Granulocyte Macrophage progenitor (GMP)	0.000117	0.00715	Stem cells
Low group (Hu et al.)	Smooth muscle	0.000125	0.00751	Stromal cells
High group (Hu et al.)	Natural killer T-cells (NKT)	0.000384	0.0227	Lymphoids
High group (Hu et al.)	Platelets	0.000715	0.0414	Stem cells

Table 12. Microenviroment composition (xCell) of high and low groups of Hu et al. signature

As obtained in the GSEA hallmark analysis, the KEGG pathways resulted enriched in the low-risk group only (Table 13). Even here, the up-regulated pathways were connected to cell cycle and proliferation. Indeed, the profile of low-risk group in the Nivactor cohort resulted more similar to an immune exhaustion than immune activation, and its phenotype reassimilate peculiarity of a more aggressive disease, despite those related to a the better survival. Eventually, the discerned characteristics strongly differed from those described in the original paper of Hu et al.

Group enriched	Geneset KEGG	p-value	FDR	Common with "Immune Exhausted" (Chen et al.)?	Common with "Immune Active" (Chen et al.)?
Low group (Hu et al.)	RIBOSOME	1.96e-7	0.0000241	No	No
Low group (Hu et al.)	OXIDATIVE_PHOSPHORYLATION	0.00000673	0.000821	No	No
Low group (Hu et al.)	BASE_EXCISION_REPAIR	0.0000103	0.00124	No	No
Low group (Hu et al.)	MISMATCH_REPAIR	0.0000422	0.00506	No	No
Low group (Hu et al.)	CELL_CYCLE	0.0000501	0.00596	No	No
Low group (Hu et al.)	GLYOXYLATE_AND _DICARBOXYLATE_METABOLISM	0.0000861	0.0102	No	No
Low group (Hu et al.)	NUCLEOTIDE_EXCISION_REPAIR	0.0000975	0.0114	No	No
Low group (Hu et al.)	UBIQUITIN_MEDIATED _PROTEOLYSIS	0.0000984	0.0114	No	No
Low group (Hu et al.)	SPLICEOSOME	0.000107	0.0123	No	No
Low group (Hu et al.)	PROTEIN_EXPORT	0.000121	0.0138	No	No
Low group (Hu et al.)	DNA_REPLICATION	0.000143	0.0162	No	No
Low group (Hu et al.)	N_GLYCAN_BIOSYNTHESIS	0.000186	0.0208	No	No
Low group (Hu et al.)	GLUTATHIONE_METABOLISM	0.00041	0.0456	No	No
Low group (Hu et al.)	PYRIMIDINE_METABOLISM	0.000411	0.0456	No	No
Low group (Hu et al.)	RNA_POLYMERASE	0.000416	0.0456	No	No

Table 13. KEGG pathways enriched in low group of Hu et al. signature

4. Liu et al (myeloid). Even the fourth gene expression signature resulted associated with PFS in the Nivactor cohort, but not with OS (Figure 23). It was the prognostic HNSCC myeloid gene expression signature of Liu et al.[139]. The risk score was based on the expression of 6 genes (CCL13, CCR7, CD276, IL1B, LYVE1, VEGFC), and the high-risk group presented the worse prognosis. The high score of myeloid signatures was associated by authors to an immunosuppressive status, while the group with the lower expression and lower risk exhibited enrichment in fatty acid metabolism and immune-related pathways. Indeed, in the Nivactor cohort (N=80) we observed an enrichment in the low-risk group of naïve B-cells (p=3.29e-9, FDR=2.2e-7), B-cells (p=0.0000047, FDR=0.00031), Memory B-cells (p=0.0000108, FDR=0.0007), Class-switched memory B-cells

(p=0.0000542; FDR=0.00347) and CD4+ Tem (p=0.000317; FDR= 0.02). Meanwhile, no Hallmark GSEA and KEGG resulted de-regulated in one of the two groups.

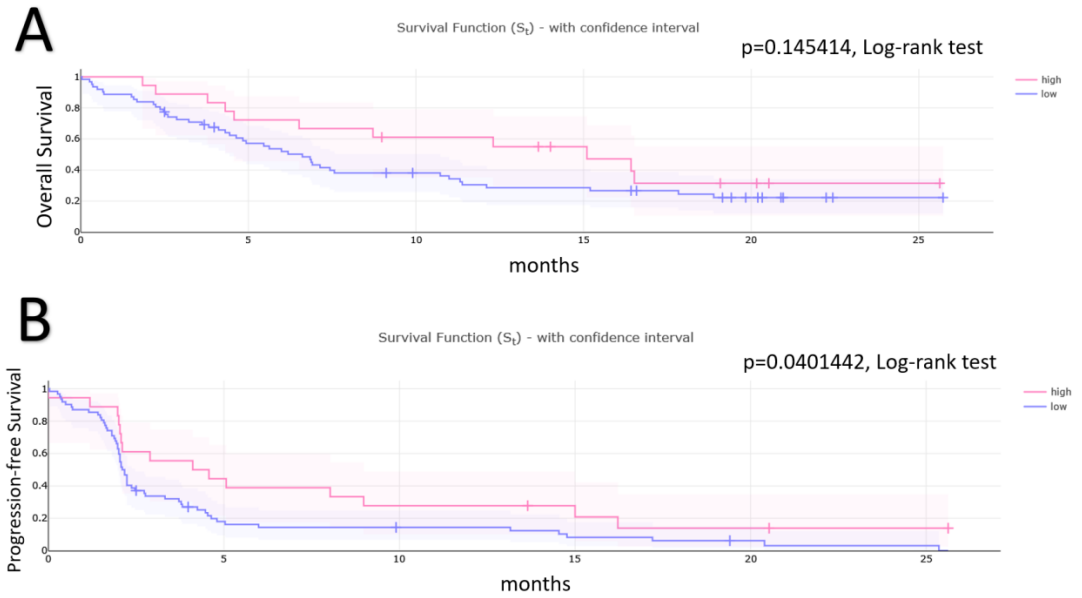


Figure 23 Kaplan-Meier analyses for OS (A) and PFS (B) testing Liu et al. prognostic myeloid immune signature on Nivactor cohort (N=80)

5. Li et al. Moreover, we observed a significant association between PFS and a four gene (PRKCQ, PLAUR, PSMD2, and BMP7) immune signature designed by Li et al.[140] (Figure 24). The signature was designed to predict immune exhaustion in thyroid cancer, and patients were stratified into two different risk groups. The low-risk group (with the better prognosis) was related to immune pathways, while the high-risk group was enriched in some signaling pathways indicative of aggressive malignant behavior.

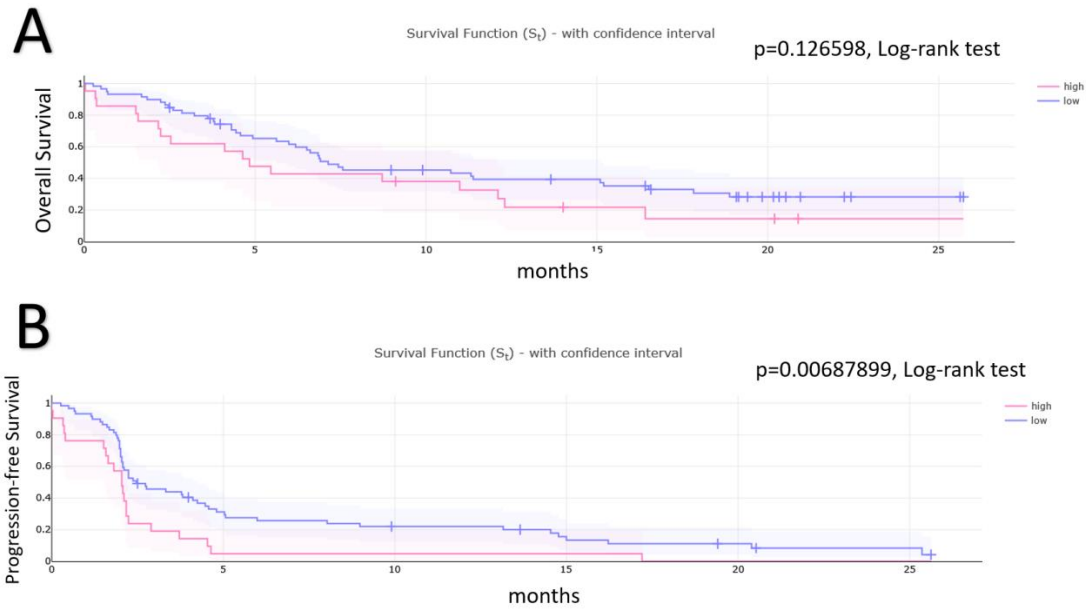


Figure 24 Kaplan-Meier analyses for OS (A) and PFS (B) testing 4-gene immune exhaustion signature of Li et al.

In the high-risk group (worse PFS) 39/50 GSEA Hallmark pathways were up-regulated, as listed in Table 14. Mostly were categorized in proliferation, signaling and pathway. Moreover, 15/39 pathways were common to the profile of “Immune exhausted” described by Chen et al, and 9/39 with the “Immune active” profile.

Group enriched	Hallmark GSEA pathway	p-value	FDR	Process category	Common with "Immune Exhausted" (Chen et al.)?	Common with "Immune Active" (Chen et al.)?
High group (Li et al.)	ANDROGEN_RESPONSE	0.00000426	0.000213	Signaling	No	No
High group (Li et al.)	PROTEIN_SECRETION	0.00000647	0.000317	Pathway	No	No
High group (Li et al.)	TGF_BETA _SIGNALING	0.0000136	0.000651	Signaling	Yes	No
High group (Li et al.)	MYC_TARGETS_V1	0.0000143	0.000673	proliferation	No	No
High group (Li et al.)	UNFOLDED_PROTEIN _RESPONSE	0.0000246	0.00113	pathway	No	No
High group (Li et al.)	G2M_CHECKPOINT	0.0000383	0.00172	proliferation	No	No
High group (Li et al.)	OXIDATIVE_ PHOSPHORYLATION	0.0000554	0.00244	metabolic	No	No
High group (Li et al.)	MITOTIC_SPINDLE	0.00009	0.00387	Proliferation	No	No
High group (Li et al.)	EPITHELIAL_MESENCHYMAL _TRANSITION	0.0000987	0.00414	development	Yes	No
High group (Li et al.)	TNFA_SIGNALING _VIA_NFKB	0.000115	0.00473	signaling	Yes	No
High group (Li et al.)	MTORC1_SIGNALING	0.000125	0.00501	signaling	No	No
High group (Li et al.)	APOPTOSIS	0.000127	0.00501	pathway	No	Yes
High group (Li et al.)	UV_RESPONSE_DN	0.000141	0.00537	DNA damage	Yes	No
High group (Li et al.)	DNA_REPAIR	0.000223	0.00826	DNA damage	No	No
High group (Li et al.)	CHOLESTEROL _HOMEOSTASIS	0.000226	0.00826	metabolic	No	No
High group (Li et al.)	HYPOXIA	0.00024	0.0084	pathway	Yes	No
High group (Li et al.)	GLYCOLYSIS	0.000276	0.00937	Metabolic	No	No
High group (Li et al.)	PI3K_AKT_MTOR _SIGNALING	0.000311	0.0103	signaling	No	Yes

High group (Li et al.)	P53_PATHWAY	0.000361	0.0115	proliferation	No	No
High group (Li et al.)	WNT_BETA_CATENIN _SIGNALING	0.000426	0.0132	Signaling	No	No
High group (Li et al.)	MYC_TARGETS_V2	0.0005	0.015	Proliferation	No	No
High group (Li et al.)	IL2_STAT5_SIGNALING	0.000548	0.0159	signaling	Yes	Yes
High group (Li et al.)	ADIPOGENESIS	0.000607	0.017	development	Yes	No
High group (Li et al.)	HEDGEHOG_SIGNALING	0.000664	0.0179	Signaling	No	No
High group (Li et al.)	INFLAMMATORY _RESPONSE	0.000742	0.0193	Immune	Yes	Yes
High group (Li et al.)	NOTCH_SIGNALING	0.000748	0.0193	Signaling	No	No
High group (Li et al.)	COMPLEMENT	0.000762	0.0193	Immune	Yes	Yes
High group (Li et al.)	ANGIOGENESIS	0.00077	0.0193	Development	Yes	No
High group (Li et al.)	REACTIVE_OXYGEN _SPECIES_PATHWAY	0.000902	0.0198	Pathway	No	Yes
High group (Li et al.)	UV_RESPONSE_UP	0.000909	0.0198	DNA damage	No	Yes
High group (Li et al.)	HEME_METABOLISM	0.00109	0.0217	Metabolic	No	No
High group (Li et al.)	E2F_TARGETS	0.00121	0.023	Proliferation	No	No
High group (Li et al.)	APICAL_JUNCTION	0.00129	0.0232	Cellular component	Yes	Yes
High group (Li et al.)	IL6_JAK_STAT3_SIGNALING	0.0015	0.0255	Immune	Yes	Yes
High group (Li et al.)	INTERFERON_GAMMA _RESPONSE	0.00192	0.0307	Immune	No	No
High group (Li et al.)	ESTROGEN_RESPONSE _LATE	0.00199	0.0307	signaling	No	No

Table 14. Enriched Hallmark GSEA pathways (Li et al.)

The evaluation of the microenvironment composition evidenced an abundance of lymphoids, myeloids and stem cells in the low-risk group, result confirmed also by the xCell cumulative ImmunoScore. Therefore, the profile of the low-risk group observed in the original paper was confirmed by xCell analyses. Meanwhile, the high-risk group was characterized by stromal and stem cells (Table 15).

Group enriched	Cell type (xCell)	p-value	FDR	xCell category
Low group (Li et al.)	CD4+ Tem	5.82e-9	3.9e-7	Lymphoids
Low group (Li et al.)	Platelets	1.75e-8	0.00000115	Stem cells
Low group (Li et al.)	CD4+ Tcm	0.00000115	0.0000746	Lymphoids
High group (Li et al.)	Smooth muscle	0.00000159	0.000102	Stromal
Low group (Li et al.)	ImmuneScore	0.00000183	0.000115	NA
Low group (Li et al.)	Hepatocytes	0.00000229	0.000142	Others
Low group (Li et al.)	Neurons	0.00000434	0.000265	Others
Low group (Li et al.)	Melanocytes	0.00000445	0.000267	Others
Low group (Li et al.)	Basophils	0.0000202	0.00119	Myeloids
Low group (Li et al.)	CD8+ naive T-cells	0.0000241	0.0014	Lymphoids
Low group (Li et al.)	Eosinophils	0.0000318	0.00181	Myeloids
Low group (Li et al.)	MicroenvironmentScore	0.0000845	0.00473	NA
Low group (Li et al.)	MPP	0.000106	0.0058	Stem cells
Low group (Li et al.)	CMP	0.000113	0.0061	Stem cells
High group (Li et al.)	CLP	0.000191	0.0101	Stem cells
Low group (Li et al.)	CD4+ T-cells	0.000269	0.014	Lymphoids

Table 15. Microenviroment composition of the groups (Li et al.)

The numerous KEGG pathways, that were all significantly enriched in Low-risk group, evidenced an upregulation in cell cycle, proliferation, and metabolism (Table 16). Seven KEGG pathways were in common with the profile of immune exhaustion defined by Chen et al, and two KEGG pathways with the profile of “Immune active”.

Group enriched	Geneset KEGG	p-value	FDR	Common with "Immune Exhausted" (Chen et al.)?	Common with "Immune Active" (Chen et al.)?
Low group (Li et al.)	GLYOXYLATE_AND_DICARBOXYLATE _METABOLISM	0.00000166	0.000204	No	No
Low group (Li et al.)	RIBOSOME	0.00000209	0.000255	No	No
Low group (Li et al.)	N_GLYCAN_BIOSYNTHESIS	0.0000116	0.0014	No	No
Low group (Li et al.)	NUCLEOTIDE_EXCISION _REPAIR	0.0000125	0.0015	No	No
Low group (Li et al.)	TGF_BETA_SIGNALING _PATHWAY	0.0000184	0.0022	Yes	No
Low group (Li et al.)	CELL_ADHESION_MOLECULES _CAMS	0.0000321	0.00379	No	No
Low group (Li et al.)	MISMATCH_REPAIR	0.0000373	0.00436	No	No
Low group (Li et al.)	CELL_CYCLE	0.0000557	0.00646	No	No
Low group (Li et al.)	UBIQUITIN_MEDIATED _PROTEOLYSIS	0.0000658	0.00757	No	No
Low group (Li et al.)	ALLOGRAFT_REJECTION	0.0000686	0.00782	No	No
Low group (Li et al.)	PATHWAYS_IN_CANCER	0.0000799	0.00903	Yes	No
Low group (Li et al.)	P53_SIGNALING_PATHWAY	0.000111	0.0124	No	No
Low group (Li et al.)	PROTEIN_EXPORT	0.000113	0.0125	No	No
Low group (Li et al.)	WNT_SIGNALING_PATHWAY	0.000141	0.0155	Yes	No
Low group (Li et al.)	ANTIGEN_PROCESSING _AND_PRESENTATION	0.000145	0.0159	No	No
Low group (Li et al.)	SPLICEOSOME	0.00016	0.0173	No	No
Low group (Li et al.)	NOTCH_SIGNALING_PATHWAY	0.000162	0.0173	No	No
Low group (Li et al.)	ECM_RECEPTOR_INTERACTION	0.000164	0.0174	Yes	No

Low group (Li et al.)	LYSOSOME	0.00019	0.02	No	No
Low group (Li et al.)	ADHERENS_JUNCTION	0.000191	0.02	No	No
Low group (Li et al.)	BASE_EXCISION_REPAIR	0.000192	0.02	No	No
Low group (Li et al.)	DNA_REPLICATION	0.000225	0.0229	No	No
Low group (Li et al.)	FOCAL_ADHESION	0.000228	0.0231	No	No
Low group (Li et al.)	TOLL_LIKE_RECEPTOR _SIGNALING_PATHWAY	0.000252	0.0252	No	Yes
Low group (Li et al.)	RNA_POLYMERASE	0.000272	0.027	No	No
Low group (Li et al.)	MTOR_SIGNALING_PATHWAY	0.000284	0.0278	No	No
Low group (Li et al.)	BASAL_CELL_CARCINOMA	0.000324	0.0315	Yes	No
Low group (Li et al.)	OXIDATIVE_PHOSPHORYLATION	0.00033	0.0316	No	No
Low group (Li et al.)	GLYCOSAMINOGLYCAN _BIOSYNTHESIS_KERATAN_SULFATE	0.000393	0.0373	Yes	No
Low group (Li et al.)	MAPK_SIGNALING_PATHWAY	0.000409	0.0384	No	No
Low group (Li et al.)	PYRIMIDINE_METABOLISM	0.000435	0.0404	No	No
Low group (Li et al.)	APOPTOSIS	0.000436	0.0404	No	Yes
Low group (Li et al.)	GLYCOSPHINGOLIPID _BIOSYNTHESIS_GLOBO_SERIES	0.000469	0.0427	No	No

Table 16. KEGG pathways de-regulated in high and low risk groups of Li et al.

6. Rooney et al. The sixth signature significantly associated survival was the cytolytic immune signature of Rooney et al. (Figure 25)[141]. The signature was designed to quantify the effective and natural anti-tumor immunity, which requires a cytolytic immune response. The more the cytolytic activity (high expression of the signature) the less was the immune response. In the original paper the investigation was biology-based, and the association with the prognosis was not performed. However,

in Nivactor cohort we interrogated whether the signatures could be translated to survival data for R/M HNSCC patients. Specifically, in Nivactor cohort the signature resulted significantly correlated to PFS (Figure 24), and high expression of the signature was correlated to a worse survival.

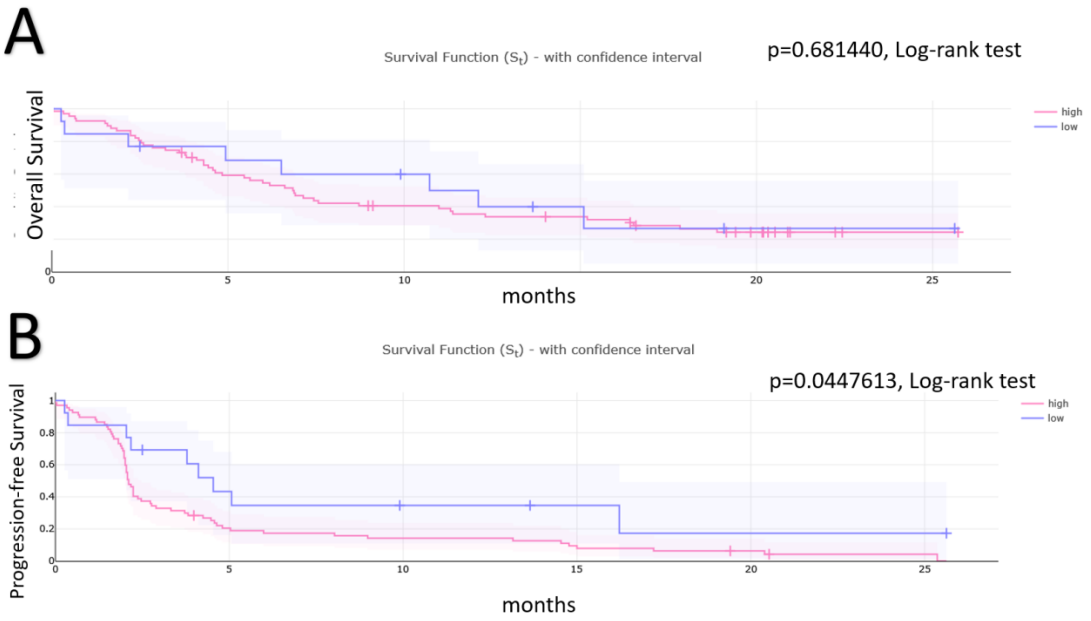


Figure 25 Kaplan-Meier analyses for OS (A) and PFS (B) testing signature of Rooney et al.

Several Hallmark GSEA pathways characterized the High group of the gene expression signature of Rooney et al. (Table 17).

Group enriched	Hallmark GSEA pathway	p-value	FDR	Process category	Common with "Immune Exhausted" (Chen et al.)?	Common with "Immune Active" (Chen et al.)?
High group (Rooney et al.)	INTERFERON_ALPHA_RESPONSE	3.11e-11	1.56e-9	immune	No	No
High group (Rooney et al.)	INTERFERON_GAMMA_RESPONSE	4.88e-11	2.39e-9	immune	No	No
High group (Rooney et al.)	ALLOGRAFT_REJECTION	5.63e-10	2.7e-8	immune	Yes	Yes
High group (Rooney et al.)	COMPLEMENT	1.3e-7	0.00000612	immune	Yes	Yes
High group (Rooney et al.)	IL6_JAK_STAT3_SIGNALING	2.85e-7	0.0000131	immune	Yes	Yes
High group (Rooney et al.)	INFLAMMATORY_RESPONSE	8.96e-7	0.0000403	immune	Yes	Yes
High group (Rooney et al.)	IL2_STAT5_SIGNALING	0.0000084	0.00037	signaling	Yes	Yes
High group (Rooney et al.)	KRAS_SIGNALING_UP	0.0000108	0.000466	signaling	No	Yes
High group (Rooney et al.)	OXIDATIVE_PHOSPHORYLATION	0.0000242	0.00102	metabolic	No	No
High group (Rooney et al.)	DNA_REPAIR	0.000046	0.00188	DNA damage	No	No

High group (Rooney et al.)	PI3K_AKT_MTOR_SIGNALING	0.0000657	0.00263	signaling	No	Yes
High group (Rooney et al.)	APOPTOSIS	0.000087	0.00339	pathway	No	Yes
High group (Rooney et al.)	TNFA_SIGNALING_VIA_NFKB	0.000112	0.00425	signaling	Yes	No
High group (Rooney et al.)	COAGULATION	0.000141	0.00521	immune	Yes	No
High group (Rooney et al.)	ADIPOGENESIS	0.000159	0.00571	development	Yes	No
High group (Rooney et al.)	PROTEIN_SECRETION	0.000173	0.00605	pathway	No	No
High group (Rooney et al.)	UV_RESPONSE_DN	0.000183	0.00621	DNA damage	Yes	No
High group (Rooney et al.)	TGF_BETA_SIGNALING	0.000191	0.00629	signaling	Yes	No
High group (Rooney et al.)	APICAL_JUNCTION	0.000241	0.00773	Cellular component	Yes	Yes
High group (Rooney et al.)	ANDROGEN_RESPONSE	0.000255	0.00791	signaling	No	No
High group (Rooney et al.)	MITOTIC_SPINDLE	0.000259	0.00791	proliferation	No	No
High group (Rooney et al.)	CHOLESTEROL_HOMEOSTASIS	0.00027	0.00791	metabolic	No	No
High group (Rooney et al.)	MYC_TARGETS_V1	0.000277	0.00791	proliferation	No	No
High group (Rooney et al.)	UNFOLDED_PROTEIN_RESPONSE	0.000311	0.00841	pathway	No	No
High group (Rooney et al.)	REACTIVE_OXYGEN _SPECIES_PATHWAY	0.000342	0.00889	pathway	No	Yes
High group (Rooney et al.)	APICAL_SURFACE	0.000407	0.0102	Cellular component	No	Yes
High group (Rooney et al.)	MTORC1_SIGNALING	0.000467	0.0112	signaling	No	No
High group (Rooney et al.)	XENOBIOTIC_METABOLISM	0.000472	0.0112	metabolic	No	Yes
High group (Rooney et al.)	HEME_METABOLISM	0.000511	0.0112	metabolic	No	No
High group (Rooney et al.)	WNT_BETA_CATENIN_SIGNALING	0.00052	0.0112	signaling	No	No
High group (Rooney et al.)	GLYCOLYSIS	0.000622	0.0124	metabolic	No	No
High group (Rooney et al.)	PEROXISOME	0.000674	0.0128	Cellular component	No	No
High group (Rooney et al.)	P53_PATHWAY	0.000687	0.0128	proliferation	No	No
High group (Rooney et al.)	EPITHELIAL _MESENCHYMAL_TRANSITION	0.000723	0.0128	development	Yes	No
High group (Rooney et al.)	FATTY_ACID_METABOLISM	0.000748	0.0128	metabolic	No	No
High group (Rooney et al.)	G2M_CHECKPOINT	0.00078	0.0128	proliferation	No	No
High group (Rooney et al.)	E2F_TARGETS	0.00108	0.0152	proliferation	No	No
High group (Rooney et al.)	HEDGEHOG_SIGNALING	0.00115	0.0152	signaling	No	No
High group (Rooney et al.)	UV_RESPONSE_UP	0.00117	0.0152	DNA damage	No	Yes
High group (Rooney et al.)	ESTROGEN_RESPONSE_EARLY	0.00132	0.0152	signaling	No	No
High group (Rooney et al.)	NOTCH_SIGNALING	0.0015	0.0152	signaling	No	No
High group (Rooney et al.)	ESTROGEN_RESPONSE_LATE	0.00231	0.0208	signaling	No	No
High group (Rooney et al.)	MYC_TARGETS_V2	0.00242	0.0208	proliferation	No	No
High group (Rooney et al.)	HYPOXIA	0.0037	0.0259	pathway	Yes	No

Table 17. Enriched Hallmark GSEA pathways (Rooney et al.)

While low expression group was distinguished by an enrichment in Neurons, pDC, Hepatocytes, Megakaryocytes, and ly Endothelial cells (Table 18).

Group enriched	Cell type (xCell)	p-value	FDR	xCell category
High group (Rooney et al.)	Macrophages	3.44e-10	2.3e-8	Myeloids
High group (Rooney et al.)	Macrophages M1	2.44e-7	0.0000161	Myeloids
Low group (Rooney et al.)	Neurons	0.00000795	0.000517	Others
High group (Rooney et al.)	CD4+ memory T-cells	0.00000837	0.000536	Lymphoids
Low group (Rooney et al.)	pDC	0.0000297	0.00187	Myeloids
Low group (Rooney et al.)	Hepatocytes	0.0000698	0.00433	Others
High group (Rooney et al.)	CLP	0.000079	0.00482	Stem cells
High group (Rooney et al.)	Plasma cells	0.000147	0.0088	Lymphoids
Low group (Rooney et al.)	Megakaryocytes	0.000152	0.00897	Stem cells
High group (Rooney et al.)	CD8+ Tcm	0.000334	0.0194	Lymphoids
High group (Rooney et al.)	Tgd cells	0.000399	0.0227	Lymphoids
High group (Rooney et al.)	CD8+ T-cells	0.00065	0.0364	Lymphoids
Low group (Rooney et al.)	ly Endothelial cells	0.000747	0.0411	Stromal cells

Table 18. Microenviroment composition of the groups (Rooney et al.)

The signature evidenced a strong deregulation also in several KEGG pathways for the high group, the one with the worse prognosis (Table 19).

Group enriched	Geneset KEGG	p-value	FDR	Common with "Immune Exhausted" (Chen et al.)?	Common with "Immune Active" (Chen et al.)?
High group (Rooney et al.)	ALLOGRAFT_REJECTION	7.07e-12	8.7e-10	No	Yes
High group (Rooney et al.)	ANTIGEN_PROCESSING _AND_PRESENTATION	1.06e-8	0.00000129	No	Yes
High group (Rooney et al.)	CELL_ADHESION _MOLECULES_CAMS	5.26e-7	0.0000637	No	No
High group (Rooney et al.)	LYSOSOME	0.00000209	0.000251	No	No
High group (Rooney et al.)	N_GLYCAN _BIOSYNTHESIS	0.00000293	0.000348	No	Yes
High group (Rooney et al.)	CYTOKINE_CYTOKINE _RECEPTOR_INTERACTION	0.00000302	0.000357	No	Yes
High group (Rooney et al.)	PRIMARY_IMMUNODEFICIENCY	0.00000355	0.000416	No	Yes
High group (Rooney et al.)	PROTEASOME	0.00000486	0.000564	No	No
High group (Rooney et al.)	CHEMOKINE_SIGNALING _PATHWAY	0.00000839	0.000965	No	Yes
High group (Rooney et al.)	JAK_STAT_ SIGNALING_PATHWAY	0.00000854	0.000974	No	Yes
High group (Rooney et al.)	NATURAL_KILLER_CELL _MEDIATED_CYTOTOXICITY	0.0000125	0.00141	No	Yes
High group (Rooney et al.)	TOLL_LIKE_RECEPTOR _SIGNALING_PATHWAY	0.0000126	0.00141	No	Yes

High group (Rooney et al.)	NICOTINATE_AND _NICOTINAMIDE_METABOLISM	0.0000182	0.00202	No	No
High group (Rooney et al.)	NOD_LIKE_RECEPTOR _SIGNALING_PATHWAY	0.0000207	0.00228	No	No
High group (Rooney et al.)	T_CELL_RECEPTOR _SIGNALING_PATHWAY	0.000021	0.00229	No	Yes
High group (Rooney et al.)	OTHER_GLYCAN _DEGRADATION	0.0000213	0.0023	No	No
High group (Rooney et al.)	OXIDATIVE_PHOSPHORYLATION	0.0000216	0.00231	No	No
High group (Rooney et al.)	LYSINE_DEGRADATION	0.0000288	0.00305	No	No
High group (Rooney et al.)	B_CELL_RECEPTOR _SIGNALING_PATHWAY	0.0000443	0.00466	No	Yes
High group (Rooney et al.)	UBIQUITIN_MEDIATED _PROTEOLYSIS	0.0000533	0.00554	No	No
High group (Rooney et al.)	ETHER_LIPID _METABOLISM	0.0000568	0.00585	No	No
High group (Rooney et al.)	VEGF_SIGNALING_PATHWAY	0.000068	0.00694	No	No
High group (Rooney et al.)	PROTEIN_EXPORT	0.0000724	0.00731	No	No
High group (Rooney et al.)	RIBOSOME	0.000101	0.0101	No	No
High group (Rooney et al.)	GLYCOSAMINOGLYCAN _BIOSYNTHESIS_KERATAN_SULFATE	0.000106	0.0105	No	No
High group (Rooney et al.)	APOPTOSIS	0.000108	0.0106	No	Yes
High group (Rooney et al.)	P53_SIGNALING_PATHWAY	0.000121	0.0118	No	No
High group (Rooney et al.)	PANTOTHENATE_AND_COA_BIOSYNTHESIS	0.000126	0.0121	No	No
High group (Rooney et al.)	PYRIMIDINE_METABOLISM	0.00013	0.0123	No	No
High group (Rooney et al.)	CITRATE_CYCLE _TCA_CYCLE	0.000147	0.0138	No	No
High group (Rooney et al.)	BETA_ALANINE_METABOLISM	0.000166	0.0155	No	Yes
High group (Rooney et al.)	GLYCOSAMINOGLYCAN _DEGRADATION	0.000167	0.0155	No	No
High group (Rooney et al.)	SPLICEOSOME	0.000168	0.0155	No	No
High group (Rooney et al.)	GLYCEROPHOSPHOLIPID _METABOLISM	0.000172	0.0155	No	No
High group (Rooney et al.)	ENDOCYTOSIS	0.000173	0.0155	No	No
High group (Rooney et al.)	CELL_CYCLE	0.00018	0.0158	No	No
High group (Rooney et al.)	PATHWAYS_IN_CANCER	0.000248	0.0216	Yes	No
High group (Rooney et al.)	GLYOXYLATE_AND _DICARBOXYLATE_METABOLISM	0.000253	0.0218	No	No
High group (Rooney et al.)	RNA_DEGRADATION	0.000256	0.0218	No	No
High group (Rooney et al.)	ADHERENS_JUNCTION	0.000263	0.0221	No	No
High group (Rooney et al.)	COMPLEMENT_AND _COAGULATION_CASCADES	0.000293	0.0243	No	No
High group (Rooney et al.)	GNRH_SIGNALING_PATHWAY	0.000303	0.0249	No	No
High group (Rooney et al.)	REGULATION_OF_AUTOPHAGY	0.000325	0.0264	No	No
High group (Rooney et al.)	REGULATION_OF _ACTIN_CYTOSKELETON	0.000421	0.0337	No	No
High group (Rooney et al.)	MAPK_SIGNALING_PATHWAY	0.000459	0.0363	No	No
High group (Rooney et al.)	MTOR_SIGNALING_PATHWAY	0.000481	0.0375	No	No
High group (Rooney et al.)	RNA_POLYMERASE	0.000539	0.0415	No	No
High group (Rooney et al.)	RIG_I_LIKE_RECEPTOR _SIGNALING_PATHWAY	0.000542	0.0415	No	No
High group (Rooney et al.)	NUCLEOTIDE_EXCISION_REPAIR	0.000564	0.0423	No	No

High group (Rooney et al.)	GLYCOSAMINOGLYCAN _BIOSYNTHESIS_CHONDROITIN_SULFATE	0.00061	0.0451	No	No
High group (Rooney et al.)	AMINO_SUGAR_AND _NUCLEOTIDE_SUGAR_METABOLISM	0.000664	0.0485	No	No
High group (Rooney et al.)	TGF_BETA_SIGNALING_PATHWAY	0.00068	0.0489	Yes	No
High group (Rooney et al.)	PHOSPHATIDYLINOSITOL _SIGNALING_SYSTEM	0.000691	0.0491	No	Yes

Table 19. KEGG pathways de-regulated in the high-risk groups of Rooney et al.

7. Ayers et al. The seventh and last gene expression signature associated with PFS in Nivactor cohort was the signature of Ayers et al. The signature was named "IFN-γ signature" and it was composed by 6 genes (IDO1, CXCL10, CXCL9, HLA-DRA, STA1, IFNG). Differently from the other signatures, the signature of Ayers et al. was built on clinical trials (KEYNOTE-001, KEYNOTE-012, and KEYNOTE-028) of oncologic patients treated with immunotherapy, also including R/M HNSCC patients. In the three trials the signature demonstrated to be associated with response (the patients with the higher expression were those that experienced more response); nevertheless, its prognostic value was not investigated. Testing the signature in the Nivactor cohort we observed a correlation with survival (Figure 26). However, the correlation between the signature and survival evidence that patients with a higher expression of the IFN-γ signature was prognostic to a worse prognosis.

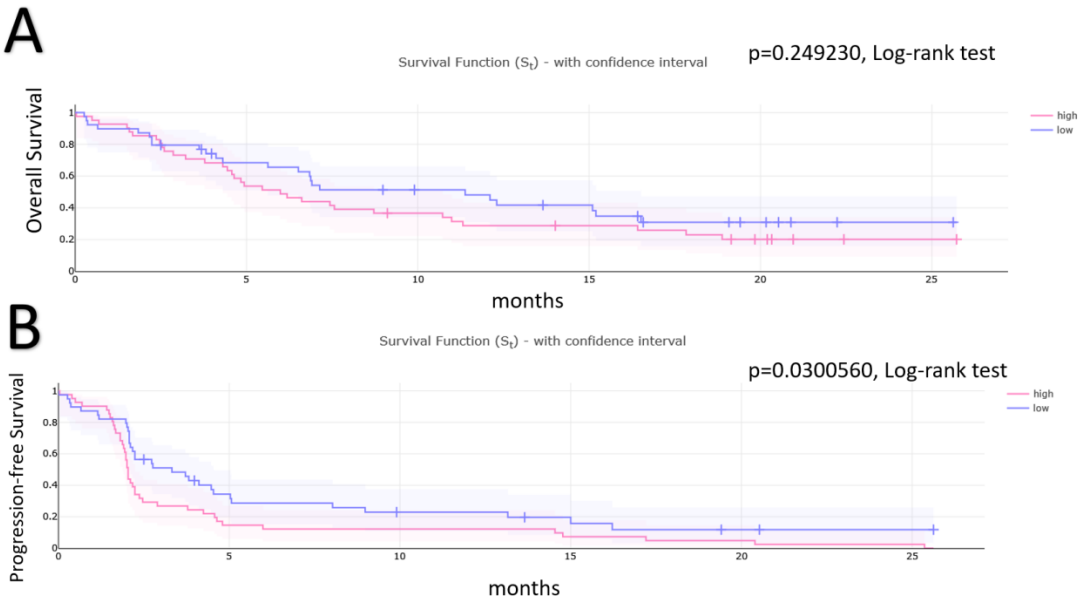


Figure 26 Kaplan-Meier analyses for OS (A) and PFS (B) testing 6-gene signature of Ayers et al.

Even if the statistic power of the signature in the survival analysis was limited, and the high expression group evidenced a worse prognosis, we observed a strong enrichment in biology characteristics. Notably, all the gene-set hallmark were significantly enriched in the high group, except for KRAS_SIGNALING_DN and PANCREAS_BETA_CELLS pathways, and 103 KEGG resulted de-regulated. Moreover, the low and high group were characterized by several cell populations, both comprising lymphoid, myeloid, stromal and other cell types (Table 20).

Group enriched	Cell type (xCell)	p-value	FDR	xCell category
High group (Ayers et al.)	Macrophages	1.76e-16	1.18e-14	Myeloids
High group (Ayers et al.)	Macrophages M1	1.07e-10	7.03e-9	Myeloids
High group (Ayers et al.)	CD4+ memory T-cells	3.52e-9	2.29e-7	Lymphoids
Low group (Ayers et al.)	Neurons	7.34e-9	4.7e-7	Others
High group (Ayers et al.)	CLP	1.68e-8	0.00000106	Lymphoids
High group (Ayers et al.)	CD8+ T-cells	5.43e-8	0.00000337	Lymphoids
Low group (Ayers et al.)	Hepatocytes	1.61e-7	0.0000098	Others
Low group (Ayers et al.)	Eosinophils	0.00000177	0.000106	Myeloids
High group (Ayers et al.)	aDC	0.00000428	0.000253	Myeloids

High group (Ayers et al.)	pDC	0.00000471	0.000273	Myeloids
Low group (Ayers et al.)	pro B-cells	0.00000597	0.000341	Lymphoids
High group (Ayers et al.)	Plasma cells	0.0000194	0.00109	Lymphoids
High group (Ayers et al.)	Tregs	0.0000222	0.00122	Lymphoids
Low group (Ayers et al.)	ly Endothelial cells	0.0000226	0.00122	Stromal
High group (Ayers et al.)	Smooth muscle	0.0000541	0.00287	Stromal
Low group (Ayers et al.)	CMP	0.0000631	0.00328	Stem cells
High group (Ayers et al.)	CD4+ T-cells	0.000153	0.0078	Lymphoids
Low group (Ayers et al.)	Basophils	0.000188	0.0094	Myeloids
Low group (Ayers et al.)	Megakaryocytes	0.000548	0.0268	Stem cells
High group (Ayers et al.)	B-cells	0.000757	0.0363	Lymphoids
Low group (Ayers et al.)	Melanocytes	0.000969	0.0455	Others
High group (Ayers et al.)	Class-switched memory B-cells	0.00102	0.0468	Lymphoids

Table 20. Microenviroment composition of the groups (Ayers et al.)

8. Liu et al. (immune) Contrary to the other signatures, the eighth and last signature resulted associated to OS in Nivactor cohort (p=0.00988619, Log-rank test Figure 26). The signature was a prognostic six-gene immune-risk model constructed on HNSCC databases by Liu et al.[142] and it was composed by DKK1, HBEGF, RNASE7, TNFRSF12A, INHBA and PIK3R3 genes. The signatures stratified patients into two risk group: high with the worse prognosis, and low with the better prognosis. The stratification was confirmed in Nivactor cohort; however, the biology discerned in Nivactor was discordant from the one published by Liu et al in the original paper.

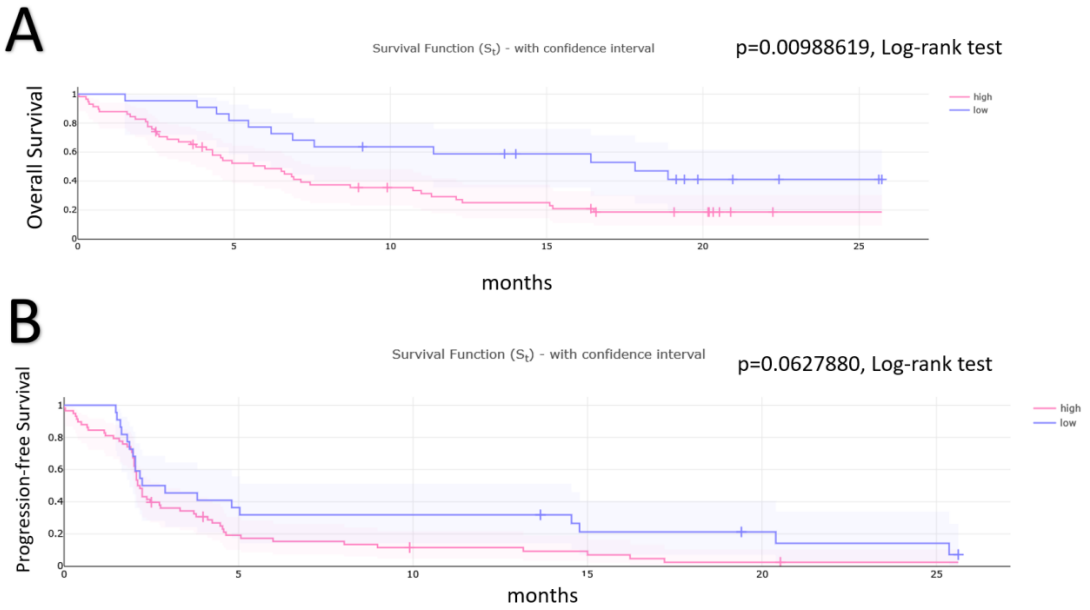


Figure 27. Kaplan-Meier analyses for OS (A) and PFS (B) testing Liu et al. prognostic immune-risk model on Nivactor cohort (N=80)

In the Nivactor cohort (N=80) the high-risk group was enriched in different pathways related to proliferation, development, and signalling, while no Hallmark GSEA resulted de-regulated in the low-risk group (Table 21). This profile resulted in line with a more aggressive phenotype and a worse prognosis.

Group enriched	Hallmark GSEA pathway	p-value	FDR	Process category	Common with "Immune Exhausted" (Chen et al.)?	Common with "Immune Active" (Chen et al.)?
High group (Liu et al.)	ANDROGEN_RESPONSE	0.00000426	0.000213	signaling	No	No
High group (Liu et al.)	PROTEIN_SECRETION	0.00000647	0.000317	pathways	No	No
High group (Liu et al.)	TGF_BETA_SIGNALING	0.0000136	0.000651	signaling	Yes	No
High group (Liu et al.)	MYC_TARGETS_V1	0.0000143	0.000673	proliferation	No	No
High group (Liu et al.)	UNFOLDED_PROTEIN _RESPONSE	0.0000246	0.00113	pathways	No	No
High group (Liu et al.)	G2M_CHECKPOINT	0.0000383	0.00172	proliferation	No	No
High group (Liu et al.)	OXIDATIVE _PHOSPHORYLATION	0.0000554	0.00244	metabolic	No	No

High group (Liu et al.)	MITOTIC_SPINDLE	0.00009	0.00387	proliferation	No	No
High group (Liu et al.)	EPITHELIAL_MESENCHYMAL_TRANSITION	0.0000987	0.00414	development	Yes	No
High group (Liu et al.)	TNFA_SIGNALING_VIA_NFKB	0.000115	0.00473	signaling	Yes	No

Table 21. Hallmark GSEA pathways de-regulated in the high group of Liu et al.

Meanwhile, in the low-risk group we observed a significant enrichment of platelets, hepatocytes, adipocytes, pro B-cells and neurons and the high-risk group was characterized by an enrichment in Smooth muscle and Common Lymphoid progenitor (CLP), while the low group by (Table 22). The enrichment in lymphoid components observed by Liu et al was not confirmed in Nivactor cohort.

Group enriched	Cell type (xCell)	p-value	FDR	XCELL CATEGORY
Low group (Liu et al.)	Platelets	8.6e-11	5.76e-9	Stem cells
Low group (Liu et al.)	Hepatocytes	0.00000124	0.0000816	Others
High group (Liu et al.)	Smooth muscle	0.00000154	0.0000998	Stromal
Low group (Liu et al.)	Adipocytes	0.00000313	0.0002	Stromal
Low group (Liu et al.)	pro B-cells	0.0000126	0.000795	Lymphoids
Low group (Liu et al.)	Neurons	0.0000296	0.00184	Others
High group (Liu et al.)	Common Lymphoid progenitor (CLP)	0.0000305	0.00186	Stem cells

Table 22. Microenvironment composition of the groups of Liu et al.

Few KEGG pathways resulted enriched in low-risk group, and the pathways were associated with metabolism, cell cycle, DNA repair and immune response (Table 23).

Group enriched	Geneset KEGG	p-value	FDR	Common with "Immune Exhausted" (Chen et al.)?	Common with "Immune Active" (Chen et al.)?
Low group (Liu et al.)	GLYOXYLATE_AND _DICARBOXYLATE_METABOLISM	0.00000166	0.000204	No	No

Low group (Liu et al.)	RIBOSOME	0.00000209	0.000255	No	No
Low group (Liu et al.)	N_GLYCAN_BIOSYNTHESIS	0.0000116	0.0014	No	Yes
Low group (Liu et al.)	NUCLEOTIDE_EXCISION_REPAIR	0.0000125	0.0015	No	No
Low group (Liu et al.)	TGF_BETA_SIGNALING_PATHWAY	0.0000184	0.0022	No	No
Low group (Liu et al.)	CELL_ADHESION_MOLECULES_CAMS	0.0000321	0.00379	No	No
Low group (Liu et al.)	MISMATCH_REPAIR	0.0000373	0.00436	No	No

Table 23. KEGG pathways de-regulated in high and low risk groups of Liu et al.

Despite the correlation of survival (either PFS or OS) and the 8 gene expression signatures, we did not observe any significant association between response or disease control rate and all the signatures, with the only exception of the signature of Hu et al. that resulted significantly associated with DCR (p=0.012, Wilcoxon test).

4.6.3 Head and neck cancer subtypes (De Cecco, Oncotarget 2015)

To understand the molecular HNSCC heterogeneity, in 2015 De Cecco et al. established a large meta-analysis of a multitude of publicly available gene expression datasets (used as training and testing set). By a consensus unsupervised clustering proposed six different and well-defined HNSCC subtypes. The subtypes were strongly investigated by the biological point of view, and they summarized the aberrant alterations occurring during HNSCC progression (in terms of de-regulated pathways and tumor microenvironment composition). Clusters were named based on their biological features, and specifically were defined as: Cl-1 “HPV-like” (features related to HPV infection; highly proliferative; enriched by patients with oropharyngeal cancer); Cl-2 “Mesenchymal” (enriched in EMT, angiogenesis WNT NOTCH signaling, tumor growth

factor β (TGFβ) and rat sarcoma (RAS) related pathways); Cl-3 “Hypoxia” (enrichment in hypoxia features, and drug metabolism-related pathways, TGFβ pathways); Cl-4 “defense-response” (enriched in interferon and immune response)”; Cl-5 “Classical/Smoking” (characterized by up-regulation of genes related to smoking and xenobiotic metabolism; characterized by a large number of smoker patients); Cl-6 “Immunoreactive” (up-regulation of all the immune system related pathways and cellular homeostasis). The robustness of this classification was based on a precise analysis that correlated the 6 subtypes to the normal state, and what was observed was a continuous progression from subtypes closer to the normal state (Cl-6 and Cl-4) and subtypes more distant from the normal state (Cl3 and Cl2). However, authors declared that the subtypes, biologically well-defined, should be corroborated by further analyses to understand the relationship with clinical outcomes.

Testing the six HNSCC subtypes, 3 resulted associated with survival our gene expression cohort: Cluster-4 “Defense response” (43 low group; 37 high group, Figure 28A), Cluster-5 “Classical” (12 low group; 68 high group, Figure 28B) and Cluster-6 “Immunoreactive” (65 low group; 15 high group, Figure 29). The Cluster-4 “Defense response” subtype showed a significance in PFS analysis (p=0.036, Log-rank test), but not in OS (p=0.14, Log-rank test).

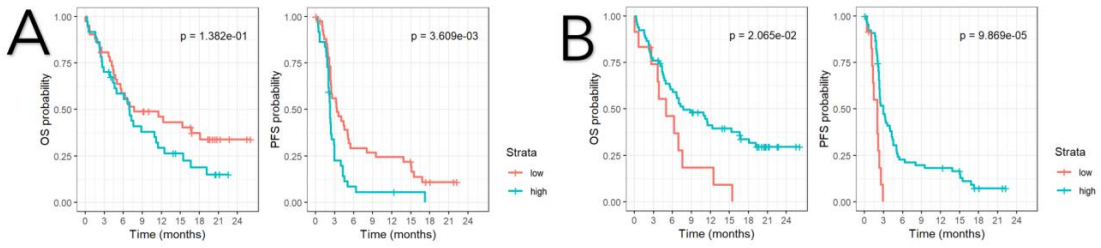


Figure 28 Kaplan-Meier of PFS and OS showing the prognostic role of 2/3 de-regulated De Cecco HNSCC subtypes; A. Cluster-4 “Defense response” B. Cluster-5 “Classical”

Meanwhile, Cluster-5 “Classical” subtype and Cluster-6 “Immunoreactive” subtype were the only two signatures (including both literature and 6 HNSCC subtypes signatures) that resulted significantly associated with both OS and PFS (Figure 28).

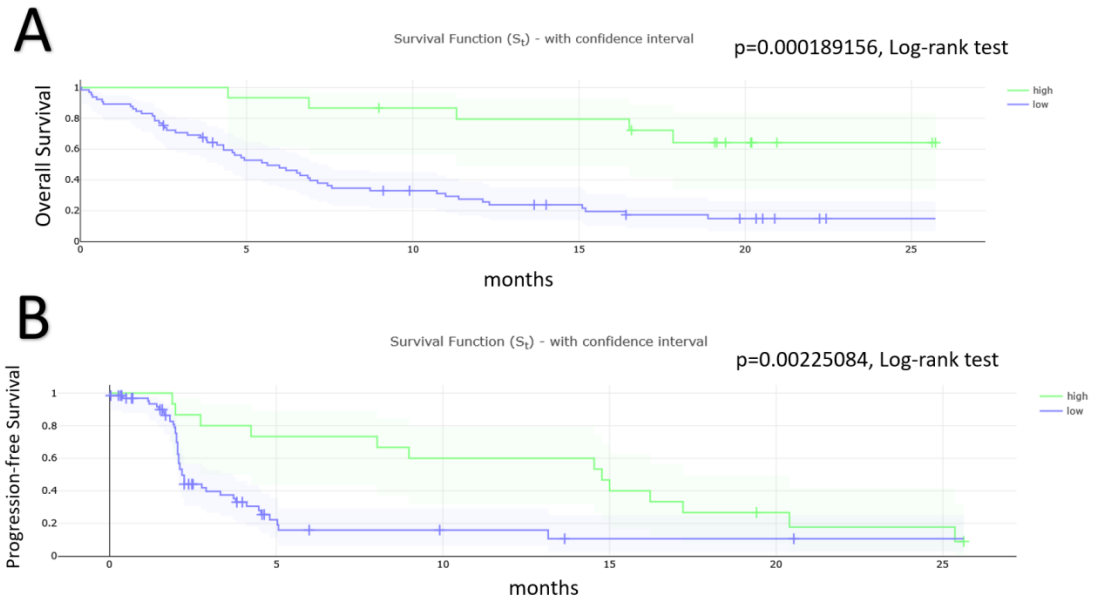


Figure 29 Kaplan-Meier of PFS and OS showing the prognostic role of Cluster-6 “Immunoreactive” HNSCC subtype

Investigating the correlation between the two signatures and response, a strong correlation between DCR and Cluster-5 (p=0.000043, Wilcoxon test, Figure 30 A) was found, meanwhile we did not observe a significant correlation between response and Cluster-5 (p=0.077, Wilcoxon test, Figure 30 B). Nevertheless, we seen a strong correlation between disease control rate and response with Cluster-6 (Figure 29 C and Figure 30 D).

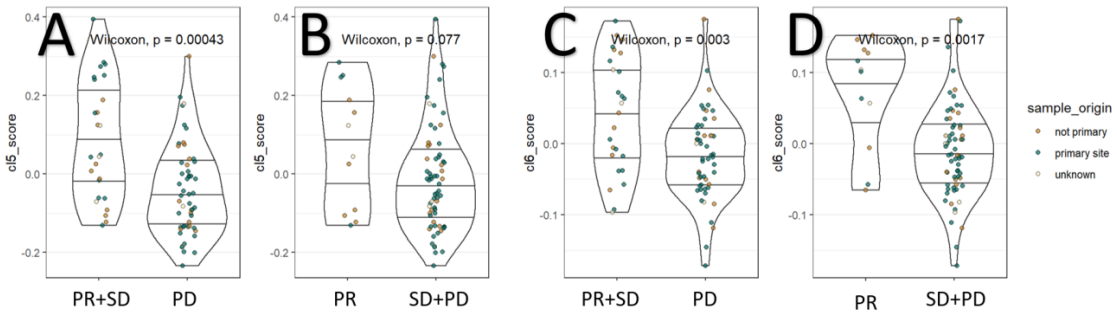


Figure 30 Violin plot showing the association of RECIST and De Cecco signatures scores in continuum. A) Association between Cl-5 score and disease control rate (PR+SD vs PD); B) Association between Cl-5 score and response (PR vs SD+PD); C) Association between Cl-6 score and disease control rate (PR+SD vs PD); D) Association between Cl-5 score and response (PR vs SD+PD). No association between the source of tumor tissue (primary, recurrence, metastasis, unknown and signatures score was observed).

4.6.4 Signature correlations

We investigated the potentially correlation between the signatures (we selected only those significantly associated with survival: 8 literature signatures and 3 HNSCC subtypes) and clinical-pathological parameters, such as age, gender, smoking status, and tumor subsite. We did not observe a correlation between the signatures and age (in continuum or using the cut-off of 65 years old), gender, tumor subsite and smoking status. Biological parameters, such as PD-L1, using in continuum CPS and TPS score, associated with two signatures: *Rooney et al* with TPS and CPS score used in continuum, while *Ayers et al* with TPS and CPS using the cut-off of 1. Moreover, we investigated the correlation within the all the 15 signatures (including 12 literature signatures and 3 HNSCC subtypes, Figure 31). Among the 8 significant prognostic signatures from literature and the three HNSCC subtypes, we observed: a direct correlation between the signatures of *Rooney et al* and *Ayers et al* (IFN-gamma); an inverse correlation between the signatures of *Bai et al* and *Liu et al* (immune) and *Bai et al* and *Hu et al*; a direct correlation between *Bai et al* and *Liu et al* (myeloid) and *Bai et al* and *Li et al;* an inverse correlation between the signatures of *Hu et al* and Cluster-5 and Cluster-6 subtypes. Notably,

among signatures significantly associated with survival one of the strongest correlations was observed between Cluster-6 and Cluster-5, that were directly correlated (Figure 30, p=2.7e-13, Pearson's chi-squared test)

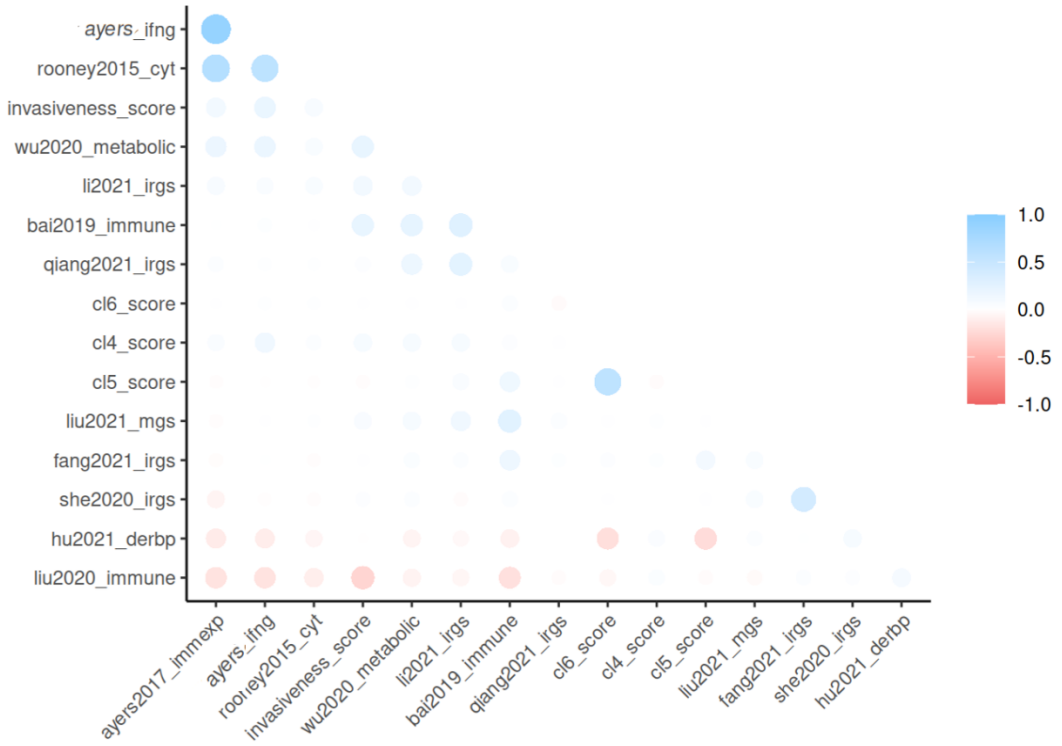


Figure 31 Correlation between gene expression signatures

4.6.5 Biological characteristics of Cluster-5 and Cluster-6 subtypes

We focus our attention on the two signatures that well-performed for both OS and PFS analyses, and that well-correlated together: Cluster-5 "Classical" and Cluster-6 "Immunoreactive". Biological investigation of de-regulated pathways (Hallmark GSEA, KEGG pathways) and microenvironment composition (xCell and TIMER) was performed. We observed a strong de-regulation of several biological Hallmark GSEA pathways and for patients in the group of high expression of Cluster-5 (Cl-5 High, Table 24). Specifically, the patients with the high score of Cluster-5 (better prognosis) were confirmed to be related to xenobiotic metabolism

(and consequently to reactive oxygen species) and WNT beta catenin signaling, as described by *De Cecco et al* (Table 24). Moreover, they expressed new features, such as an up-regulation in pathways related to proliferation (i.e., G2M checkpoint, P53 pathway and E2F targets), and up-regulation of metabolic pathways (i.e., oxidative phosphorylation).

Group enriched	Hallmark GSEA pathway	p-value	FDR	Process category
High Cl-5	REACTIVE_OXYGEN_SPECIES_PATHWAY	0.00000117	0.0000585	pathway
High Cl-5	G2M_CHECKPOINT	0.00000133	0.0000652	proliferation
High Cl-5	E2F_TARGETS	0.00000138	0.0000663	proliferation
High Cl-5	MTORC1_SIGNALING	0.00000343	0.000161	signaling
High Cl-5	MYC_TARGETS_V1	0.00000542	0.000249	proliferation
High Cl-5	MYC_TARGETS_V2	0.00000584	0.000263	proliferation
High Cl-5	CHOLESTEROL_HOMEOSTASIS	0.00000633	0.000279	metabolic
High Cl-5	DNA_REPAIR	0.0000505	0.00217	DNA damage
High Cl-5	UNFOLDED_PROTEIN_RESPONSE	0.0000516	0.00217	pathway
High Cl-5	OXIDATIVE_PHOSPHORYLATION	0.0000575	0.00236	metabolic
High Cl-5	GLYCOLYSIS	0.000067	0.00268	metabolic
High Cl-5	MITOTIC_SPINDLE	0.0000771	0.00301	proliferation
High Cl-5	WNT_BETA_CATENIN_SIGNALING	0.000162	0.00616	signaling
High Cl-5	PEROXISOME	0.000208	0.0077	Cellular component
High Cl-5	ESTROGEN_RESPONSE_LATE	0.000299	0.0108	signaling
High Cl-5	NOTCH_SIGNALING	0.000303	0.0108	signaling
High Cl-5	XENOBIOTIC_METABOLISM	0.000314	0.0108	metabolic
High Cl-5	FATTY_ACID_METABOLISM	0.000347	0.0114	metabolic
High Cl-5	PROTEIN_SECRETION	0.000398	0.0127	pathway
High Cl-5	ANDROGEN_RESPONSE	0.000402	0.0127	signaling
High Cl-5	ADIPOGENESIS	0.00042	0.0127	development
High Cl-5	HEDGEHOG_SIGNALING	0.000498	0.0144	signaling
High Cl-5	P53_PATHWAY	0.000544	0.0152	proliferation
High Cl-5	PI3K_AKT_MTOR_SIGNALING	0.000584	0.0158	signaling
High Cl-5	UV_RESPONSE_UP	0.000632	0.0164	DNA damage
High Cl-5	HEME_METABOLISM	0.000974	0.0244	metabolic

Table 24. Enriched Hallmark GSEA pathways (Cluster-5)

Summarizing, the Cluster-5 high group (better prognosis) resulted enriched in 1/6 development pathways, 2/3 DNA damage, 0/7 immune-related, 6/7 metabolic pathways, 6/6 proliferation-related pathways and 11/13 signalling-related pathways. While patients with low-expression of Cluster-5 (worse prognosis) exhibited an enriched immune

microenvironment, evidenced by Microenvironment Score (which is the sum score of ImmuneScore+StromaScore) and Immune Score. Specific lymphoid populations were enriched, such as NKT and CD4+ Tem (Table 25).

Group enriched	Cell type (xCell)	p-value	FDR	XCELL CATEGORY
Cl-5 low	MicroenvironmentScore	0.000000392	0.0000262	NA
Cl-5 low	ImmuneScore	0.00000755	0.000498	NA
High Cl-5	Smooth muscle	0.0000283	0.00184	Stromal cells
Cl-5 low	Eosinophils	0.0000775	0.00496	Myeloids
Cl-5 low	NKT	0.000136	0.00857	Lymphoids
Cl-5 low	CD4+ Tem	0.000328	0.0203	Lymphoids
Cl-5 low	Adipocytes	0.000354	0.0216	Stromal cells
Cl-5 low	Platelets	0.000425	0.0255	Stem cells
Cl-5 low	Neurons	0.000598	0.0353	Others
Cl-5 low	ly Endothelial cells	0.000636	0.0369	Stromal cells
High Cl-5	CLP	0.000671	0.0383	Stem cells
High Cl-5	Tgd cells	0.000689	0.0386	Lymphoids

Table 25. Microenviroment composition of the groups (Cluster-5)

KEGG pathways were de-regulated only in High group Cl-5, and showed similar results compared to GSEA Hallmarks: a metabolic reprogramming and a high proliferative activity (Table 26).

Group enriched	Geneset KEGG	p-value	FDR
High Cl-5	GLUTATHIONE_METABOLISM	1.77E-08	0.00000217
High Cl-5	STEROID_BIOSYNTHESIS	0.00000133	0.000162
High Cl-5	GLYCOSYLPHOSPHATIDYLINOSITOL _GPI_ANCHOR_BIOSYNTHESIS	0.00000381	0.000461
High Cl-5	CELL_CYCLE	0.00000998	0.0012
High Cl-5	BASE_EXCISION_REPAIR	0.0000228	0.00271
High Cl-5	GLYOXYLATE_AND_DICARBOXYLATE_METABOLISM	0.0000281	0.00332
High Cl-5	DNA_REPLICATION	0.0000307	0.00359
High Cl-5	MISMATCH_REPAIR	0.0000346	0.00401
High Cl-5	SPLICEOSOME	0.0000385	0.00442
High Cl-5	BASAL_CELL_CARCINOMA	0.0000456	0.0052
High Cl-5	NUCLEOTIDE_EXCISION_REPAIR	0.0000552	0.00623
High Cl-5	RNA_POLYMERASE	0.0000611	0.00684
High Cl-5	GLYCEROPHOSPHOLIPID_METABOLISM	0.0000759	0.00842
High Cl-5	BASAL_TRANSCRIPTION_FACTORS	0.0000829	0.00912
High Cl-5	OXIDATIVE_PHOSPHORYLATION	0.0000847	0.00924
High Cl-5	PYRIMIDINE_METABOLISM	0.0000865	0.00934

High Cl-5	UBIQUITIN_MEDIATED_PROTEOLYSIS	0.000119	0.0128
High Cl-5	NOTCH_SIGNALING_PATHWAY	0.000125	0.0132
High Cl-5	BIOSYNTHESIS_OF_UNSATURATED_FATTY_ACIDS	0.000144	0.0151
High Cl-5	LYSINE_DEGRADATION	0.000165	0.0171
High Cl-5	RNA_DEGRADATION	0.000193	0.0199
High Cl-5	PROTEIN_EXPORT	0.00022	0.0225
High Cl-5	VALINE_LEUCINE_AND_ISOLEUCINE_DEGRADATION	0.000232	0.0234
High Cl-5	RIBOSOME	0.000244	0.0244
High Cl-5	WNT_SIGNALING_PATHWAY	0.000315	0.0312
High Cl-5	FRUCTOSE_AND_MANNOSE_METABOLISM	0.000332	0.0326

Table 26. KEGG pathways de-regulated in the high Cluster-5

Meanwhile, few differences were observed comparing the two expression groups of Cl-6. In detail, the Cluster-6 high (better prognosis) was enriched in Hallmark GSEA "Cholesterol_Homeostasis" (p=0.000441; FDR=0.0221) and KEGG pathway "Glutathione_metabolism" (p=0.0000509; FDR=0.00626). Instead, the microenvironment composition characterized the Cluster-6 low group (patients with the worse prognosis), that was significantly enriched in NKT (p=0.0000108, FDR= 0.000727, Figure 32) and Myocytes (p=0.000463, FDR= 0.0306).

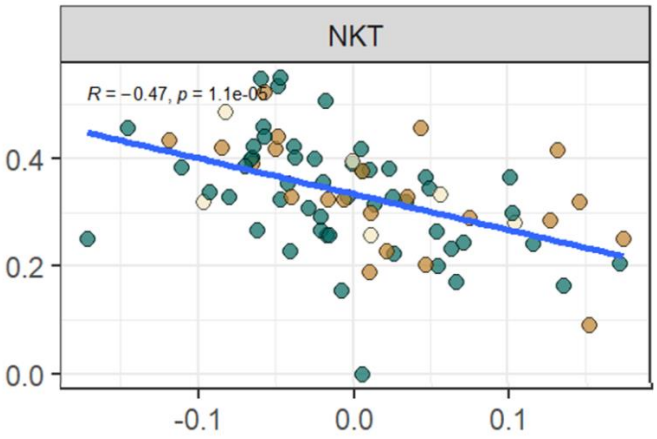


Figure 32 Scatterplot of NKT score (y-axis) correlated with Cluster-6 score (x-axis). P-value was 0.0000108 and FDR= 0.000727

4.6.6 Immuno-related genes

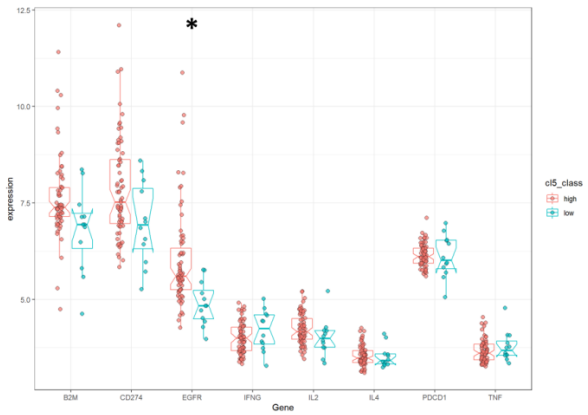
To discern the immune factors that could had influenced the relation with response to therapy/survival and Cluster-5 or Cluster-6, we investigated the differences in expression of specific immune-related genes, that indicate an immune activation in TME (i.e., CD274 also known as PD-L1; Interleukin-2 (IL-2); Interleukin-4 (IL-4)) or a negative regulation of immune populations (i.e., Beta-2-Microglobulin (B2M); Epidermal growth Factor Receptor (EGFR); Programmed cell death protein 1 (PDCD1); Tumor Necrosis Factor alpha (TNF). Their specific roles are indicated in Table 27.

Gene	Immune-related function
Beta-2-microglobulin (B2M)	It encodes a protein found in association with the major histocompatibility complex (MHC) class I. Evidence showed that alterations of B2M gene and B2M proteins contribute to poor reaction to cancer immunotherapies [143].
CD274 (commonly referred to as PD-L1)	It is a ligand that binds with the receptor PD1, commonly found on T-cells, and acts to block T-cell activation. Its expression has been observed in a variety of cancers. High levels of PD-L1 have been associated to an active immune system and a better response to immunotherapy.
Epidermal Growth Factor Receptor (EGFR)	It has been observed that tumours overexpressing EGFR grow autonomously and become "addicted" to growth factor signalling. Moreover, it has been seen that the overexpression, specifically of mutant forms of the EGFR, may create an immune-suppressive and lymphocyte depleted microenvironment within tumour. Such a microenvironment may explain the resistance of EGFR overexpressing cancers to tumour therapies, particularly to check-point inhibitor treatment [144]
Interleukin 2 (IL-2)	IL-2 has an immunoregulatory role, promoting the growth and development of peripheral immune cells in the initiation of the (defensive) immune response, and keeping them alive as effector cells. It has been showed to be correlated to durable response in patients treated with immunotherapy[145].

Interleukin 4 (IL-4)	In addition, the effect of IL-4 in stimulating both transcription factors, increased by T cell activation, is also involved with its ability to promote the cytotoxic activity of $CD8^+$ T cells[146]
Programmed cell death protein 1 (PDCD1), also known as PD-1	It is an immune-inhibitory receptor expressed in activated T-cells, B-cells and macrophages; it is involved in the negative regulation of their functions, including those of effector CD8+ T cells. High levels of PD-1 could mediate resistance to immunotherapy[147].
Tumor Necrosis Factor alpha (TNF)	This gene encodes a multifunctional proinflammatory cytokine, which is mainly secreted by macrophages. Tumor necrosis factor alpha (TNF)-dependent modulation of immune responses and cell death processes has long been the subject of intense research. It is paradoxically involved in both pro-inflammatory mechanism and anti-immunomodulatory effects. For this reason, its role in cancer progression and its influence in the response to immunotherapy is still a matter of debate[148,149]

Table 27. Immune-related genes and their function

Observing the expression of these specific genes, comparing the high and low groups of Cluster-5 we did not detect any significant differences in terms of expression for all the genes, except for EGFR (p=0.000436; FDR=0.00349, Figure 33), which expression resulted higher in the Cl-5 high (better prognosis).

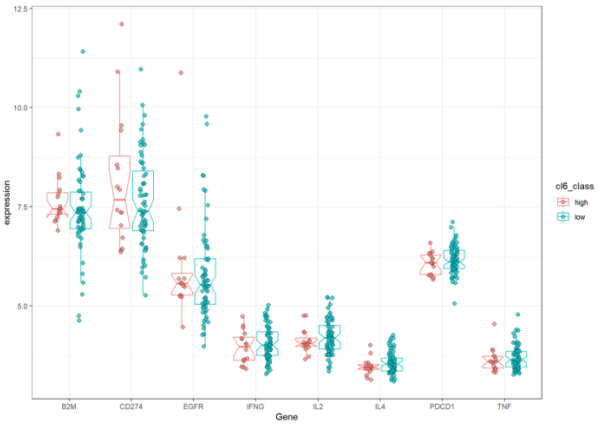


Gene	p-value	FDR
B2M	0.0134	0.0939
CD274	0.0372	0.223
EGFR	0.000436	0.00349
IFNG	0.214	0.855
IL-2	0.0769	0.384
IL-4	0.440	1
PDCD1	0.816	1
TNF	0.349	1

Figure 33 immune-related genes in Cluster-5 groups

While by the comparison of the Cluster-6 groups we did not obtain any significant result, and the genes resulted equally expressed in both high and low groups (Figure 34), indicating that some other factors were

contributing to the functionality of the relation between survival of Cluster-6.



Gene	p-value	FDR
B2M	0.186	1
CD274	0.403	1
EGFR	0.639	1
IFNG	0.390	1
IL-2	0.407	1
IL-4	0.251	1
PDCD1	0.428	1
TNF	0.354	1

Figure 34 Expression of the immune-related genes in Cluster-6 groups

4.7 DNA sequencing experiments

DNA was extracted and nucleic acids were checked. In detail, 7 and 13 samples were excluded for quality and quantity. Performing library construction 7 libraries resulted inadequate for quality or quantity. Moreover, after the alignment pipeline and the applying of specific filters, 4 libraries did not have any specific mutation and they were not considered. A total of 63 libraries were produced with annotated mutations, and of them 59 had informative gene expression (Figure 35). After the alignment of the 63 libraries, a total of 1260 variant calls were available for the analyses. The 63 DNA sequencing cohort was composed of 41 samples deriving from primary tumor, 17 samples deriving from recurrence/metastasis, and 5 without the specific definition. The 63 patients experienced different responses: 8 were partial responders (PR), 13 with stable disease (SD) and 41 with progressive disease (PD). For 1 patient the response was not annotated.

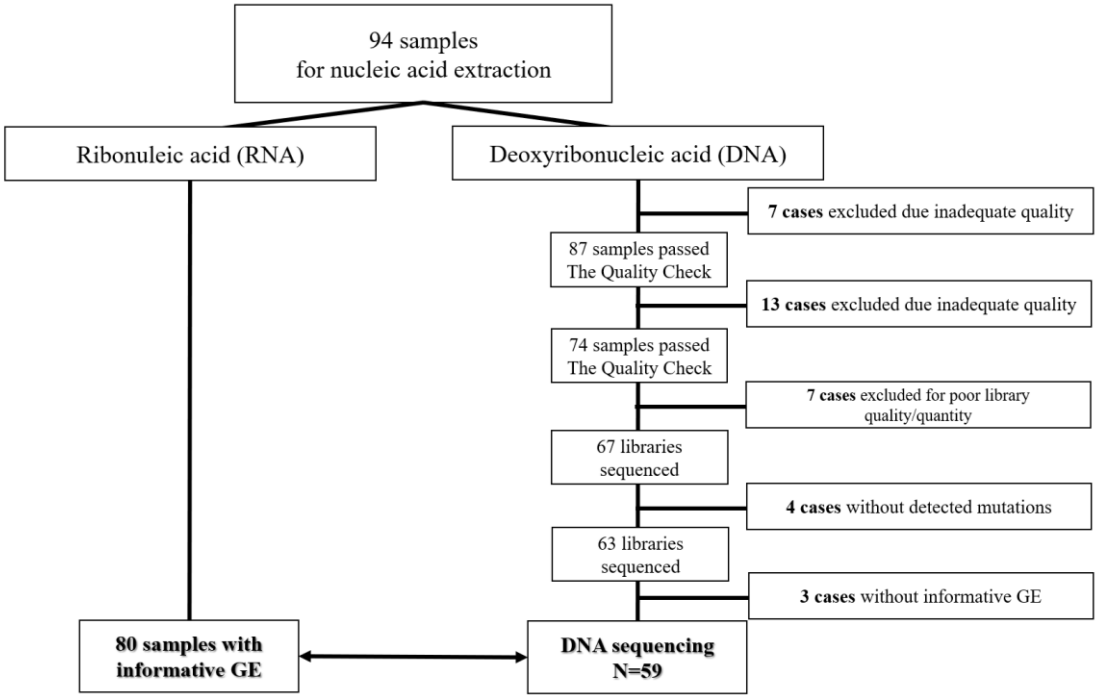


Figure 35 Consort diagram of DNA sequenced samples

Comparing the clinical-pathological characteristics of the two cohorts (i.e., gene expression dataset, DNA sequencing dataset) we did not observe any significant difference (Table 28). The median overall-survival seen for the gene expression dataset was 6.72 months [0.03-25.72], while was 7.43 months [0.26-25.62] for the mutation dataset. The median progression free survival was 2.20 months [0.01-25.62] for the gene expression database, and 2.23 months [0.01-25.62] for the mutation dataset.

Clinical-pathological characteristics		**GE dataset (N=80)**	**Mutation dataset (=63)**	**p value**
Age, years	**median [range]**	65.5 [33-84]	65 [33-84]	NA
Gender	**male**	65 (81%)	51 (81%)	1
	female	15 (19%)	12 (19%)	
Smoking status	**present**	16 (20%)	13 (21%)	.975582
	past	50 (62%)	40 (63%)	
	never	11 (14%)	8 (13%)	

	N.A.	3 (4%)	2 (3%)	
Site of primary disease	**Oral cavity**	30 (37%)	26 (41%)	.949895
	Oropharynx HPV positive	7 (9%)	6 (9%)	
	Oropharynx HPV negative	5 (6%)	3 (5%)	
	Oropharynx N.A.	7 (9%)	5 (8%)	
	Larynx	18 (22%)	13 (21%)	
	Hypopharynx	11 (14%)	8 (13%)	
	N.A.	2 (3%)	2 (3%)	
Performance status	**0**	27 (34%)	26 (41%)	.465314
	1	50 (62%)	36 (57%)	
	2	3 (4%)	1 (2%)	
Response	**CR**	0 (0%)	0 (0%)	.824204
	PR	12 (15%)	8 (12%)	
	SD	14 (18%)	13 (21%)	
	PD	53 (66%)	41 (65%)	
	NA	1 (1%)	1 (2%)	

Table 28. Clinical and pathological characteristic of the 63 patients with annotated DNA mutation. P-values were tested using Chi-square

4.8 Targeted sequencing

The totality of 1260 mutations, observed in the 63 samples, were mostly missense mutation (Figure 36 A). The most observed event was C>T transition, followed by T>C transition, and C>A transversion and C>G transversion (Figure 36 B). Every sample had a median of 20 mutations, with a range of 40 mutations in one sample, and 13 mutations for those samples with less mutations (Figure 36 C). The most 10 mutated genes were MDC1 (57% of the cohort, mutated in 36 samples and with 72 mutations), followed by PLCG2 (mutated in the 40% of samples) and LRP1B (mutated in the 37% of the cohort, Figure 36 D). The most observed event was transversion (Figure 36 E). Comparing responders vs non responders or patients with DCR vs patients without DCR, we were not able to observe any mutually exclusive mutation.

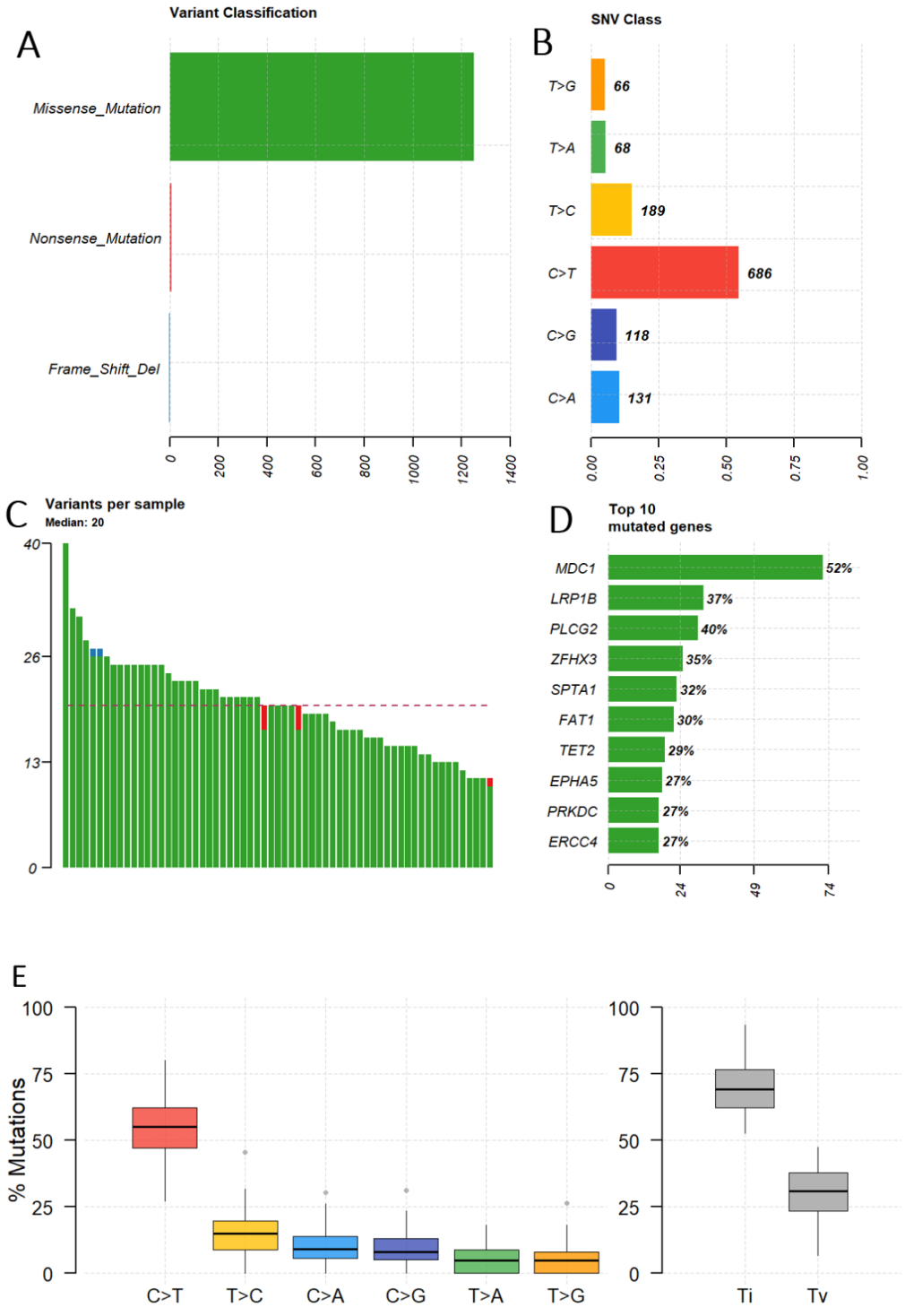


Figure 36 Most frequent characteristics of annotated somatic mutations A) 1260 variant were annotated and classified; B) Type of single nucleotide variant classification; C) number of variants for each sample; D) List of the 10 mutated genes; E) Boxplot showing the comprehensive % of transition (Ti) vs tranversion (Tv) events

The function of the 10 most mutated genes was investigated and summarized in the Table 29.

Gene (%)	Function	doi
MDC1, mutated in 52% of the 63/124 Nivactor patients	Mediator of DNA damage checkpoint 1 (MDC1) is a gene that encodes for a protein essential for the correct functionality of DNA damage response (DDR) system, which repairs DNA double-strand break. Consequently, MDC1 is vital for the maintain of genomic stability. It has been found mutated in HNSCC from 4-15% (COSMIC, TCGA, ICGD databases); However, mutations involving MDC1 gene have been observed elevated in OSCC and in R/M cancers. Associated to unfavourable prognosis is OSCC	https://doi.org/10.1016/j.dnarep.2022.103330 [150] https://doi.org/10.1111/eos.12662 [151] doi:10.1111/jop.12558 [152]
LRP1B, mutated in 37% of the 63/124 Nivactor patients	Low-density lipoprotein receptor-related protein 1B (LRP1B) is one of the most altered genes in human cancers. It has been found frequently inactivated by several genetic and epigenetic mechanisms, and it is involved in several biological processes (such as cell migration, tumorigenesis and tumor progression, and DNA damage response). LRP1B is also described as a common target gene for viral integration for human papillomavirus (HPV).	https://doi.org/10.3390/ph14090836 [153]
PLCG2, mutated in 40% of the 63/124 Nivactor patients	Phospholipase C, gamma 2 (phosphatidylinositol-specific) (PLCG2) is a gene that encodes a protein that functions as a transmembrane signaling enzyme. The dysfunction of the protein is associated with a variety of diseases including cancer, neurodegeneration, and immune disorders. Its involvement in immunotherapy response is under evaluation.	https://doi.org/10.1016/j.jbc.2021.100905 [154] https://doi.org/10.3390/curroncol28050347 [155] doi:10.1097/MD.0000000000025008 [156] https://doi.org/10.1016/j.ccell.2021.09.008 [157]

ZFHX3	Zinc finger homeobox 3 (ZFHX3) encodes for a large transcription factor, involved in proliferation, tumor growth and angiogenesis in different cancer types. Its role and involvement in immunotherapy response is studying in NSCLC.	doi: 10.3390/cancers12113415 [158] https://doi.org/10.1074/jbc.RA119.012131 [159] https://doi.org/10.1007/s00262-020-02668-8 [160]
SPTA1	Nonerythroid spectrin αII (SPTAN1) is involved in cell adhesion, cell-cell contact, and apoptosis. SPTAN1 was shown to interact with different proteins involved in DNA repair, chromatin remodeling, and fanconi anemiax.	doi:10.1155/2019/7079604 [161] doi:10.1177/1535370216662714 [162]
FAT1 mutated in 30% of the 63/124 Nivactor patients	FAT atypical cadherin 1 (FAT1) is one of the most frequently mutated genes in many type of cancers. Its highest mutation rate is found in HNSCC (29.8%). In many cancers mutations in FAT1 promote EMT, cancer progression and increase the functionality of immune infiltrating cells.	https://doi.org/10.1186/s13046-022-02461-8 [163]
TET2	Tet methylcytosine dioxygenase 2 (TET2) is frequently mutated in several solid and hematopoietic cancers. The protein Tet2 is essential for the development, differentiation, function of several lymphoid cells (i.e., B-cell, CD8+T-cell memory etc). Protein loss of function reduces anticancer immune response.	https://doi.org/10.1038/s42003-020-01391-5 [164]
EPHA5	EPH Receptor A5 (EPHA5) mutations have been annotated in different cancer types, suggesting an important role in tumorigenesis. Moreover, EphA5 protein have been associated (in lung cancer cells) to defects in G_1/S cell cycle checkpoint, making the cell unable to resolve DNA damage. The gene mutation has additionally been associated to immunosuppressive TME and worse survival for NSCLC.	https://doi.org/10.3892/ol.2019.10167 [165] https://doi.org/10.3389/fonc.2021.619949 [166] doi:10.1074/jbc.M114.630525 [167] https://doi.org/10.1016/j.lungcan.2020.11.006 [168]
PRKDC	The PRKDC gene encodes the DNA-dependent protein kinase catalytic subunit (DNA-PKcs) protein. The protein plays an important role in DNA repair, and it is strictly related to immune tolerance and the	https://doi.org/10.1186/s12935-021-02229-8 [169] doi:10.1177/0300891620950472 [170]

	maintenance of chromosome stability. The impairment of the protein is correlated with tumor initiation and progression.	
ERCC4	Excision repair cross-complementation group 4 (ERCC4) encodes for a protein involved in DNA repair. Mutations in this gene have been widely observed in cancer and they are associated with an increased risk of tumorigenesis.	doi: 10.3389/fonc.2022.951193 [171]

Table 29. Functions of the most 10 mutated genes

In literature, specific mutations have been associated to immunotherapy response or resistance in different cancer types (Table 30). We investigated their presence in the cohort of 63 patients, and if possible, their representation in the groups of responders vs non responders.

Gene	Function	Cancer type	doi
BRAF, KRAS, and TP53	Response	Melanoma, Bladder, Non-small Cell Lung Cancer	https://doi.org/10.1038/s41467-022-31055-3 [172]
EGFR	Resistance	Non-Small Cell Lung Cancer	https://doi.org/10.3389/fonc.2021.635007 [173]
Beta-2-microglobulin (B2M) and JAK1	Resistance	Melanoma	10.1056/NEJMoa1604958 [174]
Phosphatase and tensin homolog on chromosome 10 (PTEN)	Resistance	Glioblastoma	10.1038/s41591-019-0349-y [175]

Table 30. Mutations associated to immunotherapy resistance.
It was possible to consider only mutations in genes sequenced using the TSO500 panel (Illumina)

Notably, no point mutations in BRAF, KRAS, EGFR were annotated in our database, while no differences between responders vs non responders were observed in the frequency of mutations in specific genes, such as TP53, B2M and PTEN (Figure 37).

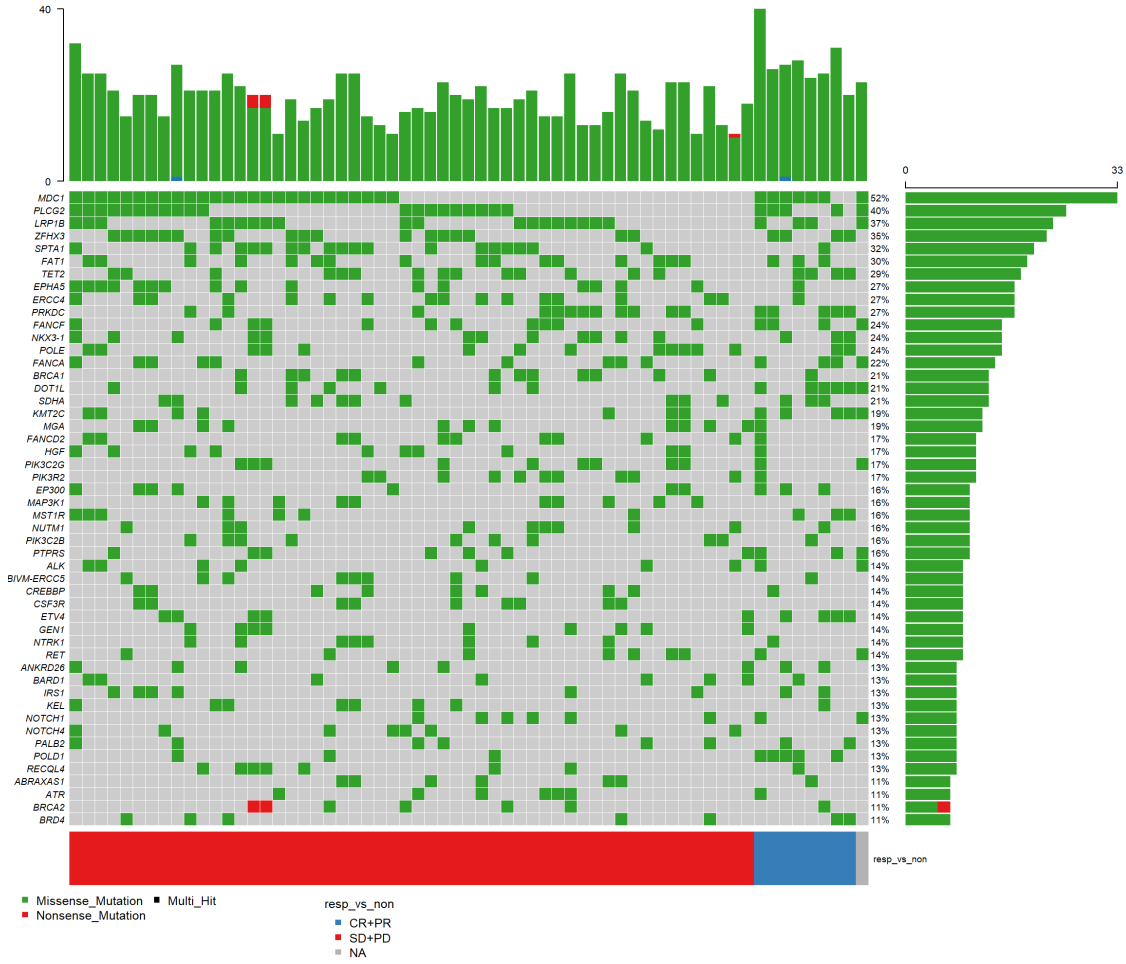


Figure 37 Waterfall plot comparing the most frequent mutations between Partial Responders (blue bar) vs Stable disease and Progressive disease (red bar)

Moreover, the integrative analysis of mutational status and the two De Cecco subtypes, revealed no differences in the top 50 mutations between high and low groups of Cluster-5 and Cluster-6. Indeed, we did not observe a specific pattern of mutations for each of the categories (Figure 38).

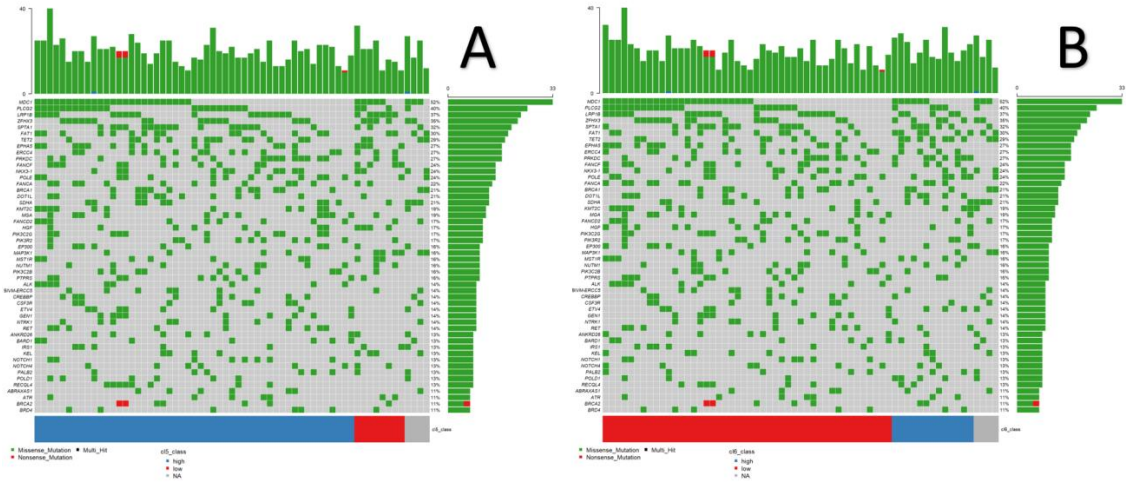


Figure 38 Waterfall plots of the most frequent somatic mutations in (A) Cluster-5 high (blue bar) vs low (red bar) and (B) Cluster-6 high (blue bar) vs low (red bar)

4.8.1 TMB and MSI

The evaluation of the TMB was possible in 63 samples. The median TMB was 5 [0-29.16]. Using TMB=10 as the cut-off, 56 samples resulted "TMB low" and 7 samples "TMB high". For 48 samples it was possible to assess the MSI, divided in 46 samples "MSI stable" and 2 "MSI high", while for 15 samples not enough microsatellites loci were sequenced to allow the evaluation of instability. For all the 48 samples with MSI score also TMB was available. However, we observed no correlation between the MSI and TMB scores (Figure 39). Four samples (with MSI available) resulted TMB high (≥ 10) and of them, two samples were classified also as "MSI high". The two patients with both TMB and MSI classified as "high" were both male, past smokers and the primary disease was oral cavity cancer. One of the two patients was 66 years old, with stable disease (SD) assessed as response to ICI, CPS=20 and TPS=0.5. The FFPE tissue profiled for this patient was the primary tumor. The other patient was 64 years old, with a progressive disease (PD) after immunotherapy, and CPS=100 and TPS=80. The available FFPE of the tumor for this patient was the tissue of the metastasis.

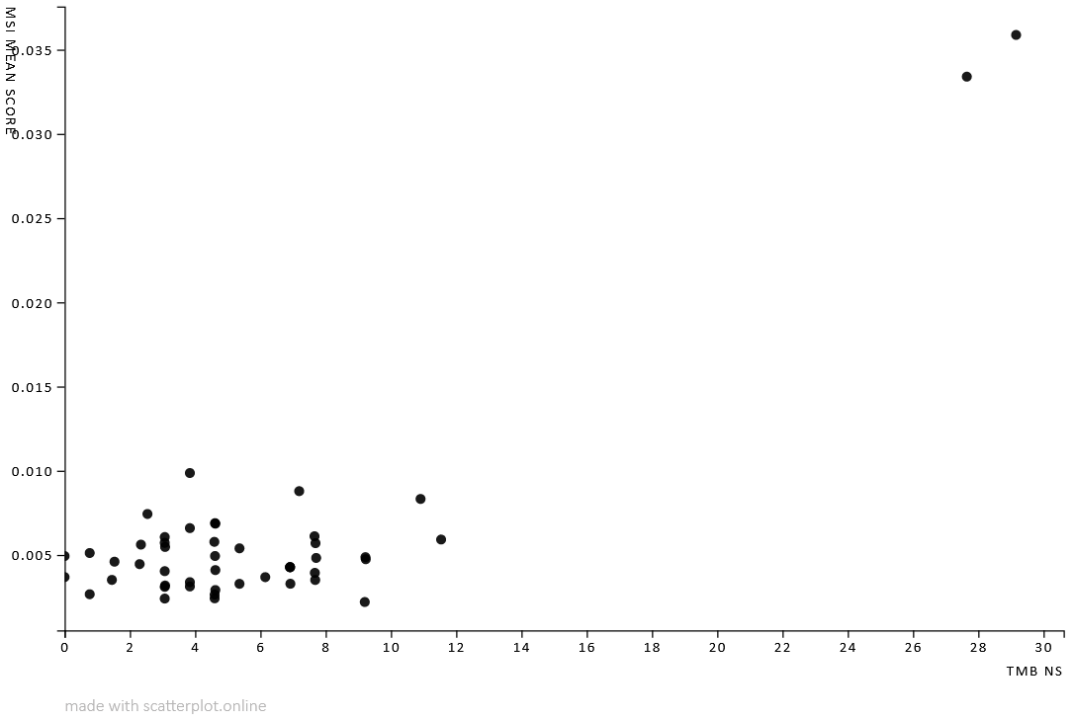


Figure 39 Correlation of MSI and TMB scores; MSI is expressed as mean score and TMB comprised only the non-synonymous mutations (NS); in the image are shown 48 samples with both values available

The survival analyses revealed no differences between patients with TMB < 10 or patients with TMB ≥ 10 (Figure 40).

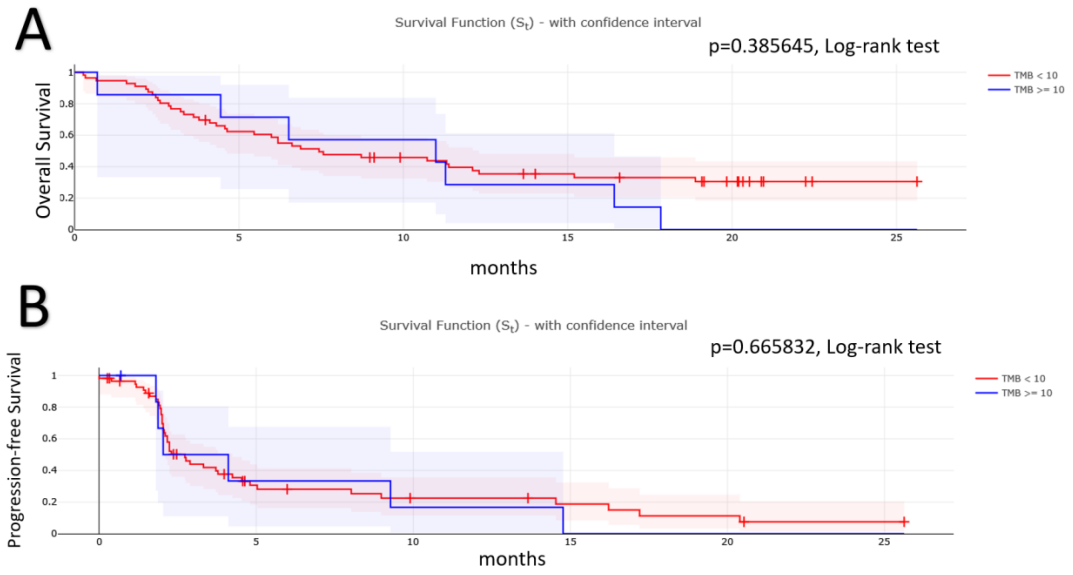


Figure 40 Kaplan Meier curves considering TMB ≥ 10 (N=63). A) OS analysis of patients with TMB < 10 (N=66, red curve) vs patients with TMB ≥ 10 (N=7, blue curve); B) PFS analysis of patients with TMB < 10 (N=66, red curve) vs patients with TMB ≥ 10 (N=7, blue curve). P-values were calculated using Log-rank test, significance was set at $p \leq 0.05$.

Due to the scarcity of patients in the group of "MSI-high" survival analyses were not investigated. Furthermore, we explored a possible correlation between TMB and MSI score, using the scores in continuum. However, we did not observe a correlation between the two scores and response to therapy (Figure 41).

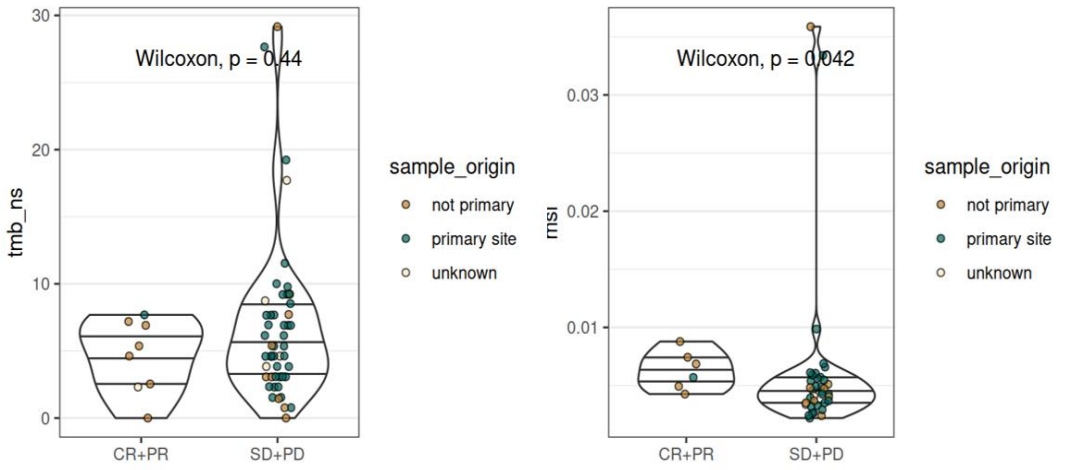


Figure 41 Violin plots of the correlation between TMB and MSI scores in continuum with response (classified in responders CR+PR vs non-responders SD+PD)

Moreover, we did not observe a correlation between TMB and MSI score and disease control rate (Figure 41).

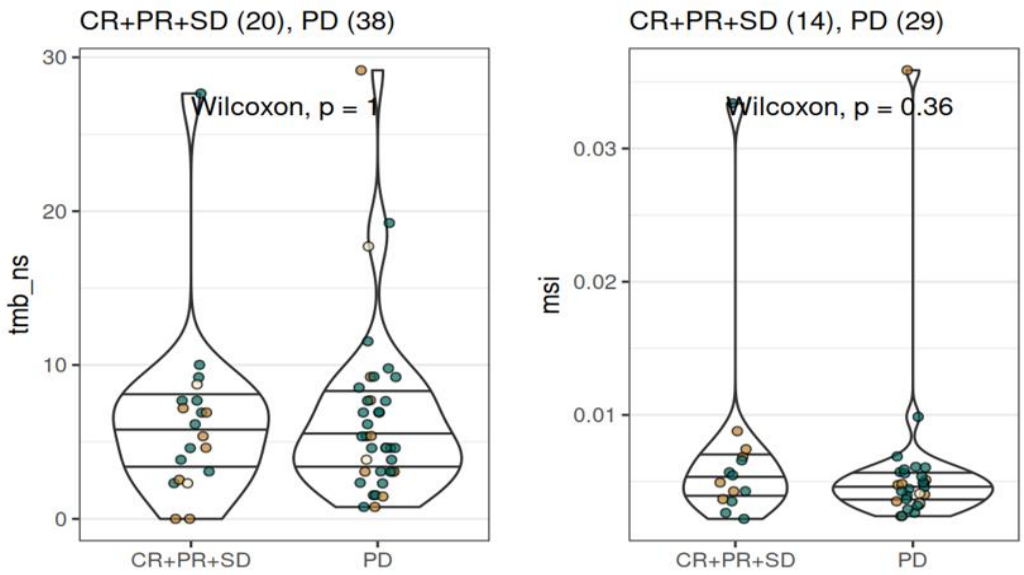


Figure 42 Violin plots investigating the correlation between TMB and MSI in continuum with disease control rate categories

Investigating the correlation between the 8 significantly observed signatures and 3 HNSCC subtypes with TMB and MSI, we observed the correlation between the signature of Bai et al. and TMB (score used in continuum, p=0.037).

4.8.2 Mutations in repair systems' related genes

In the cohort of 63 patients with annotated DNA mutations, we investigated the mutations present in genes related to repair systems in the two categories of responders (PR, N=8) and non-responders (SD+PD, N=53), while 2 patients were excluded because of lack of information about response status. The mutations investigated comprised several repair systems, such as Base excision repair system, DNA checkpoint, Homologous recombination, Fanconi anemia, Mismatch repair system, and nucleotide excision repair system (Table 31).

Type of repair system	Total number of mutations	PR (N=8)	Frequency In responders	SD+PD (N=53)	Frequency In non-responders	p-value
Base excision repair system (total number of mutations)	**7 hits in 7 pts**	**2 hits in 2 pts**	**2/8 (25%)**	**5 hits in 5 pts**	**5/53 (9%)**	**.000192**
MUTYH	6	2		4		
PARP1	1	0		1		
DNA checkpoint (total number of mutations)	**57 hits in 48 pts**	**9 hits in 8 pts**	**8/8 (100%)**	**48 hits in 40 pts**	**40/53 (75%)**	
MDC1	33	7		26		
BRCA1	13	1		12		
ATR	7	1		6		
ATM	4	0		4		
Homologous recombination (total number of mutations)	**16 hits in 16 pts**	**4 hits in 4 pts**	**4/8 (50%)**	**12 hits in 12 pts**	**12/53 (23%)**	
RAD51B	7	3		4		
MRE11	3	1		2		
RAD52	3	0		3		
RAD54L	2	0		2		
RAD50	1	0		1		
Fanconi Anemia (total number of mutations)	**45 hits in 38 pts**	**10 hits in 8 pts**	**8/8 (100%)**	**35 hits in 30 pts**	**30/53 (57%)**	
FANCA	14	4		10		
PALB2	8	2		6		
BRCA2	7	1		6		
SLX4	6	1		5		
RAD51C	4	1		3		
BRIP1	3	0		3		
FANCC	3	1		2		
Mismatch repair system (total number of mutations)	**18 hits in 15 pts**	**6 hits in 3 pts**	**3/8 (38%)**	**12 hits in 12 pts**	**12/53 (23%)**	
MSH2	7	3		4		
PMS1	4	2		2		
MSH5	3	0		3		
MLH1	3	0		3		
MSH3	1	1		0		
Nucleotide excision repair system (total number of mutations)	**17 hit in 17 pts**	**1 hit in 1 pt**	**1/8 (13%)**	**16 hit in 16 pts**	**16/53 (30%)**	
ERCC4	17	1		16		

Table 31. Specific DNA mutation annotated for each of the genes considered. patients were divided in base of response; NB: the same patient could have more than one gene mutated in the same repair system. P-value was calculated comparing the two groups (percentage of patients having mutations in all the repair systems) with Chi-Square test. The significance was set at $p < .05$.

Analyzing the same result through a graphical and more intuitive representation, it could be observed a specific and significant trend for which patients that experienced partial response had more mutations in Base excision, DNA checkpoint, Homologous recombination, Fanconi anemia and mismatch repair systems (all together, Figure 43).

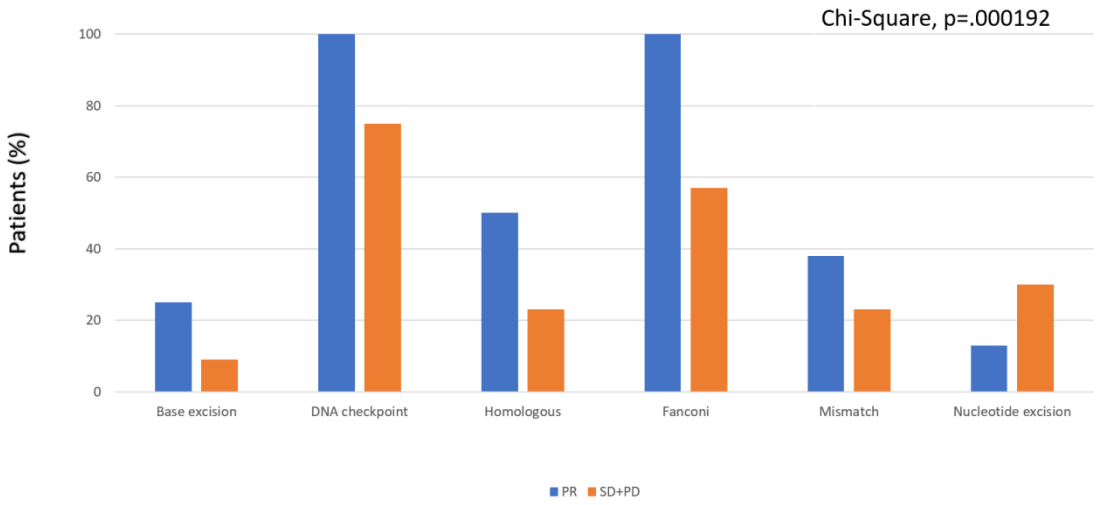


Figure 43 Bar plots showing the percentage of somatic mutations annotated for each repair system

4.9 Case report – The peripheral blood profile of a complete responder

Noteworthy, one of the two patients who obtained a complete response in Nivactor trial presented an interesting and peculiar clinical history, and for this reason we published the case report on Frontiers in Oncology in January 2022[176]. The patient (HPV-positive OPSCC), after a first-line treatment and a complete remission, experienced a Human Polyomavirus (JCV) infection in the brain, experiencing a progressive multifocal leukoencephalopathy, that was resolved in few months. However, blood test revealed a boost in lymphoid components during and at the resolution of JCV infection (even if cells count remained largely under the normal threshold). Subsequently, the patient metastasized, and he was enrolled in the clinical trial Nivactor. During the immunotherapy treatment, brain MRI evidenced the presence of small punctuate areas of contrast enhancement, reflecting a mild immune response in perivascular spaces. Nivolumab treatment was withdrawn after 10 infusions, G2 diarrhea and a syndrome of inappropriate antidiuretic hormone secretion (probably drug-related events). However, a complete and durable response (more 3 years) was observed. No tumor tissue could be retrieved for gene expression analysis. Recognizing the importance of profiling the immunological characteristics of this patient, blood samples collected before, during and after the single agent nivolumab treatment were employed to investigate, through a de-convolution gene expression method (xCell), the immune cells populations present in the peripheral blood. Lymphoid cells underwent through evident changes during and after nivolumab. Before the treatment, lymphoid cells (with exclusion of NKT and CD8+ Tem), exhibited a lower or absent expression compared to controls, while during the treatment a decrease for CD8+ Tem and NKT was recorded, and an

increase for Memory B-cells CD4+ T cells, CD4+ memory T cells, CD8+ T cells were seen. Interestingly, the immunological boost observed after JCV infection, and moreover, during the nivolumab treatment was maintained during time, suggesting the possible role of viral infection (JCV and HPV together, or alone) in the achievement of the complete response obtained by this specific patient.

5. DISCUSSION

The concept behind the immunotherapy is based on the belief that the immune system alone possesses the potential ability to fight cancer. In many tumor types it has been observed that therapies based on immune checkpoint inhibitors could result in durable tumor regression. Indeed, immunotherapy has transformed the treatment for R/M HNSCC patients, and for some patients had led to long-lasting response, never observed with other treatments. However, its anti-cancer ability was observed in only a small subset of patients and predictive biomarkers are still needed for a more precise and personalized therapy. The molecular determinants driving the response remained still a critical open question. To address this issue, we performed a throughout analysis, including well-annotated clinical-pathological data and omic profiling of tumor tissues, investigating a cohort of R/M HNSCC patients prospectically enrolled in a phase IIIb trial, and treated with nivolumab.

Comparing responding and non-responding patients in the entire cohort of 124 treated patients, we observed specific differences. Among all, smoking status. Although smoking rates continue to decrease, smokers are particularly high among HNSCC patients. We know that tobacco smoking is a major cause of HNSCC, remaining a significant cause of morbidity, and it is well established that HNSCC patients with a significant tobacco smoking history have a poorer prognosis compared to never-smokers, and an increased risk for second primary cancers in other sites (such as lung or esophagus). Tobacco consumption is also associated with inferior treatment-related outcomes, including radiation efficacy[177]. In Nivactor cohort we observed an abundance of non-smokers and past smokers in responders compared to non-responders, that were instead characterized by an increased number of current smokers. Similar data had already been

reported in literature, and smoker patients were classified as patients with less probability to benefit from anti-PD-1 inhibitors [69,178] . The possibility to understand the complexity behind the correlation between smoking and non-response was limited in our trial, given by no differences observed in de-regulated pathways and tumor microenvironment composition between the two categories of non-smokers (never-smokers and past smokers) versus smokers. Moreover, no one of the GE signatures tested resulted associated with smoking habits, and particular DNA mutations associated with smoking status were not recorded. Nevertheless, *de la Iglesia et al* performed an exhaustive report in 2019, underlying the importance of understanding the interaction between smoking status and TME [179]. Even if it is known that carcinogens in tobacco smoke are expected to cause permanent DNA damage and that result should be reflected in TMB value, they did not observe an association between smoking status and TMB status. Moreover, they did not find any significant differences in the mutational profile based on smoking status. Based on their finding, the authors stated that the biological differences between smokers and non-smokers could not be driven by simply smoking-related genomic changes, instead by a more complex TME interaction, driven by gene or protein expression modulated by tobacco exposure. Indeed, they described a clear picture of immunosuppression in active tobacco users with HNSCC, that could explain the difficulty in obtaining a satisficing ICI treatment response for patients with smoking habits.

Considering the subsite of the primary disease, we observed higher percentage of oropharyngeal cancer in responding patients and in oral cavity cancer for non-responding patients. The hypothesis behind this correlation could be partially explained by HPV-infection. It is known that HPV-infection is mostly associated to OPSCC, and that HPV-positive OPSCC represents a specific and distinct disease with its own molecular,

pathological, and clinical features. Moreover, HPV-positive OPSCCs have an improved prognosis compared to HPV-negative OPSCCs [180]. Unfortunately, in Nivactor study data about HPV status were limited, precluding a comparison between HPV-positive OPSCC patients and HPV-negative OPSCC patients (only 28 OPSCC patients were enrolled in Nivactor cohort and data about HPV-infection were not available for 10/28 patients with OPSCC disease). Nevertheless, in Checkmate-141 sustained benefit in long-term (OS=2 years) were observed irrespective of HPV status, and similar results were observed in KEYNOTE-040 and KEYNOTE-048[69,71,72]. However, literature extensively indicates that patients with HPV-positive OPSCC has various mechanisms that contribute to a unique TME, with an increased CD8+ T-cell activation and larger markers of immune infiltration, describing a strong and active immune response. HPV positive and HPV negative OPSCC strongly differ by their strategies to evade the immune recognition, by qualitative and quantitative composition of immune cells, microenvironment composition as well peripheral blood cells[136,181–183]. In Nivactor trial we had the possibility to profile the peripheral blood cells (estimated by xCell method) of a complete responder (it is to note that only 2/ of the 124 treated patients experienced complete response and for both no FFPE material was available) before, during and after nivolumab treatment. The patient had a primary disease in oropharynx and the tumor tested positive to HPV infection. Moreover, patient experienced a JCV infection in the brain, and we hypothesised that the concomitant infections might have a possible role as immune boosters to immunotherapy[176]. Nevertheless, this preliminary hypothesis requires further investigations. Recently, specific clinical trials, focused on HPV-positive patients only, are emerging investigating the role of immunotherapies in this subset of patients. To date, the contribution of HPV infection for the response to immunotherapy remains controversial

[142,184]. After epidemiological-clinical variables, we focused our attention on biological and tumor-related biomarkers. The first biomarker that we investigated was PD-L1, which is the only biomarker validated in prospective trials and approved by FDA. Currently, expression of PD-L1 is assessed by IHC; however, a multitude of IHC antibodies and different score are currently utilized in literature. At present, the dominant scores are TPS (percentage that quantifies the stain of PD-L1 in tumor cells) and CPS (score that quantifies PD-L1 expression in tumor cells, lymphoid cells, and macrophages). Nevertheless, these two scoring methos are not directly comparable and a univocal cut-off level has not been yet established[185]. The lack of a clear definition of what a "positive" PD-L1 tumor is, and the various cut-off considered (ranging from 1 to 50) mirrored the existence of an extreme heterogeneity of PD-L1 in the tumor microenvironment. Various cut-off levels have been proposed for R/M HNSCC, such as CPS ≥ 1 or CPS ≥ 20 in KEYNOTE-048, in which was observed a substantial survival advantage in patients with both CPS ≥ 1 or CPS ≥ 20 (in the arm of patients treated with pembrolizumab monotherapy). However, responders were observed even in the categories of patients with CPS < 1 or CPS < 20. In KEYNOTE-040 the patients that experienced response to pembrolizumab was higher in patients with CPS ≥ 1, and survival longer in patients with TPS≥50%. However, in Checkmate-141, (the most similar trial to Nivactor, to be compared) patients treated with nivolumab appeared to have a longer survival to those treated with SoC, regardless of PD-L1 expression. In Nivactor cohort, a clear correlation between the TPS (% considered in continuum) and response, and CPS (considered in continuum) and disease control rate was observed. However, deepening insight using various and extremely variables cut-offs in the survival analysis, we observed as only significant result when we used CPS ≥ 1. The exploration of novel cut-offs, and cut-

offs higher than those proposed in the present work was limited, due to the median observed and the restricted sample size. Moreover, the inter-samples wide variability of PD-L1 scores reported in Nivactor and the inter-tumors and inter-histologies variability observed in literature (ranging from 14% to 100% even in the same histology type) additionally highlighted the issue with PD-L1 as single predictive biomarker of immunotherapy. This extreme variability made problematic the use of PD-L1 both for determining the responsiveness by for example histology type or for each individual patient. Moreover, despite the broad utility of this biomarker, it is important to underlying that PD-L1 should be considered an imperfect tool: patients with a negative baseline PD-L1 might respond to ICI, while tumors with high PD-L1 expression could be resistant to the therapy[186]. The extreme heterogeneity in PD-L1 expression could indicate that this biomarker has certain limitation and should not be considered as the sole determinant to understand the response to immunotherapy [187]. Therefore, other, and various biomarkers must be considered to understand the ICI efficacy. In Nivactor we investigated another tumor strictly related biomarker: TMB. In the last years, TMB generated by non-synonymous mutations had been proposed as predictive biomarker for immunotherapy in different tumor types and a relationship between high TMB and response to ICI was observed [188]. However, different concerns were raised, such as the demonstration that even patients with TMB < 10 obtained benefits from ICI treatments [189] , or the challenge to reproduce the data and the difficulty in migrating TMB into clinical decision making, due to the numerous procedures and bioinformatic pipelines used in literature. However, concerns were particularly focused on cut-offs used (especially 10 mut/Mb) for defining the patients with “TMB-high” across different tumor types, taking in mind the extremely variability of TMB among cancers, which ranges from 0.01 mut/Mb to more than 4000 mut/Mb[190] .

In literature, other generic cut-offs were proposed[95,191], and among various cut-offs for "high TMB", no one has been reported and so far we lack of standardisation in several cancers. Although, considering the HNSCC reality, it was already indicated that the median number of coding somatic mutations per MB in HNSCC was around 8[179,192], making impossible the evaluation of alternative cut-offs. In Nivactor cohort we confirmed what already found in literature, or rather that is extremely infrequent found out HNSCC patients with TMB-high, and consequently correlations between TMB and response or TMB and survival are difficult to make. A similar discussion could be faced considering microsatellites, that have been observed instable in only 5% of cancer types[193], and which instability is rarely present in HNSCC[98]. Moreover, both the accumulation of point somatic mutations, mirrored by TMB, and instability of microsatellites mostly relied on impairment on mismatch repair system. Wang et al. in 2018 were among the firsts to propose the investigation of a more complex picture of the DNA damage, comprising 8 different repair systems[194], suggesting that co-mutations in different repair pathways could be a predictor of clinical benefit to ICIs, rather than the use of a single biomarker. Even considering the small sample size of Nivactor patients with available DNA sequencing (63/124 patients), we observed a specific and significant trend correlating the response and a more comprehensive picture of mutations in DNA repair systems. Similar remarks could be addressed considering the utility of single point mutations, such as mutations in TP53, KRAS, EGFR and others. Even if associations with response have been proposed and observed, it seems unreliable and vulnerable to translate the result based on a single gene to clinical practice. Indeed, scientists are mostly investigating the possibility to understand mechanisms of resistance through point mutations, instead to highlight a precise predictive biomarker[195]. Because of the intricacy of tumor

biological behavior and immune response, at date, it is unrealistic considering to apply only one single biomarker to fully depict and clarify the prognosis and prediction to immune-therapeutic response. Accordingly, we aimed to investigate the potential role of a multi-biomaker, that could stratify patients into at least two groups, in order to identify those patients that might achieve clinical benefits or response to immunotherapy. Gene expression signatures have rapidly expanded and are considered a more comprehensive tool in order to evaluate the biological contribution of the tumor and its TME and its correlation to response to ICI. Different trials evaluated the expression patterns of immune components and observed a correlation between the signature and response/survival[196]. Among literature all GE signatures, in the present analysis the priority was given to those signatures that was immune-related, such as IFN-γ signature, and those signatures that were centred on HNSCC disease. However, for a multitude of signatures retrieving the methods was problematic, which made us unable to understand their potentiality in our study. Furthermore, the thirteen accessible and tested literature signatures revealed weaknesses in being significantly associated with survival. The eight signatures significantly correlated with PFS or OS (any of those resulted associated with both survivals) were all constructed and (*when*) validated on TCGA or few public dataset (mainly GSE65858 and GSE41613), with the only exception of Ayers et al's signatures, which instead were based on immunotherapy trials, involving patients with different cancer types, and a scarce HNSCC cohort. Even if we support the use of TCGA database (the largest database published since nowadays, containing the genomic, epigenomic, transcriptomic and proteomic data of cancer patients), the recent excessive usage of this rich (but single) and well-annotated resource could provide redundant and unnecessary results, leading to confusion more than understanding. This becomes reality even

more when the achievements are not followed and validated in appropriate prospective/retrospective cohorts, specifically design for the proposed biomarker(s). In the present study, the scarcity of HNSCC recent and public databases challenged us to understand the depth of our results and their predictive validity, for a future and possible clinical application. The selected immune-related signatures were a fundamental tool to understand, more than the translation to their clinical application, the biological characteristic of Nivator patients. All the eight gene expression signatures stratify patients in two different categories, and they especially individuated those patients that did not experience benefits in terms of survival. As an example, six of the eight signatures were de-regulated when we observed the GSEA hallmark pathways, and five (*Bai, Li, Rooney, Ayers IFN-gamma, Liu immune*) of them were associated with a worse survival. All the patients with worse survival of these five signatures were enriched in 5 GSEA Hallmark pathways ("Androgen_response"; "G2M_Checkpoint"; "MYC_targets_V1"; "Protein secretion"; "TGF_Beta_Signaling"). Four signatures were enriched in 10 GSEA Hallmark pathways ("Apical_Junction"; "Epithelial_Mesenchymal_Transition"; "Estrogen_response_early"; "Hypoxia"; "Interferon_Alpha_response"; "Mitotitc_spindle"; "MTORC1_signaling"). Also, other enriched pathways evidenced metabolic re-programming (such as "Adipogenesis"; "Cholesterol_Homeostasis"; "Glycolysis"). It has already been observed tumor shapes mirroring TME and TME adapts itself to tumor biology [197]. A highly diverse spectrum of mechanism, involving TME, have been described and involved in immune escape of HNSCC. For example, TME is influenced by tumor cells and increased hypoxia has been observed. Nevertheless, hypoxic conditions influence tumor cells as well infiltrating immune cells. In reaction to hypoxia (fueled by oxidative phosphorylation

and the crosstalk with WNT beta catenin signaling pathways [198]) HNSCC switches to glycolysis, which leads to the production of lactic acid and further reduce pH. Consequently, the pH alterations influence the repertoire and the activity of immune modulator molecules and cells functionality. As an example, T-cells activation, proliferation, and cytotoxicity is reduced, macrophages and regulatory T-cells (Tregs) are accumulated and create an immune suppressive environment. Expression of tumor necrosis factor (TNF) receptor family members is increased, and IFN-gamma pathways related are downregulated. Additionally, hypoxic conditions induce the accumulation of TAMs, which are responsible for epithelial mesenchymal transition and angiogenesis. Angiogenesis is moreover correlated with MYC pathways, that are described as key factors for cell de-differentiation and for immune suppression in the microenvironment[199]. All these pathways additionally contribute to lower the pH[197]. Moreover, another example includes the alteration in cholesterol homeostasis, which is frequently observed in cancer. Cholesterol has immunomodulatory properties (for both innate and adaptive immunity), and its accumulation in TME has been associated (among others) with CD8+ T-cell exhaustion[200]. Additionally, we observed different proliferation and DNA damage related pathways enriched in patients with the worse prognosis, such as “G2M_Checkpoint”, “E2F_targets”, “P53 pathway”, “PI3K_AKT_MTOR_signaling”, “DNA_repair”, “UV response”. These pathways clearly indicated a perturbation of response mechanisms (intrinsic of healthy cells) to stressful conditions (such as hypoxia and oxygen deprivation). The de-regulation of proliferative pathways is typically described in cancer, and it is usually correlated with a more aggressive phenotype. These results were moreover confirmed observing the function of the 10 most mutated genes in all the subgroup of patients with the availability DNA annotations.

Hallmark gene-sets enriched in patients with worse survival (*Bai, Li, Rooney, Ayers IFN-gamma, Liu immune*)		
Enriched in 5/5	**Enriched in 4/5**	**Enriched in 3/5**
ANDROGEN_RESPONSE	APICAL_JUNCTION	ADIPOGENESIS
G2M_CHECKPOINT	EMT	APOPTOSIS
MYC_TARGETS_V1	ESTROGEN_RESPONSE_EARLY	CHOLESTEROL_HOMEOSTASIS
PROTEIN_SECRETION	HYPOXIA	COMPLEMENT
TGF_BETA_SIGNALING	INTERFERON_ALPHA_RESPONSE	DNA_REPAIR
	MITOTIC_SPINDLE	E2F_TARGETS
	MTORC1_SIGNALING	GLYCOLYSIS
	OXIDATIVE PHOSPHORILATION	HEDGEHOG_SIGNALING
	TNFA_SIGNALING_VIA_NFKB	HEME_METABOLISM
	UNFOLDED_PROTEIN_RESPONSE	IL2_STAT5_SIGNALING
	UV_RESPONSE_DN	IL6_JAK_STAT3_SIGNALING
		INFLAMMATORY_RESPONSE
		MYC_TARGETS_V2
		NOTCH_SIGNALING
		P53_PATHWAY
		PI3K_AKT_MTOR_SIGNALING
		REACTIVE_OXYGEN_SPECIES_PATHWAY
		UV_RESPONSE_UP
		WNT_BETA_CATENIN_SIGNALING

Interestingly, Chen et al. described[136] three immune conditions (non-immune; active immune; exhausted immune) in the TME of HNSCC patients. Among the mechanisms identified in poor survival patients according to the 5 signatures tested, several pathways were related to immune-exhaustion by Chen et al., such as WNT and TGF-beta signalling pathways. Moreover, they individuated a situation in which an immune phenotype did not absolutely predict activity of the immune cells and consequently immunotherapy response. For instance, they observed that

IFN signaling, T-cell-related signatures, high immune cell infiltration, enhanced cytolysis did not contribute to understand the differences between an active immune system and an exhausted immune system. Indeed, in Nivactor we observed an enrichment in several lymphoid populations for groups with the worse survival (5 literature signatures: *Liu myeloid*, *IFN-gamma, Rooney*, *Hu*, *Li*): B-cells (2/5), CD4+ Tem, CD4+ T-cells and CD4+ memory T-cells (3/5), CD8+ T-cells and CD8+ Tcm (2/5), Class-switched memory B-Cells (2/5), Common Lymphoid progenitor (CLP, 4/5). These immune cells usually play an important role in promoting anti-tumor immune response. Nevertheless, TME of the groups with the worse prognosis was also enriched in myeloid-derived suppressor cells, cuch as Tregs (2/5), Macrophages (4/5) and stromal components, confirming the hypothesis of an immune exhaustion in this subset of patients. When we used cluster subtypes, already associated to biology and prognosis of HNSCC, asimilar biological profile (observed for GSEA Hallmark, KEGG and xCell) was observed with Cluster-5 subtype (*De Cecco et al.*[39]). Differently from the immune signatures the Cluster-5 resulted significant in survival analysis and presented a significant correlation with response. Specifically, Cluster-5 seemed to strongly identify those patients that did not obtain any benefit from immunotherapy. The two *De Cecco et al* (Cluster-5 and Cluster-6) subtypes appeared to be more reliable tool with a strong prognostic significance and a correlation with response. The possible explanation could be based on the enormous amount of gene expression data considered for the subtype construction in the original paper (11 public datasets were used) and a specific method used for the parallelism and subtraction of normal tissue features. However, the two clusters, biologically well characterized in 2015, struggled to maintain their identity in the Nivactor analysis. While the Cluster-5 acquired new traits (highly similar to traits of the literature

signatures), the Cluster-6 did not evidence any specific biological characteristics, with the exception of NKT (highly represented in the Cluster-6 low group, the one with the worse prognosis). The up-regulation of NKT was moreover confirmed in the poorer prognosis evidenced by Cluster-5 low. It is already known that NKT cells are among of the most highly infiltrated cells in HNSCC TME. However, at date, the role of NKT for HNSCC is still debatable in literature, and their role seems to be various. Moreover, the debate is accentuated when NKT abundance and function are studied in peripheral blood instead in the TME [201]. While some authors declared that these immune cells are a critical component in the early phase of immune response against tumor in the TME, and they are associated with better prognosis[202], other authors individuated a correlation of NK cells and worse prognosis and tumor progression[203], due to the NK role that negatively regulates T-cell activity, leading to immunosuppression [204]. Moreover, the two signatures did not express their potentiality investigating some key points of immune-related genes, interrogating whether their capability was associated to the mere immune cell infiltration or rather to the complexity of the TME. The biological characterization of Cluster-6 in the original paper, highlighted that this subtype was the closest to the normal state, expressing similarity with the airway epithelium and the one maintaining active cellular homeostasis. The possible explanation behind the performance of Cluster-6 in Nivactor (small group of patients with better PSF and OS survival), might be associated to the mirroring of several characteristics of normal state, thus enabling to identify the patients more prone to respond to immunotherapy. However, this preliminary hypothesis needs to be validated in a larger prospective cohort of R/M HNSCC patients treated with immunotherapy or in a selected population with extremized long and short survival patients.

An important limitation of the present study was the lack of a public available dataset focused on immunotherapy real life data. Nevertheless, the only use of Nivactor dataset generated promising results. The two *De Cecco et al.* subtypes identified two small subsets of patients that could have beneficial results in terms of survival: 15 patients were classified as Cluster-6 high (associated to better survival) and 12 patients were classified as Cluster-5 low (associated to worse survival). The two signatures highly and directly correlated together, suggesting the interesting prospective of their integration. However, despite the possible and speculative interpretation of biological results, several open questions about the biological contribution of the two subtypes remained unanswered and further investigations to individuate the biological contribution of each of the two clusters are required.

5.1 Strengths, limitations, and future directions

The unsatisfactory results obtained by DNA profiling (even when integrated with gene expression results) and the promising results observed only by the gene expression analyses are concordant to the emerging evidence that described epigenetic events highly important in HNSCC, probably even more than the mere genetic alterations[205,206]. DNA methylation (hypo- and hyper-), histone modifications (acetylation, methylation, phosphorylation, etc), non-coding RNAs (microRNAs, lncRNAs, etc) are all epigenomic mechanisms that are highly involved in carcinogenesis, tumor progression (and consequently in patients' prognosis), and TME regulation in HNSCC (i.e., immunosuppression mediated by myeloid-derived cells and CAFs). Future studies involving the profiling of epigenomic characteristics, to understand the modifications and the crosstalk between tumor and TME could provide answers to specific open questions related to the functionality of

immunotherapy in HNSCC, aimed to a more personalized medicine[207]. Moreover, the complex immune composition of tissue from HNSCC patients could be better discerned by additional analyses through single cell approaches (i.e., single cell RNA sequencing). This high-resolution transcriptomic analysis will give the opportunity to characterize the spatial localization of immune cells and the other cellular component of TME, their crosstalk and transcriptional spatial states. The deep understanding will allow the characterization the intra-tumoral heterogeneity, and to build a specific molecular model of exhaustion[208] to correlate with HNSCC subtypes. We should consider that the characterization of biology of the tumor sample is currently achieved through the only profiling of biopsy (mostly FFPE > fresh/frozen; mostly biopsy > surgery). Thus, the use of these specific materials could be associated to several problems, such as: i) the invasive nature of the biopsy; ii) the small size of the tissue and therefore, the limited amount of nucleic acid for omic profiling; iii) the lack of the tissue due to precedent use for the diagnostic testing; iv) bias due to intrinsic tumor heterogeneity (that, using a small biopsy, risks to be missed or misunderstood,) that means tissue from a certain area may have different mutations and gene expression than other areas (providing inaccurate information); v) preservation methods such as formalin fixation, that cause C > T transition through deamination of cytosine and could lead to false positive to DNA profiling; vi) heterogeneity of tumor tissue samples collected (frequently derived from or primary tumor or recurrence or metastasis); vii) biopsy allows the detection of a dominant biological pattern at a precise time point (usually before treatment) and do not concede to observe the high dynamism of tumors (which is also caused by selective pressure during drug treatment)[209]. Nevertheless, the reality of translational studies based on clinical trials for HNSCC allows the possibility to usually retrieve formalin fixed tumor tissue (instead of

fresh/frozen tissue; range of collection years goes from 1 year from the first diagnosis to 10 years from the first diagnosis) from various stages of disease (mostly before treatment; however, only one sample for patient usually is available, not allowing the paired comparison between primary and metastasis or primary and normal tissue) and small biopsies (due to the anatomical sites considered). In Nivactor cohort we had the possibility to collect the tumor tissue from 75% of the patients (94/124) and the right amount of nucleic acid for omic profiling from 65% (80/124, gene expression) and 51% (63/124, DNA sequencing) of the samples. Unfortunately, we did not dispose of the FFPE tumor tissue of two complete responders, losing a relevant piece of the puzzle. However, data about the observed dropouts are in line with the ones reported in literature in general cancer studies[210]. Additionally, the lack of tumor material is even increased when the profiling includes tumor tissue of R/M patients treated in second line, according to what we already observed in other HNSCC studies[211] . The Nivactor study was an academic study, that demonstrated solidity and strength (the clinical-pathological data resulted in line with previous clinical trial published) and allowed the discerning of several biological features. The biological dataset of the current study will be deposited and published, to allow the utilize of these important data to the academic community. Even if we were able and proud to obtain satisfactory results in the Nivactor trial, we aim to improve these limitations. For instance, recent technological improvements had been achieved for the accurate and highly sensitive profiling of liquid biopsy (i.e., blood, urine, stool, and saliva samples). In oncology, the use of liquid biopsy is currently allowing patient screening, prognostic stratification, monitoring treatment response (before, during and after the therapy) and detection of minimal residual disease after surgery/recurrence. The use of liquid biopsies is growing in relevance, mostly because it is a non-invasive

method, which allows overcoming the traditional tissue-based bias. Moreover, the use of liquid biopsy permits longitudinal monitoring of patients' disease (gaining a richer molecular understanding of the specific cancer considered). Even if in liquid biopsy the relative amount of nucleic acids and cells (small RNAs, mRNAs, lncRNAs, cfDNA, exosomes and others) derived from a tumor (against the background) is low, in the majority of cases the amount of material collected is adequate to extract the nucleic acids for the molecular profiling. Even if tissue biopsies are irreplaceable for identify the specific genes' expression and mutations of the tumor, the liquid biopsy could be a good candidate to understand the mechanisms at the base of tumor relapse- metastatization, resistance to therapy or for early detection of cancer. Moreover, they relevance as predictive factor for immunotherapy is growing[212,213]. Specifically, for a cohort of patients in the Nivactor trial we collected the blood samples before and during therapy. The collection permitted us to acquire important knowledge about specific and peculiar characteristics of a responding patient during a follow-up of more than 3 years[176]. As a future direction we aim to profile for a subset of Nivactor patients, some of the known circulating biomarkers, such as cfDNA (for detecting its abundance and specific mutations), DNA (for T-cell receptors sequencing), mRNAs (allowing the inference of peripheral immune-cell population), miRNAs (and other epigenomic regulators) to better comprehend the longitudinal biology during the nivolumab treatment. Moreover, the peripheral immune profiling will allow to correlate the profile of the tumor biology and the molecular features observed in liquid biopsy. Consequently, the liquid biopsy could function as well as surrogate for depict the profile of those patients for which we lacked the FFPE material. Another important aspect that needs to be underlying is the background of HNSCC patients. Head and neck cancers have been exhaustively described as an immune

suppressed tumor, in which from the very first steps different escaping mechanisms are activated by cancer cells. Moreover, increasing evidence are suggesting that HNSCC TME is immunosuppressive, and play a vital role in ICI resistance[214–216]. The study of a cohort of patients with rare disease with a strong immunosuppression that hardly responds to immunotherapy had generated several challenges, firstly due to the impossibility to profile the tumor of only two complete responders. Also, reasoning in numbers, the profiling of the partial responding patients resulted reduced if compared to patients with a stable disease or for which a progression was observed. However, the overall numbers, even if in line with other studies in literature, strongly reduced the power of our analyses and the exhaustive biological deepening. Nevertheless, the suggested picture underlies that cancer immunotherapy strongly differs from other therapy approaches, such as chemotherapy or targeted therapies, in which cancer is flighted by a single (or more than one) targeted oncogenic feature(s). Cancer immunotherapy involves several, dynamic and simultaneous biological targets indicating that immune-cell infiltration is necessary to induce a response to treatment, but it is not sufficient alone. Different preclinical studies are now evaluating the possibility to combine immunotherapy with other drugs, such as those that inhibits oxidative phosphorylation (to alleviate hypoxia condition)[217] or WNT-signaling pathways, that regulates above all cell proliferation and tissue homeostasis, with indirect effects on T-cell functionality[218].

6. CONCLUSIONS

The present translational work, built up on a solid phase IIIb Italian clinical trial, corroborated the intrinsic and complex TME heterogeneity in R/M HNSCC patients. Thanks to these efforts, several features of immunosuppression have been identified. These results allowed the possible speculation of the TME correlation with low response rate to immunotherapy for these patients and a possible prospective of new biological targets for the treatment of R/M HNSCC patients. Although the use of a single predictive biomarker for individuate the response to immunotherapy seems unrealistic, a more complex tool, such gene expression signatures, appeared to be more reliable for this hard task. Future analyses and datasets are required to validate the prognostic/predictive role of Cluster-5 and Cluster-6 (integrated or considered as alone). Additional biological exploration and functional analyses will allow to untangle the complexity of HNSCC biology, to create a specific and detailed picture of the intrinsic TME heterogeneity.

7. REFERENCES

1. Haridy, Y. *et al.* Triassic Cancer - Osteosarcoma in a 240-Million-Year-Old Stem-Turtle. *JAMA Oncology* vol. 5 Preprint at https://doi.org/10.1001/jamaoncol.2018.6766 (2019).

2. Hajdu, S. I. A note from history: Landmarks in history of cancer, part 1. *Cancer* vol. 117 Preprint at https://doi.org/10.1002/cncr.25553 (2011).

3. The global challenge of cancer. *Nature cancer* vol. 1 Preprint at https://doi.org/10.1038/s43018-019-0023-9 (2020).

4. Casás-Selves, M. & Degregori, J. How Cancer Shapes Evolution and How Evolution Shapes Cancer. *Evolution: Education and Outreach* vol. 4 Preprint at https://doi.org/10.1007/s12052-011-0373-y (2011).

5. Yuan, Y. Spatial heterogeneity in the tumor microenvironment. *Cold Spring Harb Perspect Med* **6**, (2016).

6. Guy, G. P. *et al.* Economic burden of cancer survivorship among adults in the United States. *Journal of Clinical Oncology* **31**, (2013).

7. Luengo-Fernandez, R., Leal, J., Gray, A. & Sullivan, R. Economic burden of cancer across the European Union: A population-based cost analysis. *Lancet Oncol* **14**, (2013).

8. Bray, F. *et al.* Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* **68**, (2018).

9. Fitzmaurice, C. *et al.* Global, regional, and national cancer incidence, mortality, years of life lost, years lived with disability, and disability-Adjusted life-years for 29 cancer groups, 1990 to 2017: A systematic analysis for the global burden of disease study. *JAMA Oncol* **5**, (2019).

10. Siegel, R. L., Miller, K. D., Fuchs, H. E. & Jemal, A. Cancer statistics, 2022. *CA Cancer J Clin* **72**, (2022).

11. 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* **71**, (2021).

12. Pai, S. I. & Westra, W. H. Molecular pathology of head and neck cancer: Implications for diagnosis, prognosis, and treatment. *Annual Review of Pathology: Mechanisms of Disease* vol. 4 Preprint at https://doi.org/10.1146/annurev.pathol.4.110807.092158 (2009).

13. van Liew, J. R. *et al.* Weight loss after head and neck cancer: A dynamic relationship with depressive symptoms. *Head Neck* **39**, (2017).

14. Callahan, C. Facial disfigurement and sense of self in head and neck cancer. *Soc Work Health Care* **40**, (2004).

15. Ferlay, J. *et al.* Global cancer observatory: cancer today. *International Agency for Research on Cancer* (2020).

16. Stepnick, D. & Gilpin, D. Head and Neck Cancer: An Overview. *Semin Plast Surg* **24**, (2010).

17. Ojeda, D., Huber, M. A. & Kerr, A. R. Oral Potentially Malignant Disorders and Oral Cavity Cancer. *Dermatologic Clinics* vol. 38 Preprint at https://doi.org/10.1016/j.det.2020.05.011 (2020).

18. Chamoli, A. *et al.* Overview of oral cavity squamous cell carcinoma: Risk factors, mechanisms, and diagnostics. *Oral Oncology* vol. 121 Preprint at https://doi.org/10.1016/j.oraloncology.2021.105451 (2021).

19. Conway, D. I., Purkayastha, M. & Chestnutt, I. G. The changing epidemiology of oral cancer: Definitions, trends, and risk factors. *Br Dent J* **225**, (2018).

20. Bradford, C. R., Ferlito, A., Devaney, K. O., Mäkitie, A. A. & Rinaldo, A. Prognostic factors in laryngeal squamous cell carcinoma. *Laryngoscope Investigative Otolaryngology* vol. 5 Preprint at https://doi.org/10.1002/lio2.353 (2020).

21. Myers, J. E. *et al.* Detecting episomal or integrated human papillomavirus 16 DNA using an exonuclease V-qPCR-based assay. *Virology* **537**, (2019).

22. Shah, A. *et al.* Oral sex and human papilloma virus-related head and neck squamous cell cancer: A review of the literature. *Postgraduate Medical Journal* vol. 93 Preprint at https://doi.org/10.1136/postgradmedj-2016-134603 (2017).

23. Taberna, M. *et al.* Human papillomavirus-related oropharyngeal cancer. *Annals of Oncology* vol. 28 Preprint at https://doi.org/10.1093/annonc/mdx304 (2017).

24. Garneau, J. C., Bakst, R. L. & Miles, B. A. Hypopharyngeal cancer: A state of the art review. *Oral Oncol* **86**, (2018).

25. Lydiatt, W. M. *et al.* Head and neck cancers-major changes in the American Joint Committee on cancer eighth edition cancer staging manual. *CA Cancer J Clin* **67**, (2017).

26. Rothschild, U. *et al.* Immunotherapy in head and neck cancer - scientific rationale, current treatment options and future directions. *Swiss medical weekly* vol. 148 Preprint at https://doi.org/10.4414/smw.2018.14625 (2018).

27. Cancer Research UK. Head and neck cancers statistics. *Cancer Research UK* **14**, (2019).

28. Lu, M. & Zhan, X. The crucial role of multiomic approach in cancer research and clinically relevant outcomes. *EPMA Journal* vol. 9 Preprint at https://doi.org/10.1007/s13167-018-0128-8 (2018).

29. Serafini, M. S. *et al.* Transcriptomics and Epigenomics in head and neck cancer: available repositories and molecular signatures. *Cancers Head Neck* **5**, (2020).

30. Vogelstein, B. & Kinzler, K. W. Cancer genes and the pathways they control. *Nature Medicine* vol. 10 Preprint at https://doi.org/10.1038/nm1087 (2004).

31. Hanahan, D. & Weinberg, R. A. Hallmarks of cancer: The next generatio1. Hanahan D, Weinberg RA. Hallmarks of cancer: The next generation. Cell [Internet]. 2011;144(5):646–74. Available from: http://dx.doi.org/10.1016/j.cell.2011.02.013n. *Cell* **144**, (2011).

32. Yates, L. R. & Campbell, P. J. Evolution of the cancer genome. *Nature Reviews Genetics* vol. 13 Preprint at https://doi.org/10.1038/nrg3317 (2012).

33. Patmore, H. S., Cawkwell, L., Stafford, N. D. & Greenman, J. Unraveling the chromosomal aberrations of head and neck squamous cell carcinoma: A review. *Annals of Surgical Oncology* vol. 12 Preprint at https://doi.org/10.1245/ASO.2005.09.017 (2005).

34. Leemans, C. R., Braakhuis, B. J. M. & Brakenhoff, R. H. The molecular biology of head and neck cancer. *Nature Reviews Cancer* vol. 11 Preprint at https://doi.org/10.1038/nrc2982 (2011).

35. Tsimberidou, A. M., Fountzilas, E., Bleris, L. & Kurzrock, R. Transcriptomics and solid tumors: The next frontier in precision cancer medicine. *Seminars in Cancer Biology* vol. 84 Preprint at https://doi.org/10.1016/j.semcancer.2020.09.007 (2022).

36. Lawrence, M. S. *et al.* Comprehensive genomic characterization of head and neck squamous cell carcinomas. *Nature* **517**, (2015).

37. Walter, V. *et al.* Molecular Subtypes in Head and Neck Cancer Exhibit Distinct Patterns of Chromosomal Gain and Loss of Canonical Cancer Genes. *PLoS One* **8**, (2013).

38. Keck, M. K. *et al.* Integrative analysis of head and neck cancer identifies two biologically distinct HPV and three non-HPV subtypes. *Clinical Cancer Research* **21**, (2015).

39. de Cecco, L. *et al.* Head and neck cancer subtypes with biological and clinical relevance: Meta-analysis of gene-expression data. *Oncotarget* **6**, (2015).

40. Tonella, L., Giannoccaro, M., Alfieri, S., Canevari, S. & de Cecco, L. Gene Expression Signatures for Head and Neck Cancer Patient Stratification: Are Results Ready for Clinical Application? *Current Treatment Options in Oncology* vol. 18 Preprint at https://doi.org/10.1007/s11864-017-0472-2 (2017).

41. Hanahan, D. & Coussens, L. M. Accessories to the Crime: Functions of Cells Recruited to the Tumor Microenvironment. *Cancer Cell* vol. 21 Preprint at https://doi.org/10.1016/j.ccr.2012.02.022 (2012).

42. Lorusso, G. & Rüegg, C. The tumor microenvironment and its contribution to tumor evolution toward metastasis. *Histochemistry and Cell Biology* vol. 130 Preprint at https://doi.org/10.1007/s00418-008-0530-8 (2008).

43. Farc, O. & Cristea, V. An overview of the tumor microenvironment, from cells to complex networks (Review). *Exp Ther Med* **21**, (2020).

44. Baghban, R. *et al.* Tumor microenvironment complexity and therapeutic implications at a glance. *Cell Communication and Signaling* vol. 18 Preprint at https://doi.org/10.1186/s12964-020-0530-4 (2020).

45. Gonzalez, H., Hagerling, C. & Werb, Z. Roles of the immune system in cancer: From tumor initiation to metastatic progression. *Genes and Development* vol. 32 Preprint at https://doi.org/10.1101/GAD.314617.118 (2018).

46. Jin, M. Z. & Jin, W. L. The updated landscape of tumor microenvironment and drug repurposing. *Signal Transduction and Targeted Therapy* vol. 5 Preprint at https://doi.org/10.1038/s41392-020-00280-x (2020).

47. Schneider, G. *et al.* Cross talk between stimulated NF-κB and the tumor suppressor p53. *Oncogene* **29**, (2010).

48. de Stefani, A. *et al.* Improved survival with perilymphatic interleukin 2 in patients with resectable squamous cell carcinoma of the oral cavity and oropharynx. *Cancer* **95**, (2002).

49. Nagarsheth, N., Wicha, M. S. & Zou, W. Chemokines in the cancer microenvironment and their relevance in cancer immunotherapy. *Nature Reviews Immunology* vol. 17 Preprint at https://doi.org/10.1038/nri.2017.49 (2017).

50. Näsman, A. *et al.* Tumor infiltrating CD8 + and Foxp3 + Lymphocytes correlate to clinical outcome and human papillomavirus (HPV) status in Tonsillar cancer. *PLoS One* **7**, (2012).

51. Xu, Q., Wang, C., Yuan, X., Feng, Z. & Han, Z. Prognostic value of tumor-infiltrating lymphocytes for patients with head and neck squamous cell carcinoma. *Transl Oncol* **10**, (2017).

52. Sharonov, G. v., Serebrovskaya, E. O., Yuzhakova, D. v., Britanova, O. v. & Chudakov, D. M. B cells, plasma cells and antibody repertoires in the tumour microenvironment. *Nature Reviews Immunology* vol. 20 Preprint at https://doi.org/10.1038/s41577-019-0257-x (2020).

53. Kinker, G. S. *et al.* B Cell Orchestration of Anti-tumor Immune Responses: A Matter of Cell Localization and Communication. *Frontiers in Cell and Developmental Biology* vol. 9 Preprint at https://doi.org/10.3389/fcell.2021.678127 (2021).

54. Yuen, G. J., Demissie, E. & Pillai, S. B Lymphocytes and Cancer: A Love–Hate Relationship. *Trends in Cancer* vol. 2 Preprint at https://doi.org/10.1016/j.trecan.2016.10.010 (2016).

55. Fridman, W. H. *et al.* B cells and cancer: To B or not to B? *Journal of Experimental Medicine* vol. 218 Preprint at https://doi.org/10.1084/JEM.20200851 (2021).

56. Gavrielatou, N., Vathiotis, I., Economopoulou, P. & Psyrri, A. The role of b cells in head and neck cancer. *Cancers* vol. 13 Preprint at https://doi.org/10.3390/cancers13215383 (2021).

57. Arruebo, M. *et al.* Assessment of the evolution of cancer treatment therapies. *Cancers* vol. 3 Preprint at https://doi.org/10.3390/cancers3033279 (2011).

58. Krzyszczyk, P. *et al.* The growing role of precision and personalized medicine for cancer treatment. *Technology (Singap World Sci)* **06**, (2018).

59. Marur, S. & Forastiere, A. A. Head and Neck Squamous Cell Carcinoma: Update on Epidemiology, Diagnosis, and Treatment. *Mayo Clinic Proceedings* vol. 91 Preprint at https://doi.org/10.1016/j.mayocp.2015.12.017 (2016).

60. Winquist, E., Agbassi, C., Meyers, B. M., Yoo, J. & Chan, K. K. W. Systemic therapy in the curative treatment of head and neck squamous cell cancer: a systematic review. *Journal of Otolaryngology - Head and Neck Surgery* vol. 46 Preprint at https://doi.org/10.1186/s40463-017-0199-x (2017).

61. Anderson, G. *et al.* An updated review on head and neck cancer treatment with radiation therapy. *Cancers* vol. 13 Preprint at https://doi.org/10.3390/cancers13194912 (2021).

62. Saddawi-Konefka, R. *et al.* Defining the Role of Immunotherapy in the Curative Treatment of Locoregionally Advanced Head and Neck Cancer: Promises, Challenges, and Opportunities. *Frontiers in Oncology* vol. 11 Preprint at https://doi.org/10.3389/fonc.2021.738626 (2021).

63. Elmusrati, A., Wang, J. & Wang, C. Y. Tumor microenvironment and immune evasion in head and neck squamous cell carcinoma. *International Journal of Oral Science* vol. 13 Preprint at https://doi.org/10.1038/s41368-021-00131-7 (2021).

64. Sim, F., Leidner, R. & Bell, R. B. Immunotherapy for Head and Neck Cancer. *Hematology/Oncology Clinics of North America* vol. 33 Preprint at https://doi.org/10.1016/j.hoc.2018.12.006 (2019).

65. Kao, H. F. & Lou, P. J. Immune checkpoint inhibitors for head and neck squamous cell carcinoma: Current landscape and future directions. in *Head and Neck* vol. 41 (2019).

66. Peng, Y. P. *et al.* Combination of Tumor Mutational Burden and Specific Gene Mutations Stratifies Outcome to Immunotherapy Across Recurrent and Metastatic Head and Neck Squamous Cell Carcinoma. *Front Genet* **12**, (2021).

67. Shibata, H., Saito, S. & Uppaluri, R. Immunotherapy for Head and Neck Cancer: A Paradigm Shift From Induction Chemotherapy to Neoadjuvant Immunotherapy. *Frontiers in Oncology* vol. 11 Preprint at https://doi.org/10.3389/fonc.2021.727433 (2021).

68. Botticelli, A. *et al.* Anti–PD-1 and Anti–PD-L1 in Head and Neck Cancer: A Network Meta-Analysis. *Frontiers in Immunology* vol. 12 Preprint at https://doi.org/10.3389/fimmu.2021.705096 (2021).

69. Ferris, R. L. *et al.* Nivolumab vs investigator's choice in recurrent or metastatic squamous cell carcinoma of the head and neck: 2-year long-term survival update of CheckMate 141 with analyses by tumor PD-L1 expression. *Oral Oncol* **81**, (2018).

70. Mehra, R. *et al.* Efficacy and safety of pembrolizumab in recurrent/metastatic head and neck squamous cell carcinoma: Pooled analyses after long-term follow-up in KEYNOTE-012. *Br J Cancer* **119**, (2018).

71. Cohen, E. E. W. *et al.* Pembrolizumab versus methotrexate, docetaxel, or cetuximab for recurrent or metastatic head-and-neck squamous cell carcinoma (KEYNOTE-040): a randomised, open-label, phase 3 study. *The Lancet* **393**, (2019).

72. Burtness, B. *et al.* Pembrolizumab alone or with chemotherapy versus cetuximab with chemotherapy for recurrent or metastatic squamous cell carcinoma of the head and neck (KEYNOTE-048): a randomised, open-label, phase 3 study. *The Lancet* **394**, (2019).

73. Nagasaka, M. *et al.* PD1/PD-L1 inhibition as a potential radiosensitizer in head and neck squamous cell carcinoma: A case report. *J Immunother Cancer* **4**, (2016).

74. Pai, S. I. *et al.* Comparative analysis of the phase III clinical trials of anti-PD1 monotherapy in head and neck squamous cell carcinoma patients (CheckMate 141 and KEYNOTE 040). *Journal for ImmunoTherapy of Cancer* vol. 7 Preprint at https://doi.org/10.1186/s40425-019-0578-0 (2019).

75. Cramer, J. D., Burtness, B. & Ferris, R. L. Immunotherapy for head and neck cancer: Recent advances and future directions. *Oral Oncology* vol. 99 Preprint at https://doi.org/10.1016/j.oraloncology.2019.104460 (2019).

76. Argiris, A. *et al.* A randomized, open-label, phase 3 study of nivolumab in combination with ipilimumab vs extreme regimen (cetuximab + cisplatin/carboplatin + fluorouracil) as first-line therapy in patients with recurrent or metastatic squamous cell carcinoma of the head and neck-CheckMate 651. *Annals of Oncology* **27**, (2016).

77. Muzaffar, J., Bari, S., Kirtane, K. & Chung, C. H. Recent advances and future directions in clinical management of head and neck squamous cell carcinoma. *Cancers* vol. 13 Preprint at https://doi.org/10.3390/cancers13020338 (2021).

78. Conroy, M. & Naidoo, J. Immune-related adverse events and the balancing act of immunotherapy. *Nature Communications* vol. 13 Preprint at https://doi.org/10.1038/s41467-022-27960-2 (2022).

79. Łasińska, I. *et al.* Immunotherapy in Patients with Recurrent and Metastatic Squamous Cell Carcinoma of the Head and Neck. *Anticancer Agents Med Chem* **19**, (2018).

80. Przybylski, K., Majchrzak, E., Weselik, L. & Golusiński, W. Immunotherapy of head and neck squamous cell carcinoma (HNSCC). Immune checkpoint blockade. *Otolaryngologia Polska* **72**, (2018).

81. Veigas, F. *et al.* Immune checkpoints pathways in head and neck squamous cell carcinoma. *Cancers* vol. 13 Preprint at https://doi.org/10.3390/cancers13051018 (2021).

82. Kugel, C. H. *et al.* Age correlates with response to anti-PD1, reflecting age-related differences in intratumoral effector and regulatory T-cell populations. *Clinical Cancer Research* **24**, (2018).

83. Bupp, M. R. G., Potluri, T., Fink, A. L. & Klein, S. L. The confluence of sex hormones and aging on immunity. *Frontiers in Immunology* vol. 9 Preprint at https://doi.org/10.3389/fimmu.2018.01269 (2018).

84. Jain, V. *et al.* Association of Age with Efficacy of Immunotherapy in Metastatic Melanoma. *Oncologist* **25**, (2020).

85. Oliva, M. *et al.* Immune biomarkers of response to immune-checkpoint inhibitors in head and neck squamous cell carcinoma. *Annals of Oncology* vol. 30 Preprint at https://doi.org/10.1093/annonc/mdy507 (2019).

86. Cavalieri, S. *et al.* Immuno-oncology in head and neck squamous cell cancers: News from clinical trials, emerging predictive factors and unmet needs. *Cancer Treatment Reviews* vol. 65 Preprint at https://doi.org/10.1016/j.ctrv.2018.03.003 (2018).

87. Duffy, M. J. & Crown, J. Biomarkers for predicting response to immunotherapy with immune checkpoint inhibitors in cancer patients. *Clinical Chemistry* vol. 65 Preprint at https://doi.org/10.1373/clinchem.2019.303644 (2019).

88. Seiwert, T. Y. *et al.* Abstract LB-339: Biomarkers predictive of response to pembrolizumab in head and neck cancer (HNSCC). *Cancer Res* **78**, (2018).

89. Lee, M., Samstein, R. M., Valero, C., Chan, T. A. & Morris, L. G. T. Tumor mutational burden as a predictive biomarker for checkpoint inhibitor immunotherapy. *Hum Vaccin Immunother* **16**, (2020).

90. McNamara, M. G. *et al.* Impact of high tumor mutational burden in solid tumors and challenges for biomarker application. *Cancer Treatment Reviews* vol. 89 Preprint at https://doi.org/10.1016/j.ctrv.2020.102084 (2020).

91. Büttner, R. *et al.* Implementing TMB measurement in clinical practice: Considerations on assay requirements. *ESMO Open* vol. 4 Preprint at https://doi.org/10.1136/esmoopen-2018-000442 (2019).

92. Meléndez, B. *et al.* Methods of measurement for tumor mutational burden in tumor tissue. *Translational Lung Cancer Research* vol. 7 Preprint at https://doi.org/10.21037/tlcr.2018.08.02 (2018).

93. Klempner, S. J. *et al.* Tumor Mutational Burden as a Predictive Biomarker for Response to Immune Checkpoint Inhibitors: A Review of Current Evidence. *Oncologist* **25**, (2020).

94. Savill, K. M. Z., Zettler, M. E., Feinberg, B. A., Jeune-Smith, Y. & Gajra, A. Awareness and utilization of tumor mutation burden (TMB) as a biomarker for administration of immuno-oncology (I-O) therapeutics by practicing community oncologists in the United States (U.S.). *Journal of Clinical Oncology* **39**, (2021).

95. Zheng, M. Tumor mutation burden for predicting immune checkpoint blockade response: The more, the better. *J Immunother Cancer* **10**, (2022).

96. Mouw, K. W., Goldberg, M. S., Konstantinopoulos, P. A. & D'Andrea, A. D. DNA damage and repair biomarkers of immunotherapy response. *Cancer Discovery* vol. 7 Preprint at https://doi.org/10.1158/2159-8290.CD-17-0226 (2017).

97. Chalmers, Z. R. *et al.* Analysis of 100,000 human cancer genomes reveals the landscape of tumor mutational burden. *Genome Med* **9**, (2017).

98. Bonneville, R. *et al.* Landscape of Microsatellite Instability Across 39 Cancer Types. *JCO Precis Oncol* (2017) doi:10.1200/po.17.00073.

99. Baretti, M. & Le, D. T. DNA mismatch repair in cancer. *Pharmacology and Therapeutics* vol. 189 Preprint at https://doi.org/10.1016/j.pharmthera.2018.04.004 (2018).

100. Le, D. T. *et al.* PD-1 Blockade in Tumors with Mismatch-Repair Deficiency. *New England Journal of Medicine* **372**, (2015).

101. Li, K., Luo, H., Huang, L., Luo, H. & Zhu, X. Microsatellite instability: A review of what the oncologist should know. *Cancer Cell International* vol. 20 Preprint at https://doi.org/10.1186/s12935-019-1091-8 (2020).

102. Suraweera, N. *et al.* Evaluation of tumor microsatellite instability using five quasimonomorphic mononucleotide repeats and pentaplex PCR. *Gastroenterology* **123**, (2002).

103. Liu, J. *et al.* MA03.05 DNA Damage Response (DDR) Gene Mutations and Correlation With Immunotherapy Response in NSCLC Patients. *Journal of Thoracic Oncology* **16**, (2021).

104. Rizzo, A. *et al.* Predictive Biomarkers for Checkpoint Inhibitor-Based Immunotherapy in Hepatocellular Carcinoma: Where Do We Stand? *Frontiers in Oncology* vol. 11 Preprint at https://doi.org/10.3389/fonc.2021.803133 (2021).

105. Bai, R., Lv, Z., Xu, D. & Cui, J. Predictive biomarkers for cancer immunotherapy with immune checkpoint inhibitors. *Biomarker Research* vol. 8 Preprint at https://doi.org/10.1186/s40364-020-00209-0 (2020).

106. Huang, Q., Lei, Y., Li, X., Guo, F. & Liu, M. A Highlight of the Mechanisms of Immune Checkpoint Blocker Resistance. *Frontiers in Cell and Developmental Biology* vol. 8 Preprint at https://doi.org/10.3389/fcell.2020.580140 (2020).

107. Ayers, M. *et al.* IFN-γ-related mRNA profile predicts clinical response to PD-1 blockade. *Journal of Clinical Investigation* **127**, (2017).

108. Jorgovanovic, D., Song, M., Wang, L. & Zhang, Y. Roles of IFN-γin tumor progression and regression: A review. *Biomarker Research* vol. 8 Preprint at https://doi.org/10.1186/s40364-020-00228-x (2020).

109. Fasano, M. *et al.* Translational insights and new therapeutic perspectives in head and neck tumors. *Biomedicines* vol. 9 Preprint at https://doi.org/10.3390/biomedicines9081045 (2021).

110. de Keukeleire, S. J. *et al.* Immuno-oncological biomarkers for squamous cell cancer of the head and neck: Current state of the art and future perspectives. *Cancers* vol. 13 Preprint at https://doi.org/10.3390/cancers13071714 (2021).

111. Gavrielatou, N., Doumas, S., Economopoulou, P., Foukas, P. G. & Psyrri, A. Biomarkers for immunotherapy response in head and neck cancer. *Cancer Treatment Reviews* vol. 84 Preprint at https://doi.org/10.1016/j.ctrv.2020.101977 (2020).

112. Lei, Y., Li, X., Huang, Q., Zheng, X. & Liu, M. Progress and Challenges of Predictive Biomarkers for Immune Checkpoint Blockade. *Frontiers in Oncology* vol. 11 Preprint at https://doi.org/10.3389/fonc.2021.617335 (2021).

113. Feng, B. & Hess, J. Immune-related mutational landscape and gene signatures: Prognostic value and therapeutic impact for head and neck cancer. *Cancers* vol. 13 Preprint at https://doi.org/10.3390/cancers13051162 (2021).

114. Eisenhauer, E. A. *et al.* New response evaluation criteria in solid tumours: Revised RECIST guideline (version 1.1). *Eur J Cancer* **45**, (2009).

115. Hendry, S. *et al.* Assessing Tumor-infiltrating Lymphocytes in Solid Tumors: A Practical Review for Pathologists and Proposal for a Standardized Method from the International Immunooncology Biomarkers Working Group: Part 1: Assessing the Host Immune Response, TILs in Invasive Breast Carcinoma and Ductal Carcinoma in Situ, Metastatic Tumor Deposits and Areas for Further Research. *Advances in Anatomic Pathology* vol. 24 Preprint at https://doi.org/10.1097/PAP.0000000000000162 (2017).

116. Dai, P. *et al.* Calibration-free NGS quantitation of mutations below 0.01% VAF. *Nat Commun* **12**, (2021).

117. Strom, S. P. Current practices and guidelines for clinical next-generation sequencing oncology testing. *Cancer Biology and Medicine* vol. 13 Preprint at https://doi.org/10.28092/j.issn.2095-3941.2016.0004 (2016).

118. He, M. M. *et al.* Variant Interpretation for Cancer (VIC): A computational tool for assessing clinical impacts of somatic variants. *Genome Med* **11**, (2019).

119. Auton, A. *et al.* A global reference for human genetic variation. *Nature* vol. 526 Preprint at https://doi.org/10.1038/nature15393 (2015).

120. Montgomery, N. D. *et al.* Identification of Germline Variants in Tumor Genomic Sequencing Analysis. *Journal of Molecular Diagnostics* vol. 20 Preprint at https://doi.org/10.1016/j.jmoldx.2017.09.008 (2018).

121. Li, M. M. *et al.* Standards and Guidelines for the Interpretation and Reporting of Sequence Variants in Cancer: A Joint Consensus Recommendation of the Association for Molecular Pathology, American Society of Clinical Oncology, and College of

American Pathologists. *Journal of Molecular Diagnostics* vol. 19 Preprint at https://doi.org/10.1016/j.jmoldx.2016.10.002 (2017).

122. Mayakonda, A., Lin, D. C., Assenov, Y., Plass, C. & Koeffler, H. P. Maftools: Efficient and comprehensive analysis of somatic variants in cancer. *Genome Res* **28**, (2018).

123. Ritchie, M. E. *et al.* Limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res* **43**, (2015).

124. Wilkinson, L. ggplot2: Elegant Graphics for Data Analysis by WICKHAM, H. *Biometrics* **67**, (2011).

125. Liberzon, A. *et al.* The Molecular Signatures Database Hallmark Gene Set Collection. *Cell Syst* **1**, (2015).

126. Aran, D., Hu, Z. & Butte, A. J. xCell: Digitally portraying the tissue cellular heterogeneity landscape. *Genome Biol* **18**, (2017).

127. Therneau, T. M. & Grambsch, P. M. The Cox Model BT - Modeling Survival Data: Extending the Cox Model. in *Statistics for Biology and Health* (2000).

128. Kassambara, A., Kosinski, M. & Biecek, P. survminer: Drawing Survival Curves using 'ggplot2'. R package version 0.4.8. *https://cran.r-project.org/web/packages/survminer/* Preprint at (2020).

129. Budczies, J. *et al.* Cutoff Finder: A Comprehensive and Straightforward Web Application Enabling Rapid Biomarker Cutoff Optimization. *PLoS One* **7**, (2012).

130. Hsiehchen, D. *et al.* DNA Repair Gene Mutations as Predictors of Immune Checkpoint Inhibitor Response beyond Tumor Mutation Burden. *Cell Rep Med* **1**, (2020).

131. Bi, G. *et al.* Multi-omics characterization and validation of invasiveness-related molecular features across multiple cancer types. *J Transl Med* **19**, (2021).

132. She, Y. *et al.* Immune-related gene signature for predicting the prognosis of head and neck squamous cell carcinoma. *Cancer Cell Int* **20**, (2020).

133. Qiang, W., Dai, Y., Xing, X. & Sun, X. Identification and validation of a prognostic signature and combination drug therapy for immunotherapy of head and neck squamous cell carcinoma. *Comput Struct Biotechnol J* **19**, (2021).

134. Charoentong, P. *et al.* Pan-cancer Immunogenomic Analyses Reveal Genotype-Immunophenotype Relationships and Predictors of Response to Checkpoint Blockade. *Cell Rep* **18**, (2017).

135. Bai, S. *et al.* Bioinformatic Analysis Reveals an Immune/Inflammatory-Related Risk Signature for Oral Cavity Squamous Cell Carcinoma. *J Oncol* **2019**, (2019).

136. Chen, Y. P. *et al.* Identification and validation of novel microenvironment-based immune molecular subgroups of head and neck squamous cell carcinoma: Implications for immunotherapy. *Annals of Oncology* **30**, (2019).

137. Fang, R. *et al.* A novel comprehensive immune-related gene signature as a promising survival predictor for the patients with head and neck squamous cell carcinoma. *Aging* **13**, (2021).

138. Hu, G. *et al.* Integrated Analysis of RNA-Binding Proteins Associated With the Prognosis and Immunosuppression in Squamous Cell Carcinoma of Head and Neck. *Front Genet* **11**, (2021).

139. Liu, Z. *et al.* Comprehensive Analysis of Myeloid Signature Genes in Head and Neck Squamous Cell Carcinoma to Predict the Prognosis and Immune Infiltration. *Front Immunol* **12**, (2021).

140. Li, C. W. *et al.* A 4 Gene-based Immune Signature Predicts Dedifferentiation and Immune Exhaustion in Thyroid Cancer. *Journal of Clinical Endocrinology and Metabolism* **106**, (2021).

141. Rooney, M. S., Shukla, S. A., Wu, C. J., Getz, G. & Hacohen, N. Molecular and genetic properties of tumors associated with local immune cytolytic activity. *Cell* **160**, (2015).

142. Liu, X. *et al.* Systematic Profiling of Immune Risk Model to Predict Survival and Immunotherapy Response in Head and Neck Squamous Cell Carcinoma. *Front Genet* **11**, (2020).

143. Wang, H., Liu, B. & Wei, J. Beta2-microglobulin(B2M) in cancer immunotherapies: Biological function, resistance and remedy. *Cancer Letters* vol. 517 Preprint at https://doi.org/10.1016/j.canlet.2021.06.008 (2021).

144. Kapoor, S. S. & Zaiss, D. M. W. Emerging role of egfr mutations in creating an immune suppressive tumour microenvironment. *Biomedicines* vol. 10 Preprint at https://doi.org/10.3390/biomedicines10010052 (2022).

145. Choudhry, H. *et al.* Prospects of IL-2 in Cancer Immunotherapy. *BioMed Research International* vol. 2018 Preprint at https://doi.org/10.1155/2018/9056173 (2018).

146. Silva-Filho, J. L., Caruso-Neves, C. & Pinheiro, A. A. S. IL-4: An important cytokine in determining the fate of T cells. *Biophysical Reviews* vol. 6 Preprint at https://doi.org/10.1007/s12551-013-0133-z (2014).

147. Wang, X. *et al.* Tumor cell-intrinsic PD-1 receptor is a tumor suppressor and mediates resistance to PD-1 blockade therapy. *Proc Natl Acad Sci U S A* **117**, (2020).

148. Montfort, A. *et al.* The TNF paradox in cancer progression and immunotherapy. *Frontiers in Immunology* vol. 10 Preprint at https://doi.org/10.3389/fimmu.2019.01818 (2019).

149. Salomon, B. L. *et al.* Tumor necrosis factor α and regulatory T cells in oncoimmunology. *Frontiers in Immunology* vol. 9 Preprint at https://doi.org/10.3389/fimmu.2018.00444 (2018).

150. Xie, R. *et al.* Functional defects of cancer-associated MDC1 mutations in DNA damage repair. *DNA Repair (Amst)* **114**, 103330 (2022).

151. Zhang, X., Hu, F., Liu, L. & Xu, B. Effect of silencing of mediator of DNA damage checkpoint protein 1 on the growth of oral squamous cell carcinoma in vitro and in vivo. *Eur J Oral Sci* **127**, (2019).

152. Dave, J. H., Vora, H. H., Ghosh, N. R. & Trivedi, T. I. Mediator of DNA damage checkpoint protein 1 (MDC1) as a prognostic marker for patients with oral squamous cell carcinoma. *Journal of Oral Pathology and Medicine* **46**, (2017).

153. Príncipe, C., de Sousa, I. J. D., Prazeres, H., Soares, P. & Lima, R. T. Lrp1b: A giant lost in cancer translation. *Pharmaceuticals* vol. 14 Preprint at https://doi.org/10.3390/ph14090836 (2021).

154. Jackson, J. T., Mulazzani, E., Nutt, S. L. & Masters, S. L. The role of PLCγ2 in immunological disorders, cancer, and neurodegeneration. *Journal of Biological Chemistry* vol. 297 Preprint at https://doi.org/10.1016/j.jbc.2021.100905 (2021).

155. Sayed, R. el & Blais, N. Immunotherapy in extensive-stage small cell lung cancer. *Current Oncology* vol. 28 Preprint at https://doi.org/10.3390/curroncol28050347 (2021).

156. Li, Z. *et al.* PLCG2 as a potential indicator of tumor microenvironment remodeling in soft tissue sarcoma. *Medicine* **100**, (2021).

157. Chan, J. M. *et al.* Signatures of plasticity, metastasis, and immunosuppression in an atlas of human small cell lung cancer. *Cancer Cell* **39**, (2021).

158. Dong, G. *et al.* ZFHX3 promotes the proliferation and tumor growth of er-positive breast cancer cells likely by enhancing stem-like features and MYC and TBX3 transcription. *Cancers (Basel)* **12**, (2020).

159. Fu, C. *et al.* The transcription factor ZFHX3 is crucial for the angiogenic function of hypoxia-inducible factor 1α in liver cancer cells. *Journal of Biological Chemistry* **295**, (2020).

160. Zhang, J. *et al.* ZFHX3 mutation as a protective biomarker for immune checkpoint blockade in non-small cell lung cancer. *Cancer Immunology, Immunotherapy* **70**, (2021).

161. Ackermann, A. & Brieger, A. The Role of Nonerythroid Spectrin αII in Cancer. *Journal of Oncology* vol. 2019 Preprint at https://doi.org/10.1155/2019/7079604 (2019).

162. Lambert, M. W. Nuclear alpha spectrin: Critical roles in DNA interstrand cross-link repair and genomic stability. *Exp Biol Med* **241**, (2016).

163. Chen, Z. G., Saba, N. F. & Teng, Y. The diverse functions of FAT1 in cancer progression: good, bad, or ugly? *J Exp Clin Cancer Res* **41**, 248 (2021).

164. Jiang, S. Tet2 at the interface between cancer and immunity. *Communications Biology* vol. 3 Preprint at https://doi.org/10.1038/s42003-020-01391-5 (2020).

165. Zhang, W. *et al.* Differential expression of epha5 protein in gastric carcinoma and its clinical significance. *Oncol Lett* **17**, (2019).

166. Zhang, X. The Expression Profile and Prognostic Values of EPHA Family Members in Breast Cancer. *Front Oncol* **11**, (2021).

167. Staquicini, F. I. *et al.* Receptor tyrosine kinase EphA5 is a functional molecular target in human lung cancer. *Journal of Biological Chemistry* **290**, (2015).

168. Wang, H. *et al.* PDL1 high expression without TP53, KEAP1 and EPHA5 mutations could better predict survival for patients with NSCLC receiving atezolizumab. *Lung Cancer* **151**, (2021).

169. Chen, Y. *et al.* Role of PRKDC in cancer initiation, progression, and treatment. *Cancer Cell International* vol. 21 Preprint at https://doi.org/10.1186/s12935-021-02229-8 (2021).

170. Yin, Y. *et al.* Emerging functions of PRKDC in the initiation and progression of cancer. *Tumori* vol. 107 Preprint at https://doi.org/10.1177/0300891620950472 (2021).

171. Zuo, C. *et al.* Polymorphisms in ERCC4 and ERCC5 and risk of cancers: Systematic research synopsis, meta-analysis, and epidemiological evidence OPEN ACCESS EDITED BY. *Front. Oncol* **12**, 951193 (2022).

172. Gajic, Z. Z., Deshpande, A., Legut, M., Imieliński, M. & Sanjana, N. E. Recurrent somatic mutations as predictors of immunotherapy response. doi:10.1038/s41467-022-31055-3.

173. To, K. K. W., Fong, W. & Cho, W. C. S. Immunotherapy in Treating EGFR-Mutant Lung Cancer: Current Challenges and New Strategies. *Frontiers in Oncology* vol. 11 Preprint at https://doi.org/10.3389/fonc.2021.635007 (2021).

174. Zaretsky, J. M. *et al.* Mutations Associated with Acquired Resistance to PD-1 Blockade in Melanoma. *New England Journal of Medicine* **375**, (2016).

175. Zhao, J. *et al.* Immune and genomic correlates of response to anti-PD-1 immunotherapy in glioblastoma. *Nat Med* **25**, (2019).

176. Locati, L. D. *et al.* Complete Response to Nivolumab in Recurrent/Metastatic HPV-Positive Head and Neck Squamous Cell Carcinoma Patient After Progressive Multifocal Leukoencephalopathy: A Case Report. *Front Oncol* **11**, (2022).

177. Economopoulou, P., de Bree, R., Kotsantis, I. & Psyrri, A. Diagnostic tumor markers in head and neck squamous cell carcinoma (HNSCC) in the clinical setting. *Frontiers in Oncology* vol. 9 Preprint at https://doi.org/10.3389/fonc.2019.00827 (2019).

178. Ferris, R. L. *et al.* Nivolumab for Recurrent Squamous-Cell Carcinoma of the Head and Neck. *New England Journal of Medicine* **375**, (2016).

179. de la Iglesia, J. v. *et al.* Effects of tobacco smoking on the tumor immune microenvironment in head and neck squamous cell carcinoma. *Clinical Cancer Research* **26**, (2020).

180. Dang, R. P. *et al.* Clinical outcomes in patients with recurrent or metastatic human papilloma virus-positive head and neck cancer. *Anticancer Res* **36**, (2016).

181. Wang, H. fan *et al.* The double-edged sword—how human papillomaviruses interact with immunity in head and neck cancer. *Frontiers in Immunology* vol. 10 Preprint at https://doi.org/10.3389/fimmu.2019.00653 (2019).

182. Mandal, R. *et al.* The head and neck cancer immune landscape and its immunotherapeutic implications. *JCI Insight* **1**, (2016).

183. Budhwani, M. *et al.* Immune-Inhibitory Gene Expression is Positively Correlated with Overall Immune Activity and Predicts Increased Survival Probability of Cervical and Head and Neck Cancer Patients. *Front Mol Biosci* **8**, (2021).

184. Julian, R., Savani, M. & Bauman, J. E. Immunotherapy approaches in hpv-associated head and neck cancer. *Cancers* vol. 13 Preprint at https://doi.org/10.3390/cancers13235889 (2021).

185. Fundytus, A., Booth, C. M. & Tannock, I. F. How low can you go? PD-L1 expression as a biomarker in trials of cancer immunotherapy. *Annals of Oncology* vol. 32 Preprint at https://doi.org/10.1016/j.annonc.2021.03.208 (2021).

186. Sankar, K. *et al.* Novel strategies for immuno-oncology breakthroughs with cell therapy. (2021) doi:10.1186/s40364-022-00378-0.

187. Patel, S. P. & Kurzrock, R. PD-L1 expression as a predictive biomarker in cancer immunotherapy. *Molecular Cancer Therapeutics* vol. 14 Preprint at https://doi.org/10.1158/1535-7163.MCT-14-0983 (2015).

188. Chan, T. A. *et al.* Development of tumor mutation burden as an immunotherapy biomarker: Utility for the oncology clinic. *Annals of Oncology* vol. 30 Preprint at https://doi.org/10.1093/annonc/mdy495 (2019).

189. Strickler, J. H., Hanks, B. A. & Khasraw, M. Tumor mutational burden as a predictor of immunotherapy response: Is more always better? *Clinical Cancer Research* vol. 27 Preprint at https://doi.org/10.1158/1078-0432.CCR-20-3054 (2021).

190. Addeo, A., Banna, G. L. & Weiss, G. J. Tumor Mutation Burden - From Hopes to Doubts. *JAMA Oncology* vol. 5 Preprint at https://doi.org/10.1001/jamaoncol.2019.0626 (2019).

191. Cristescu, R. *et al.* Tumor mutational burden predicts the efficacy of pembrolizumab monotherapy: A pan-tumor retrospective analysis of participants with advanced solid tumors. *J Immunother Cancer* **10**, (2022).

192. Yarchoan, M., Hopkins, A. & Jaffee, E. M. Tumor Mutational Burden and Response Rate to PD-1 Inhibition. *New England Journal of Medicine* **377**, (2017).

193. Cortes-Ciriano, I., Lee, S., Park, W. Y., Kim, T. M. & Park, P. J. A molecular portrait of microsatellite instability across multiple cancers. *Nat Commun* **8**, (2017).

194. Wang, Z. *et al.* Comutations in DNA damage response pathways serve as potential biomarkers for immune checkpoint blockade. *Cancer Res* **78**, (2018).

195. Wang, M., Yu, L., Wei, X. & Wei, Y. Role of tumor gene mutations in treatment response to immune checkpoint blockades. *Precis Clin Med* **2**, (2019).

196. McKean, W. B., Moser, J. C., Rimm, D. & Hu-Lieskovan, S. Biomarkers in Precision Cancer Immunotherapy: Promise and Challenges. *American Society of Clinical Oncology Educational Book* (2020) doi:10.1200/edbk_280571.

197. Seliger, B. *et al.* Immune escape mechanisms and their clinical relevance in head and neck squamous cell carcinoma. *International Journal of Molecular Sciences* vol. 21 Preprint at https://doi.org/10.3390/ijms21197032 (2020).

198. Xu, W. *et al.* Hypoxia activates Wnt/β-catenin signaling by regulating the expression of BCL9 in human hepatocellular carcinoma. *Sci Rep* **7**, (2017).

199. Gabay, M., Li, Y. & Felsher, D. W. MYC activation is a hallmark of cancer initiation and maintenance. *Cold Spring Harb Perspect Med* **4**, (2014).

200. King, R. J., Singh, P. K. & Mehla, K. The cholesterol pathway: impact on immunity and cancer. *Trends in Immunology* vol. 43 Preprint at https://doi.org/10.1016/j.it.2021.11.007 (2022).

201. Caruntu, A. *et al.* Persistent Changes of Peripheral Blood Lymphocyte Subsets in Patients with Oral Squamous Cell Carcinoma. *Healthcare (Switzerland)* **10**, (2022).

202. Charap, A. J. *et al.* Landscape of natural killer cell activity in head and neck squamous cell carcinoma. *Journal for ImmunoTherapy of Cancer* vol. 8 Preprint at https://doi.org/10.1136/jitc-2020-001523 (2020).

203. Jung, E. K. *et al.* Natural killer cells have a synergistic anti-tumor effect in combination with chemoradiotherapy against head and neck cancer. *Cytotherapy* **24**, 905–915 (2022).

204. Chi, H. *et al.* Natural killer cell-related prognosis signature characterizes immune landscape and predicts prognosis of HNSCC. *Front Immunol* **0**, 5879 (2022).

205. Gaździcka, J., Gołąbek, K., Strzelczyk, J. K. & Ostrowska, Z. Epigenetic Modifications in Head and Neck Cancer. *Biochemical Genetics* vol. 58 Preprint at https://doi.org/10.1007/s10528-019-09941-1 (2020).

206. Romanowska, K., Sobecka, A., Rawłuszko-Wieczorek, A. A., Suchorska, W. M. & Golusiński, W. Head and neck squamous cell carcinoma: Epigenetic landscape. *Diagnostics* vol. 11 Preprint at https://doi.org/10.3390/diagnostics11010034 (2021).

207. Gougousis, S. *et al.* Epigenetic editing and tumor-dependent immunosuppressive signaling in head and neck malignancies. *Oncol Lett* **23**, (2022).

208. Chen, D. S. & Mellman, I. Elements of cancer immunity and the cancer-immune set point. *Nature* vol. 541 Preprint at https://doi.org/10.1038/nature21349 (2017).

209. Brock, G., Castellanos-Rizaldos, E., Hu, L., Coticchia, C. & Skog, J. Liquid biopsy for cancer screening, patient stratification and monitoring. *Translational Cancer Research* vol. 4 Preprint at https://doi.org/10.3978/j.issn.2218-676X.2015.06.05 (2015).

210. Kim, E. S. *et al.* Gefitinib versus docetaxel in previously treated non-small-cell lung cancer (INTEREST): a randomised phase III trial. *The Lancet* **372**, (2008).

211. Bossi, P. *et al.* A randomized, phase 2 study of cetuximab plus cisplatin with or without paclitaxel for the first-line treatment of patients with recurrent and/or metastatic squamous cell carcinoma of the head and neck. *Annals of Oncology* **28**, (2017).

212. Nixon, A. B. *et al.* Peripheral immune-based biomarkers in cancer immunotherapy: can we realize their predictive potential? *Journal for ImmunoTherapy of Cancer* vol. 7 Preprint at https://doi.org/10.1186/s40425-019-0799-2 (2019).

213. Snow, A., Chen, D. & Lang, J. E. The current status of the clinical utility of liquid biopsies in cancer. *Expert Review of Molecular Diagnostics* vol. 19 Preprint at https://doi.org/10.1080/14737159.2019.1664290 (2019).

214. Saussez, S., Duray, A., Demoulin, S., Hubert, P. & Delvenne, P. Immune suppression in head and neck cancers: A review. *Clinical and Developmental Immunology* vol. 2010 Preprint at https://doi.org/10.1155/2010/701657 (2010).

215. Bhardwaj, N. Harnessing the immune system to treat cancer. *Journal of Clinical Investigation* vol. 117 Preprint at https://doi.org/10.1172/JCI32136 (2007).

216. Desrichard, A. *et al.* Tobacco smoking-associated alterations in the immune microenvironment of squamous cell carcinomas. *J Natl Cancer Inst* **110**, (2018).

217. Boreel, D. F., Span, P. N., Heskamp, S., Adema, G. J. & Bussink, J. Targeting oxidative phosphorylation to increase the efficacy of radio- And immune-combination therapy. *Clinical Cancer Research* vol. 27 Preprint at https://doi.org/10.1158/1078-0432.CCR-20-3913 (2021).

218. Zhou, Y. *et al.* Wnt signaling pathway in cancer immunotherapy. *Cancer Letters* vol. 525 Preprint at https://doi.org/10.1016/j.canlet.2021.10.034 (2022).