RÉSEAU EUROPÉEN **DE RECHERCHE TRANSLATIONNELLE ET D'INNOVATION EN ONCOLOGIE**

ONCONET SUDOE

USE OF DIAGNOSTIC AND TREATMENT RESPONSE BIOMARKERS IN THE SUDOE AREA

July 2019

www.onconet-sudoe.eu

TABLE OF CONTENTS

1. Background and tumours analysed

Cancer is a generic term for a large group of diseases that can affect any part of the body. Other terms used are malignant tumours and neoplasms. One defining feature of cancer is the rapid creation of abnormal cells that grow beyond their usual boundaries, and which can then invade adjoining parts of the body and spread to other organs, the latter process is referred to as metastasizing. Metastases are a major cause of death from cancer [1].

1.1 The problem

Cancer is a leading cause of death worldwide, accounting for an estimated 9.6 million deaths in 2018. The most common cancers are:

- Lung (2.09 million cases)
- Breast (2.09 million cases)
- Colorectal (1.80 million cases)
- Prostate (1.28 million cases)
- Skin cancer (non-melanoma) (1.04 million cases)
- Stomach (1.03 million cases)

The most common causes of cancer death are cancers of:

- Lung (1.76 million deaths)
- Colorectal (862 000 deaths)
- Stomach (783 000 deaths)
- Liver (782 000 deaths)
- Breast (627 000 deaths)

CANCER SITE	NO. OF NEW CASES (% OF ALL SITES)	NO. OF DEATHS (% OF ALL SITES)
Lung	2,093,876 (11.6)	1,761,007 (18.4)
Breast	2,088,849 (11.6)	626,679 (6.6)
Prostate	1,276,106 (7.1)	358,989 (3.8)
Colon	1,096,601 (6.1)	551,269 (5.8)
Nonmelanoma of skin	1,042,056 (5.8)	65,155 (0.7)
Stomach	1,033,701 (5.7)	782,685 (8.2)
Liver	841,080 (4.7)	781,631 (8.2)
Rectum	704,376 (3.9)	310,394 (3.2)
Esophagus	572,034 (3.2)	508,585 (5.3)
Cervix uteri	569,847 (3.2)	311,365 (3.3)

Table 1. Top 10 most common cancer types

Table 1 shows the top 10 cancer types for estimated cases and deaths worldwide for men and women, combined and separately, with NMSC included within the *other* category. For both sexes combined, lung cancer is the most commonly diagnosed cancer (11.6% of the total cases) and the leading cause of cancer death (18.4% of the total cancer deaths), closely followed by female breast cancer (11.6%), colorectal cancer (10.2%), and prostate cancer (7.1%) for incidence and colorectal cancer (9.2%), stomach cancer (8.2%), and liver cancer (8.2%) for mortality. By sex, lung cancer is the most commonly diagnosed cancer and the leading cause of cancer death in males, followed by prostate and colorectal cancer for incidence, and liver and stomach cancer for mortality. Among females, breast cancer is the most commonly diagnosed cancer and the leading cause of cancer death, followed by colorectal and lung cancer for incidence, and vice versa for mortality; cervical cancer ranks fourth for both incidence and mortality. Overall, the top 10 cancer types account for over 65% of newly diagnosed cancer cases and deaths [2].

1.2 Cancer of unknown primary

Almost one in every three patients with cancer has distant metastases at the time of clinical diagnosis. In most cases, the primary tumour and the metastases are identified concomitantly, but for some patients, the primary lesion cannot be found after the initial clinical assessment. In these cases, the diagnosis of cancer of unknown primary (CUP) is made, a clinical situation quite difficult to manage due to the absence of a standard-of-care for the initial therapeutic approach. Cancers of unknown primary (CUPs) represent up to

150,000 new cases diagnosed each year in the United States and the European Union, but the number may increase to 400,000. The diagnosis of CUPs is mainly done by immunohistopathology. A detailed pathologic examination of biopsied tissue is mandatory and typically consists of hematoxylin-and-eosin staining and immunohistochemical tests.

1.2.1 Immunohistochemical Tests

Immunohistochemical markers, usually peroxidase-labeled antibodies against specific tumour antigens, are helpful in determining the tumour lineage. PSA and thyroglobulin (to detect prostate and thyroid cancers, respectively) are the most specific of the currently available markers; however, prostate and thyroid cancers rarely present as CUP. Also, no test is 100% specific, including the PSA test, which can be positive in patients with salivary gland carcinoma. Communication between the pathologist and the clinician is essential to a correct diagnosis and cannot be replaced by a battery of stains.

There are 20 known subtypes of cytokeratin (CK) intermediate filaments, all of which have different molecular weights and levels of expression in different cell types and cancers. Monoclonal antibodies to specific CK subtypes have been used to help classify tumours according to their site of origin; the most commonly used CK stains in CUP adenocarcinoma cases are CK 7 and 20. CK 7 is expressed in upper gastrointestinal tract tumours, cholangiocarcinoma, and pancreas, lung, ovary, endometrium, and breast cancers, whereas CK 20 is normally expressed in the lower gastrointestinal epithelium, urothelium, and Merkel cells. The CK 20+/CK 7- phenotype suggests a colon primary tumour; 75%–95% of colon tumours show this pattern of staining. Moreover, as it can be seen in Table 2 and 3 and Figure 1, CK 20−/CK 7+ is found in several cancer types, such as lung, breast, ovarian, and endometrial cancers. Cholangiocarcinoma and pancreatic cancer can be CK 20−/CK 7+ or CK 7+ with focal positivity for CK 20. Eighty-five percent of lung cancers are positive for CK 7, and the use of thyroid transcription factor-1 (TTF-1) and surfactant apoprotein can further help distinguish lung primary tumours from other CK 7+ tumours. Approximately 68% of lung adenocarcinomas and 25% of squamous cell lung cancers stain positive for TTF-1 [3,4].

CK7+/CK20−	CK7+/CK20+	CK7−/CK20+	CK7−/CK20−
Breast carcinoma Lung adenocarcinoma Endometrial adenocarcinoma Endocervical adenocarcinoma Ovarian (serous) carcinoma Cholangiocarcinoma Small cell lung carcinoma Mesothelioma Thyroid carcinoma Salivary gland tumours Kidney (papillary) Urothelial carcinoma (subset) Pancreatic adenocarcinoma Gastric adenocarcinoma	Urothelial carcinoma Pancreatic adenocarcinoma Ovarian mucinous carcinoma Bladder adenocarcinoma Gastric adenocarcinoma Cholangiocarcinoma	Colorectal adenocarcinoma Merkel cell carcinoma Gastric adenocarcinoma	Prostate adenocarcinoma Renal (clear cells) Hepatocellular carcinoma Adrenocortical carcinoma Non-seminoma germ cell tumours Mesothelioma Small cell lung carcinoma Gastric adenocarcinoma

Table 2. Main primary origins of carcinomas of unknown primary site (CUPs) based on staining for CK7 and CK20

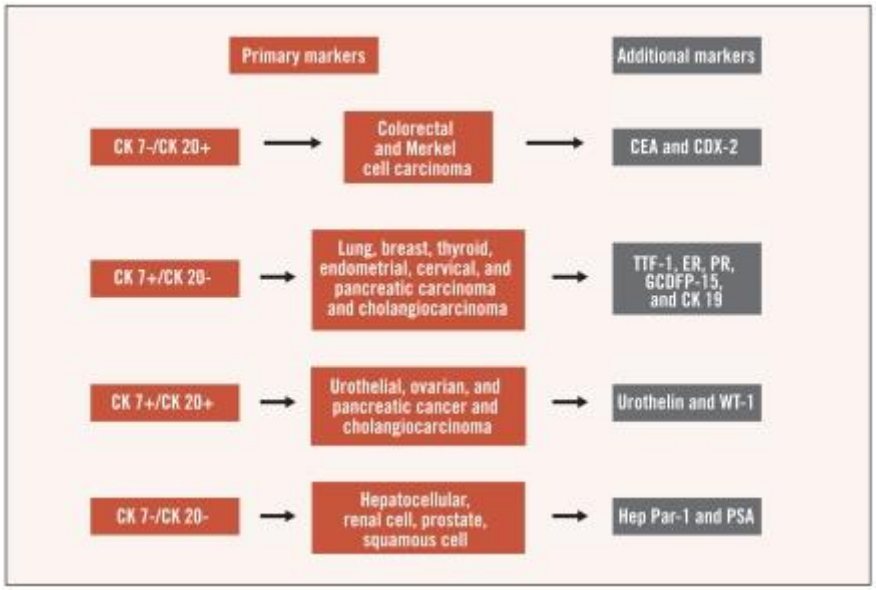

Figure 1. Immunohistochemical analysis of CUP based on cytokeratin (CK) status

Abbreviations: CDX = caudal-related homeobox; CEA = carcinoembryonic antigen; ER = estrogen receptor; GCDFP = gross cystic disease fluid protein; Hep Par = hepatocyte paraffin; PR = progesterone receptor; PSA = prostate specific antigen; WT = Wilms' tumor

Stain	Primary tumor
ER, PR, GCDFP-15, Her-2/neu	Breast cancer
TTF, CK 7, surfactant proteins	Lung cancer
Chromogranin, synaptophysin, neuron-specific enolase	Neuroendocrine tumor
β-Hcg, α-fetoprotein	Germ cell tumor
CK 7, CK 20, uroplakin III	Urothelial Malignancy
Calretinin	Mesothelioma
Hep Par-1	Hepatocellular carcinoma
CK 7, CK 20, CDX-2, CEA	Colorectal cancer

Table 3. Immunoperoxidase stains used in the differential diagnosis of CUP.

Abbreviations: CDX = caudal-related homeobox; CEA = carcinoembryonic antigen; CK = cytokeratin; ER = estrogen receptor; GCDFP = gross cystic disease fluid protein; Hcg = human chorionic gonadotropin; Hep Par = hepatocyte paraffin; PR = progesterone receptor; TTF = thyroid transcription factor.

1.3 Cancer pharmacogenomics, challenges in implementation, and patient-focused perspectives

Cancer pharmacogenomics is an evolving landscape and has the potential to significantly impact cancer care and precision medicine. While germline DNA is useful in understanding the pharmacokinetic and pharmacodynamics disposition of a drug, somatic DNA is particularly useful in identifying drug targets and predicting drug response. Molecular profiling of somatic DNA has resulted in the current breadth of targeted therapies available, expanding the armamentarium to battle cancer.

The transition from development of standard cytotoxic chemotherapies to highly targeted agents and immunotherapies has resulted in the current breadth of treatment options available. Further, increased numbers of targeted therapies are receiving accelerated drug approval alongside companion diagnostic assays, which are critical in identifying predictive biomarkers that allow for a personalized approach to therapy selection. A highly focused attack on targetable driver mutations has not only resulted in superior response rates and overall survival (OS) compared to traditional, non-targeted chemotherapy but has also

allowed for more rapid time to drug approval, ensuring timely access of life-prolonging drugs to cancer patients in dire need of more options.

Biomarkers can be categorized into two broad types: prognostic and predictive. A prognostic biomarker is a marker, or measurable trait, that provides information on the likely course of cancer, including aggressiveness of disease, regardless of treatment. Widely recognized examples include gene expression arrays such as the 70-gene profile MammaPrint or 21-gene profile Oncotype Dx for estrogen/progesterone receptor-positive, lymph node-negative breast cancer and microsatellite instability (MSI) in colorectal cancer patients. MammaPrint (from Agendia, Inc., Irvine, CA, USA) and Oncotype Dx assist in determining the risk of breast cancer recurrence in women with early stage breast cancer and provide guidance as to which high-risk patients may require additional chemotherapy. While MSI and mutations within DNA repair genes can result in increased risk of developing colorectal cancer (e.g. Lynch Syndrome), MSI-high (MSI-H) colorectal tumours also indicate a favourable prognosis compared to microsatellite stable/low-frequency MSI (MSS/MSI-L) tumours, independently of chemotherapy, in local and advanced colorectal cancer.

A predictive biomarker is a marker, or measurable trait, that can be used to identify patients most likely to benefit from treatment and/or those predisposed to toxicity. Examples of clinically relevant germline and somatic predictive biomarkers for drug response/toxicity are discussed in the next section of the paper and are the primary focus of this review. Notably, some biomarkers may be characterized as both prognostic and predictive within the same tumour type, such as overexpression of HER2 in breast cancer, which without chemotherapy is considered a poor prognostic biomarker resulting in an aggressive phenotype; however, with the development of therapies targeting HER2 (e.g. trastuzumab), this biomarker is also considered a positive predictive biomarker for therapy response.

The following table, reported by J. N. Patel [5], provides a summary of cancer therapies with pharmacogenomic information in the Food and Drug Administration (FDA)-approved drug label, and the figure provides an illustration of clinically relevant somatic mutations and drug targets in cancer.

Disease	Biomarker	Therapy	Frequency
Breast	HER2	Trastuzumab, lapatinib, pertuzumab, ado-trastuzumab emtansine	20%
	ESR1	Exemestane, letrozole, anastrozole, fulvestrant, tamoxifen,	60%
Colorectal	KRAS	Cetuximab, panitumumab	35%–40%
	EGFR	Cetuximab, panitumumab	35%–45%
	DPYD	5-Fluorouracil, capecitabine	<5%
	UGT1A1	Irinotecan	30%
Lung	ALK	Crizotinib, ceritinib	5%–7%
	EGFR	Erlotinib, gefitinib, afatinib, osimertinib	15%–20%
Melanoma	BRAF	Vemurafenib, dabrafenib, trametinib	50%–60%
Acute promyelocytic leukemia	PML-RARα	Arsenic trioxide, tretinion	>95%
Chronic myeloid leukemia	BCR-ABL	Imatinib, dasatinib, nilotinib, bosutinib, ponatinib, omacetaxine mepesuccinate	>95%
	UGT1A1	Nilotinib	30%
Cutaneous T-cell lymphoma	CD-25/IL2RA	Denileukin diftitox	75%
Chronic lymphocytic leukemia (CLL)	del(17p)	Ibrutinib	3%–8% at diagnosis; up to 30% in refractory CLL
	CD20/MS4A1	Obinutuzumab, rituximab	25%
Acute lymphocytic leukemia	TPMT	6-Mercaptopurine, thioguanine	<5%
Non-Hodgkin's lymphoma	CD20/MS4A1	Rituximab, tositumomab	>90%

Note: Data from http://www.fda.gov/drugs/scienceresearch/researchareas/pharmacogenetics/ucm083378.htm.[87]

Table 4. Summary of oncology pharmacogenomic biomarkers in FDA drug labelling

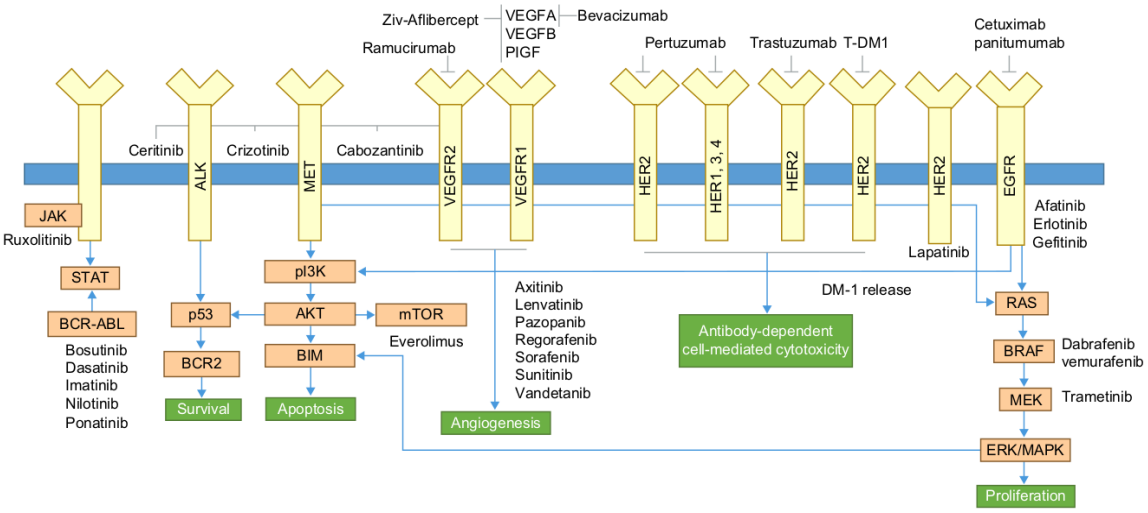

Figure 2. Summary of somatic cancer biomarkers and targeted therapies

Figure 2 depicts examples of key signalling pathways and downstream effects of mutations within somatic biomarkers, and their respective targeted therapies [5].

2. Colorectal cancer

2.1 Diagnosis

As shown above, colorectal cancer (CRC) is ranked as one of the most common types of cancer. It is clear that CRC is a rather heterogeneous disease by means of its various clinical manifestations, biological behaviour and in-tumour variety of mutations making it a true challenge, not only to detect in an early stage, but also to treat or even manage in the long term. Nowadays, it is evident that CRC is a multifactorial/polygenic disease arising both due to epigenetic as well as genetic manifestations occurring in a number of genes with an unparalleled role for the maintenance of normal cellular homeostasis.

Currently, the diagnosis of colon cancer is based on manual examination of histopathological images by a pathologist. Cancer disrupts the normal control mechanisms of cell proliferation and differentiation. This affects the structure, appearance, and activity of the cell nuclei structures within the tissue. Therefore, feature information extracted from cell nuclei structures may be utilized to detect cancer tissue within whole slide hematoxylin and eosin (H&E) stains.

Figure 3 shows the comparison of haematoxylin and eosin (H&E) stained tissue sections of normal and dysplastic colorectal tissues with magnification of 100 × , in which Fig. 3(a) is H&E staining of colorectal carcinoma tissue and Fig. 3(b) is H&E staining of normal colorectal tissue. It is clearly showed from Fig. 3(a) that there is a loss of normal glandular architecture for adenocarcinoma of colon as well as the marked nuclear atypia with prominent nucleoli and a high nucleus-to-cytoplasm ratio [6].

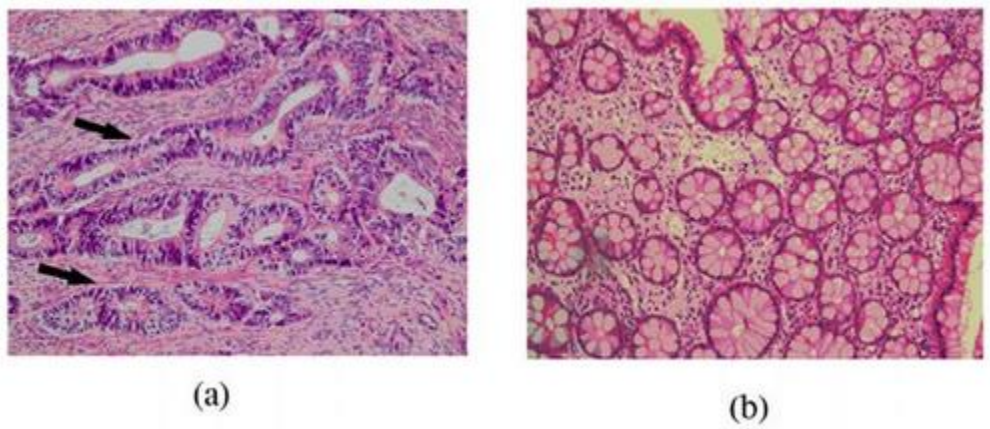

Figure 3. Comparison of haematoxylin and eosin (H&E) stained normal and carcinoma colorectal tissue

CRC researchers focus their research on innovative ideas to identify molecular markers for the development of highly accurate, non-invasive screening tests for CRC in the hope of

increasing the compliance of the population and to decrease potential unwanted side-effects which accompany the more invasive techniques. Several molecular classes have been tested for their potential use in CRC screening:

A. Cytokeratins (CKs)

CKs, proteins expressed by epithelial cells, are members of the intermediate filaments family along with vimentin, desmin, neuro-filament, and glial-filamen. Numerous studies have attempted to identify a possible expression pattern of CKs and connect it either with the organ of origin (in order to determine whether it is a primary CRC) or with tumour progression. However, as more and more studies are conducted it is becoming clear that such a connection is not likely to be identified in the near future. To begin with, CK7 and CK20 are helpful when the clinician needs to differentiate metastases from CRC, which are usually CK7−/CK20+, from other tumours. CK20 almost selectively stains the normal gland cells of the colonic mucosa and Merkel cells while its expression is rarely may be seen in the urothelium or other mucosas. By contrast, CK7 is usually expressed in urinary bladder and female genital tract epithelia, mesothelium, normal lung, and, rarely, it may be observed in gastric and intestinal normal glands. However, the majority of researchers agree that it is not found in normal colonic mucosa. Based on these findings, the immunophenotype CK7/CK20 is used as a routine in order to differentiate CK20-expressing metastasis of colorectal adenocarcinomas from lung, ovarian or bladder carcinomas, which are usually stained with CK7. However, it is reported that non-neoplastic colonic mucosa proximal to the rectum exhibits a CK7−/CK20+ phenotype, as is the case for 90% of CRCs. When CK17 is included in the diagnostic panel, the efficacy of the test is improved as less than 10% of CRCs express CK17 in contrast to other carcinomas that are more often positive for CK17 (including stomach, endometrium and urine bladder). In addition, pancreatic ductal carcinomas are consistently positive and a number of carcinomas from other sites, may exhibit CK17 expression. Furthermore, when CRC progression is studied, CK20 and CK7 can be useful. Results indicate that advanced CRCs were more often CK20+/CK7+ compared to early-stage cancers, which were predominantly CK20+/CK7−. Thus, CK7 expression may be a differentiating marker for the progression of CRC [6].

B. Caudal type homeobox 2 (CDX2)

CDX2 is a transcription factor encoded by CDX2 gene, a member of the caudal subgroup of homeobox genes. Its main role is to ensure maintenance of a cellular intestinal phenotype during the in utero and ex utero life. CDX2 presents strong expression patterns in epithelia of the normal small intestine, appendix, colon, and rectum as well as in the pancreatic centroacinar and interacinar ductal cells. It is revealed that loss of CDX2 may give birth to human CRC. CRCs, beside those exhibiting MSI, are consistently CDX2-positive. In fact, a quite interesting research recently investigated the effect of restoration of CDX2 expression

on colon cancer cell viability, colony formation, cell cycle distribution, apoptosis, invasion ability and xenograft tumour growth in nude mice. According to the researchers, CDX2 upregulation significantly reduced the life span and inhibited colony formation, and the invasion and migration ability of LoVo cells. Moreover, it was able to induce cell cycle arrest and apoptosis in vitro, especially under hypoxic culture conditions. According to data from histological studies, expression patterns of CDX2 are found in a variety of neoplastic tissues such as adenocarcinomas that exhibit intestinal-type differentiation, including adenocarcinomas of the gastroesophageal junction, bladder, urachus, small bowel, pancreas, appendix, and ovary [6].

2.2 Treatment

Hereditary nonpolyposis colorectal cancer (HNPCC), also known as Lynch syndrome, is a common autosomal dominant syndrome characterized by early age at onset, neoplastic lesions, and microsatellite instability (MSI). Because cancers with MSI account for approximately 15% of all colorectal cancers and because of the need for a better understanding of the clinical and histologic manifestations of HNPCC, the National Cancer Institute hosted an international workshop on HNPCC in 1996, which led to the development of the Bethesda Guidelines for the identification of individuals with HNPCC who should be tested for MSI [7].

The Revised Bethesda Guidelines for testing colorectal tumours for microsatellite instability (MSI) [8]:

1. Colorectal or uterine cancer diagnosed in a patient how is less than 50 years of age
2. Presence of synchronous, metachronous colorectal, or other HNPCC-associated tumours, * regardless of age.
3. Colorectal cancer with the MSI-H ** histology *** diagnosed in a patient who is less than 60 years of age^{+}.
4. Colorectal cancer diagnosed in one or more first-degree relatives with an HNPCC-related tumour, with one of the cancers being diagnosed under age 50 years.
5. Colorectal cancer diagnosed in two or more first- or second-degree relatives with HNPCC-related tumours, regardless of age.

* Hereditary nonpolyposis colorectal cancer (HNPCC)-related tumours include colorectal, endometrial, stomach, ovarian, pancreas, ureter and renal pelvis, biliary tract, and brain (usually glioblastoma as seen in Turcot syndrome) tumours, sebaceous gland adenomas and keratoacanthomas in Muir-Torre syndrome, and carcinoma of the small bowel.

** MSI-H - microsatellite instability–high in tumours refers to changes in two or more of the five National Cancer Institute-recommended panels of microsatellite markers

*** Presence of tumour infiltrating lymphocytes, Crohn's-like lymphocytic reaction, mucinous/signet-ring differentiation, or medullary growth pattern.

+ There was no consensus among the participants on whether to include the age criteria in guideline 3 above; participants voted to keep less than 60 years of age in the guidelines.

The tests are usually for MLH1 methylation, and for HNPCC germline mutations, there are three key DNA mismatch repair (MMR) genes (i.e., MSH2, MLH1 and, in attenuated cases, MSH6) that are responsible for these cancers. A few candidate genes (e.g., PMS2 and MLH3) are still awaiting additional validation regarding their role in the etiology of colorectal cancers with MSI [9].

Finally, in Table 5, it can be seen how ESMO guidelines summarize the management of hereditary colorectal cancer [10].

Syndrome	Diagnosis of index case (with cancer)			Management of the affected individual (with cancer)		Management of individuals at high risk (healthy mutation carriers or individuals at 50% risk of being mutation carrier)		
	Clinical	Molecular screening (tumour tissue)	Germline genetic testing (blood)	Treatment	Follow-up	Cancer risk	Surveillance	Germline genetic testing (blood)
Lynch	Amsterdam, Bethesda	MSI and/or IHC for MMR proteins	MLH1, MSH2 MSH6, PMS2	• Tumour resection • Discuss colectomy, especially in young patients	Yearly endoscopy of the remnant colon or rectum	High	• Colonoscopy q 1–2 years, starting age 25 (30 years in case of MSH6 or PMS2 mutations) • Annual pelvic examinations, transvaginal ultrasound, ca125, endometrial biopsy in females, starting age 30–35 years	Direct genetic testing of the mutation identified in the family
Familial CRC X	Amsterdam, Bethesda	No MMR deficiency	Unknown	As average population	As average population	Moderate only CRC	• Colonoscopy 1 3–5 years, starting 5–10 years before youngest case in the family.	None
FAP	Colonoscopy: >100 adenomas	none	APC	• Total or subtotal colectomy when adenomas occur • Endoscopic removal of duodenal adenomas	• After subtotal colectomy: rectal examination q 6–12 m • After total colectomy: pouch exam. q 1–2 years • Duodenoscopy from 6 months to 5 years according to Spigelman stage • Thyroid examination yearly	100%	• Flexible sigmoidoscopy q 2 years, starting age 12–14 years until diagnosis of adenomas • If no mutation identified in the family: Flexible sigmoidoscopy q 2 years until 40 years, then q 3–5 years until 50, then general population screening	APC
Attenuated FAP (aFAP)	Colonoscopy: a. 2 relatives 10–99 adenomas (>30 years of age) b. 1 relative of CRC patient with 10–99 adenomas (>30 years of age)		APC	• Total or subtotal colectomy when adenomas occur. • Endoscopic removal of duodenal adenomas	As above	High	• Colonoscopy q 2 years, starting age 18–20 years, lifelong in mutation carriers.	APC
MAP	As aFAP		MUTYH	As aFAP	As aFAP	High	As aFAP	MUTYH

APC, adenomatous-polyposis-coli; MSI, microsatellite instability; MMR, mismatch repair proteins; CRC, colorectal cancer; FAP, familial adenomatous polyposis; aFAP, attenuated FAP; MAP, MUTYH-associated polyposis.

Table 5. Management of hereditary colorectal cancer

3. Lung

3.1 Diagnosis

Primary lung cancer is the leading cause of cancer death, and the percentage of adenocarcinoma (AC) among lung cancers has been increasing gradually in recent decades.

Thyroid transcription factor 1 (TTF-1), also known as Nkx2.1 or thyroid-specific enhancer-binding protein, is a 38-kDa nuclear protein encoded by a gene located on chromosome 14q13. TTF1 is a master regulatory transcription factor for tissue-specific genes. TTF-1 is expressed in the thyroid, the lung and the diencephalon during embryogenesis. It has been reported that MUC5B is a target gene of TTF-1, which is involved in lung development and carcinogenesis, and strongly represses MUC5B expression it is also reported that no or low TTF-1 expression is detected in mucinous ACs. Because these ACs may express MUC5B, the diagnostic accuracy of lung AC should be increased by immunostaining with both of these factors.

TTF-1 is well known as a useful marker for lung ACs. TF-1 is most commonly used to distinguish primary lung adenocarcinoma from other sources that have metastasized to the lung. TTF-1 has also been a useful marker for differentiating primary lung adenocarcinoma from pleural mesothelioma. Positive TTF-1 staining by IHC has been described in as few as 25% to as many as 80% of primary adenocarcinomas, depending on the techniques used. However, the vast majority of squamous cell carcinomas and most extrapulmonary tumors (except for thyroid) lack TTF1 expression [11].

p40 is the best marker for diagnosing pulmonary squamous cell carcinoma. p40 staining yields high sensitivity as well as high specificity for distinguishing SQC from ADC, neuroendocrine carcinomas, and malignant mesothelioma.

The revolution in gene and other biomarker analysis has led to the identification of a number of biomarkers for which potentially active agents are already approved for other indications (e.g. crizotinib for *ALK* gene rearrangements).

These biomarkers include *ROS1* and *RET* gene rearrangements, *HER2* amplification and mutation, *BRAF* mutations and others. In addition, though no active agent has been clinically proven, *KRAS* mutation analysis has become common, given the widespread availability of validated tests. MET and PDL1 expression are other potential future biomarkers.

Numerous factors have been identified as potential predictive markers for specific chemotherapies in NSCLC, including ERCC1, BRCA1, RRM1 and thymidylate synthase (TS). For instance, low ERCC1 expression by IHC and mRNA have been associated with sensitivity to cisplatin-based chemotherapy, but IHC results have not been reproducible. Low BRCA1 mRNA expression also appears to confer sensitivity to the platinum agents, among other drugs, but confers resistance to the tubulin-targeting taxanes and vinca alkaloids.

A number of other markers are currently under evaluation for predicting benefit or resistance to chemotherapy. There is sufficient evidence to encourage research and suggest success, but at present no biomarker has been adequately evaluated for use in the clinic to select patients for any chemotherapeutic agent.

3.2 Treatment

In lung cancer, clinically relevant prognostic information is provided by staging. Staging forms the basis for the treatment options. Epidermal growth factor receptor (EGFR) and anaplastic lymphoma kinase are biomarkers used for prediction of chemotherapy and prediction of targeted treatment. Other driver biomarkers in lung cancer (point mutations and rearrangements in specific genes including *Her2*, *BRAF*, *NUT*, *MET*, *ROS1*, *DDR2*, *FGFR1*, *KRAS*, and *PTEN*) provide additional information for clinical decision making. More than 100 other lung cancer prognostic markers have been published [12].

3.2.1 PD-L1

PD-1,or 'programmed-death 1', was initially considered to be a molecule that regulates cell death, but is now identified as a key immune checkpoint inhibitory receptor, which is expressed on activated tumour-specific $CD4^+$ helper and $CD8^+$ killer T lymphocytes. PD-L1 or 'programmed-death ligand 1' (CD274), the main PD-1 ligand, is a transmembrane protein expressed on a variety of cell types, including antigen presenting cells, mainly dendritic cells and macrophages and constitutively expressed by non-lymphoid tissues including heart, lung and others. Binding of PD-L1 inhibits the function of activated T-cells, which is an important mechanism for negative feedback control of inflammation and autoimmunity in the peripheral effector phase of T-cell activation and identifies the PD-1/PD-L1 pathway as a significant immune response checkpoint. Tumour cells can co-opt this PD-1/PD-L1 regulatory mechanism. Tumour cells may express PD-L1, with subsequent PD-1 binding and inhibition of T-cell activation allowing cancer cells to evade immune attack. A number of trials validated that PD-L1 expression correlates with an increased response to PD-1 and PD-L1 immune checkpoint inhibitors. Currently, PD-L1 IHC 22C3 pharmDx (Dako) is the only FDA-approved companion diagnostic, which is used to select patients for treatment with pembrolizumab. The other three FDA-approved PD-L1 IHC assays are a complementary test that may provide physicians more information and inform patient dialogue when deciding treatment. These are PD-L1 IHC 28-8 pharmDx assay for nivolumab treatment, VENTANA PD-L1 IHC (SP142) assay for atezolizumab treatment and VENTANA PD-L1 IHC (SP263) assay for durvalumab.

4. Prostate

4.1 Diagnosis

The most important prognostic marker for prostate cancer (PCa), the Gleason score, is determined by pathologists on H&E stained

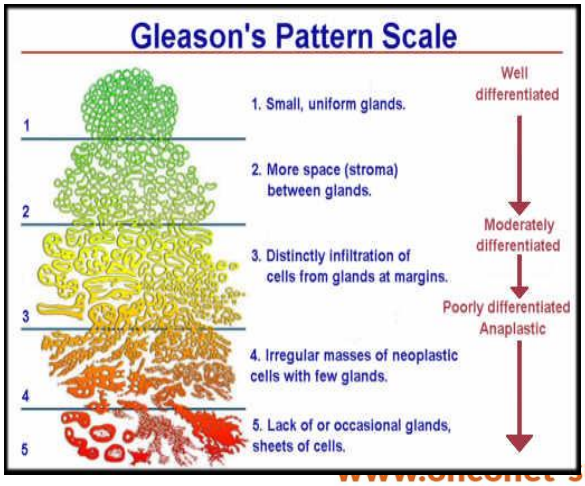

Figure 4. Gleason score [13]

specimens and is based on the architectural pattern of epithelial tissue. The Gleason Score ranges from 1-5 and describes how much the cancer from a biopsy looks like healthy tissue (lower score) or abnormal tissue (higher score). Most cancers score a grade of 3 or higher.

The new prostate grading system is an extension of the current Gleason grading scale for determining the stage of prostate cancer. This system is designed to provide a simplified and more accurate grading stratification system than the current Gleason Score. This new method is especially focused on better representing low grade disease to reduce unnecessary treatment of indolent prostate cancer. The new grading system subdivides prostate cancer into five categories using pathological characteristics as described in Table 6.

The use of serum prostate-specific antigen screening to facilitate early detection of prostate cancer has resulted in a dramatic increase in the number of prostate needle core biopsies which pathologists must examine. This has been accompanied by a strong increase in the number of biopsies with ambiguous lesions, and an unequivocal diagnosis of malignancy is difficult to render, especially in the case of limited foci or in small atypical acinar lesions.

TRADITIONAL GLEASON SCORE	NEW GRADING SYSTEM GROUP 1
GLEASON 3+3=6 Only individual discrete well-formed glands	GRADE 1
GLEASON 3+4=7 Predominantly well-formed glands with a lesser component of poorly-formed/fused/cribiform glands.	GRADE 2
GLEASON 4+3=7 Predominantly poorly-formed/ fused/cribriform glands with a lesser component of well-formed glands.	GRADE 3
GLEASON 4+4=8 Only poorly-formed/fused/cribriform glands or -Predominantly well-formed glands with a lesser component lacking or -Predominantly lacking glands with a lesser component of well-formed glands.	GRADE 4
GLEASON 9-10 Lacks gland formation (or with necrosis) with or without poorly-formed/fused/cribriform gland.	GRADE 5

Table 6. New Prostate Cancer Grading System [13]

When assessing small foci of atypical glands upon needle biopsy, the pathologist searches for differences between the benign glands and atypical glands in terms of usual morphological features and in such cases, immunohistochemical stains for basal cell markers such as 34ßE12 antibody or antibodies directed against cytokeratin 5 and 6 and more recently p63 may be a useful adjuvant to identify basal cells which are typically present in benign glands but absent in prostatic carcinoma.

However, several benign mimickers of prostate carcinoma, including atrophy, atypical adenomatous hyperplasia, nephrogenic adenoma can stain negatively with these antibodies and thus a negative basal cell marker immunostain alone does not exclude a diagnosis of benignancy. Alphamethyl-coenzyme-A-racemase (AMACR) a new sensitive marker of prostate carcinoma, can be useful in confirming ambiguous lesion suspected for malignancy. Although, as with any immunohistochemical studies, problems exist in terms of both

sensitivity and specificity. P504S, a cytoplasmic protein, is a highly sensitive and specific positive marker for prostate carcinoma.

In addition, the identification and characterization of the disease have become increasingly precise through improved risk stratification and advances in magnetic resonance and functional imaging, as well as from the emergence of new biomarkers.

Active surveillance (the serial monitoring for disease progression with the intent to cure) appears to be safe and has become the preferred approach for men with less-aggressive prostate cancer [14].

4.2 Treatment

Although there are known prognostic factors to guide management, there are no established predictive biomarkers to choose one particular treatment over another.

Surgery and radiation continue to be curative treatments for localized disease but have adverse effects such as urinary symptoms and sexual dysfunction that can negatively affect quality of life. For metastatic disease, chemotherapy as initial treatment now appears to extend survival compared with androgen deprivation therapy alone. New vaccines, hormonal therapeutics, and bone-targeting agents have demonstrated efficacy in men with metastatic prostate cancer resistant to traditional hormonal therapy.

Initial treatment with chemotherapy can improve survival compared with androgen deprivation therapy. Abiraterone, enzalutamide, and other agents can improve outcomes in men with metastatic prostate cancer resistant to traditional hormonal therapy.

Advanced disease progressing without a significant rise in PSA should be investigated for neuro-endocrine markers, using biopsy or blood analyses for neuron-specific enolase and/or chromogranin A.

5. Breast

5.1 Diagnosis

Hematoxylin and eosin (H&E) staining remains the most important and fundamental method for tumour histological examination in pathology. Histological features in H&E images are measured to evaluate tumour grade and prognosis. The Nottingham grading system (NGS) is recommended by the World Health Organization to obtain histological grade score. Histological tumour grade is based on the degree of differentiation of the tumour tissue. In breast cancer, it refers to the semi-quantitative evaluation of morphological characteristics

and is a relatively simple and low-cost method, requiring only adequately prepared hematoxylin-eosin-stained tumour tissue sections to be assessed by an appropriately trained pathologist using a standard protocol. NGS is based on the evaluation of three morphological features: (a) degree of tubule or gland formation, (b) nuclear pleomorphism, and (c) mitotic count (Figure 5) [15,16,17].

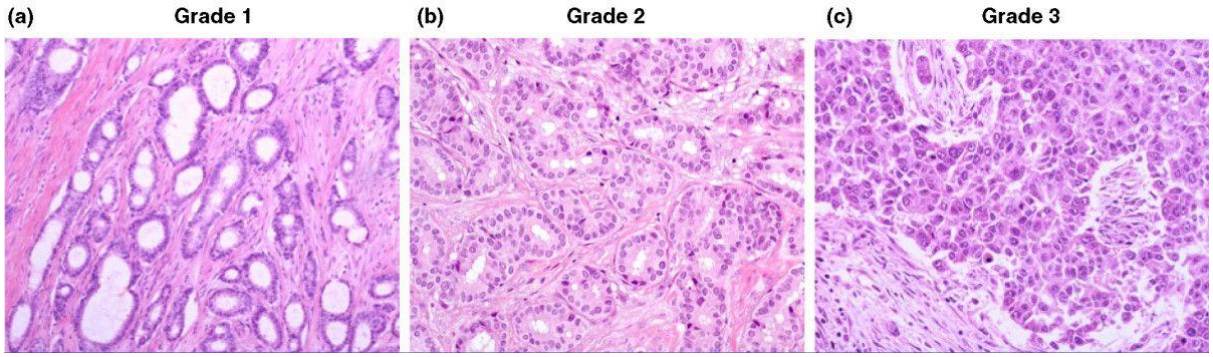

Figure 5. Histological grade of breast cancer as assessed by the Nottingham Grading System

Figure 5(a) shows a well-differentiated tumour (grade 1) that demonstrates high homology to the normal breast terminal duct lobular unit, tubule formation (>75%), a mild degree of nuclear pleomorphism, and low mitotic count. In Figure 5(b) a moderately differentiated tumour (grade 2) can be appreciated, whereas in Figure 5(c) the tumour is poorly differentiated (grade 3) and has a marked degree of cellular pleomorphism and frequent mitoses and no tubule formation (<10%).

5.2 Treatment

In breast cancer, biomarker analysis is routine practice. It originally began with testing for hormone receptor expression to guide tamoxifen therapy. This consensus statement revises and updates the recommendations for biomarkers use in the diagnosis and treatment of breast cancer, and is a joint initiative of the Spanish Society of Medical Oncology and the Spanish Society of Pathology. This expert group recommends determining in all cases of breast cancer the histologic grade and the alpha-estrogen receptor (ER), progesterone receptor, Ki-67 and HER2 status, in order to assist prognosis and establish therapeutic options, including hormone therapy, chemotherapy and anti-HER2 therapy. One of the four available genetic prognostic platforms (MammaPrint, Oncotype, Prosigna or EndoPredict) may be used in node-negative ER-positive patients to establish a prognostic category and decide with the patient whether adjuvant treatment may be limited to hormonal therapy.

5.2.1 Estrogen receptor and progesterone receptor

Expression of estrogen receptor (ER)-alpha is a favourable prognostic factor and strongly predictive of a response to hormone therapy. Approximately 30–40% of patients with ER-expressing advanced breast cancer will have an objective response to hormone treatment, and a further 20% of patients will achieve disease stabilisation. Detailed guidelines addressing methods for the immunohistochemical analysis of ERs and progesterone receptors (PRs) are available. In general, 70–75% of invasive breast carcinomas express ER-alpha. PRs are regulated by ER-alpha, so expression of PRs suggests that the oestrogen/ER-alpha pathway is functional.

Along with hormone receptors, HER2 is the most important prognostic and predictive marker in breast cancer. it has been known that breast cancers that overexpress HER2 represent highly aggressive biological subtype. The introduction of new-targeted anti-HER2 therapies, such as lapatinib, pertuzumaband trastuzumab emtansine (T-DM1), the last one administered with no requirement for simultaneous cytostatics, underlines the importance of identifying patients with HER2-positive breast cancer. Any invasive breast carcinoma should be tested for HER2 overexpression, along with ERs, PRs and Ki-67.

GATA transcription factors are zinc finger DNA binding proteins that activate transcription during development and cell differentiation. It was found that, when compared with normal tissues, GATA3 and TRPS1 were distinctly high expressed in BC patients and predict better survival in patients with BC. GATA3 is positively associated with ESR1, while TRPS1 is correlated with ERBB2 and might act as a potential modulator of chemosensitivity in breast tumour. GATA3 and TRPS1 are distinctive biomarkers and essential prognostic factors in BC.

<u>Summary of biomarkers consensus in breast cancer</u>

Conventional markers (recommended in all patients)

- ER-alpha
- PR
- HER2
- Ki-67
- Histological grade

Genetic platforms (recommended in patients with low risk of relapse)

- MammaPrint
- Oncotype DX
- Prosigna
- EndoPredict

6. Current and future perspectives

Several actions are being developed along the EU in order to make a step further in the optimization of cancer biomarkers used in the clinics. Few biomarkers progress from discovery to become validated tools or diagnostics. To bridge this gap, three European biomedical research infrastructures — EATRIS-ERIC (focused on translational medicine), BBMRI-ERIC (focused on biobanking) and ELIXIR (focused on data sharing) — are paving the way to developing and sharing best practices for biomarker validation. A COSME action named CLINIMARK: Good biomarker practice to increase the number of clinically validated biomarkers is being held since 2017, which includes National Health Institute Doutor Ricardo Jorge (INSA) at Lisbon (Portugal).

In the framework of Horizon 2020, 22 health projects are based on biomarkers for cancer and 4 of them are coordinated from Spanish entities. From the more than 96000 studies published worldwide on cancer biomarkers in the last 5 years, SUDOE area was represented in a proportion of them: France participated in more than 4200 papers, Spain in more than 3400 and Portugal in more than 700 studies in the field.

New approaches as the emergence of new imaging techniques and high-throughput molecular sequencing generates large amounts of global data, even per patient. These data can be converted to relevant clinical information allowing patient stratification and the deciphering of pathological mechanisms, thus advancing in the paradigm of personalized medicine. Among the molecular sequencing techniques, proteomics provides functional information about the activity of proteins, which is often crucial for the understanding of cell physiology. In addition, proteomic techniques will be key to push forward liquid biopsies for detection of different biomarkers.

Other growing areas include epigenetic markers. Epigenetic alterations are innovative cancer biomarkers owing to stability, frequency, reversibility and accessibility in body fluids, entailing great potential of assay development to assist in patient management. Several studies identified putative epigenetic cancer biomarkers, some of which have been commercialized. For example, EPICUP (IDIBELL and Ferrer, from Spain) is the first epigenetic diagnostic test based on the analysis of DNA methylation profiles and helps the oncologist to identify the primary tumour in patients with cancer of unknown origin (CUP). However, large multicentre validation studies are required to foster translation to the clinics [18].

Promising results from microRNA (miRNA) profiling and hypermethylation of selected genes have raised hopes of identifying new biomarkers. Some of these epigenetic biomarkers may be useful in risk assessment and for screening populations to identify who is likely to develop cancer [19].

Micro RNA (miRNAs) expression profile is being studied as part of the molecular phenotype of circulating tumour cells (CTCs), and has been associated with clinical outcome of patients with breast cancer. This can be used to enhance the prognostic accuracy of the CTC phenotype by incorporating miRNA into a combined prediction model [20].

In addition, there is increasing evidence that microRNAs (i.e., miR-34a, miR-143, miR-153, miR-27a, miR-218, and miR-520) play an essential role in tumorigenesis and chemotherapy resistance, by targeting various cellular and molecular pathways (i.e., PI3K/Akt/Wnt, EMT, p53, p21, and ATM) that are involved in the pathogenesis of colorectal cancer (CRC). Identifying the miRNAs that are involved in chemo-resistance, and their function, may help as a potential therapeutic option for treatment of CRC or as potential prognostic biomarker [21].

A crucial area in new cancer therapy in based on immunotherapy. In this sense, new microsatellite instability could be a predictive biomarkers for cancer immunotherapy [22].

As the field rapidly evolves, we must prioritise the development of biomarkers to guide the use of immunotherapies in the most appropriate patients. For instance, the predictive values of PD-L1 expressions for immunotherapy are currently under debate and need to be further developed [23].

7. Onconet Sudoe conclusions

Regarding ONCONET SUDOE diagnostic platforms, we need to be very cautions with conclusions due to the small sample of answers to the survey that was launched by the consortium (n=5). Surveys completed come from Portuguese hospitals: Centro Hospitalar e Universitário de Coimbra, Hospital de São José (Lisboa), Hospital de Braga, Hospital de Santo António (Porto) and Hospital de São João (Porto). They follow CAP (College of American Pathology) Protocols, WHO protocols and Royal College of Pathologists (UK) protocols. Practice in every hospital will depend on local organization and healthcare systems, but biomarkers used for diagnostic and therapy assessment purposes are consistent with ESMO guidelines.

Regarding the biomarkers and/or commercial kits for diagnostics used for colorectal cancer diagnosis, Portuguese hospitals mostly use haematoxylin/eosin staining of paraffin-embedded tissue of origin in first place. After this, they perform an immunohistochemistry against CK7/CK20 and CDX2 in order to assure the diagnostic. In order to assess the best therapeutic option, they perform staining against MLH1, MSH2, MSH6 and PMS2 if Bethesda criteria for testing colorectal tumours for microsatellite instability (MSI) are met. In some cases, staining against KRAS can be used for patients with advanced disease.

As for lung adenocarcinoma diagnosis, they commonly use biomarkers CK7 and TTF1, in specific cases Napsin A as well. For epidermoid carcinoma cases, they use p40 staining. For therapeutic purposes, the biomarker most widespread used is PDL1. As for adenocarcinoma, EGFR testing is performed and if negative, next step is testing for ALK and ROS1 rearrangements (by FISH).

Regarding the biomarkers used for prostate cancer diagnosis, Portuguese hospitals mostly use haematoxylin/eosin staining, as well as staining for p504s, 34bE12 and p40/p63 to confirm diagnosis, as well as PSA testing. For breast cancer diagnosis, after haematoxylin/eosin staining, diagnostic services perform testing for CK7, Estrogen receptor, Mammaglobin and/or GATA3 to confirm the diagnosis. In order to address the appropriate therapeutic strategy, levels of estrogen and progesterone receptor, HER2, EGDR, RAS, ROS1, Ki-67 and/or ck19 are analysed.

As stated above, multiplex platforms, including 'next-generation' technologies, for testing specific mutation panels are emerging as a cost- and resource-effective approach to simultaneous analysis of multiple potential targets. Although these technologies are not yet being routinely used, they hold a great promise for the future. In addition, diagnostic laboratories are tending to stablish a quality management system in order to have the appropriate ISO accreditations.

8. References

[1] "Cancer", Who.int, 2019. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/cancer. [Accessed: 17- May- 2019]

[2] F. Bray *et al.*, "Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries.," CA: a cancer journal for clinicians, vol. 68, p. 394–424, 2018.

[3] J. Sèlves *et al.*, "Immunohistochemistry for Diagnosis of Metastatic Carcinomas of Unknown Primary Site" Cancers, 2018.

[4] G. R. Varadhachary, "Carcinoma of unknown primary origin" Gastrointestinal cancer research (GCR), vol. 1, núm. 6, p. 229, 2007.

[5] J. N. Patel, "Cancer pharmacogenomics, challenges in implementation, and patient-focused perspectives" Pharmacogenomics and Personalized Medicine , vol. 9, núm. 1, pp. 65-77, 2016.

[6] T.K. Nikolouzakis et al., "Improving diagnosis, prognosis and prediction by using biomarkers in CRC patients (Review)", Oncology Reports, vol. 39, núm. 6, pp. 2455-2472, 2018.

[7] A. Umar et al., "Revised Bethesda Guidelines for hereditary nonpolyposis colorectal cancer (Lynch syndrome) and microsatellite instability", Journal of the National Cancer Institute, vol. 96, núm. 4, pp. 261-268, 2004.

[8] "Revised Bethesda guidelines for testing", Stanfordhealthcare.org, 2019. [Online]. Available: https://stanfordhealthcare.org/medical-conditions/cancer/lynch-syndrome/hnpcc-diagnosis/testing-guidelines.html. [Accessed: 22- May- 2019]

[9] A. Karahan et al., "Relationship between MLH-1, MSH-2, PMS-2,MSH-6 expression and clinicopathological features in colorectal cancer", International Journal of Clinical and Experimental Pathology, vol. 8, núm. 4, pp. 4044-4053, 2015.

[10] H. J. Schmoll et al., "ESMO Consensus Guidelines: Management of Patients with Colon and Rectal Cancer. A personalized approach to clinical decision making", Annals of Oncology, vol. 23, núm. 10, pp. 2479-2516, 2012.

[11] M. Majem, "SEOM clinical guidelines for the treatment of non-small cell lung cancer", Clinical & Translational Oncology, vol. 21, nº 1, pp. 3-17, 2018.

[12] E. Thunnissen, K. van der Oord and M. den Bakker, "Prognostic and predictive biomarkers in lung cancer (Review)", Springer Berlin Heidelberg, vol. 464, nº 3, pp. 347-358, 2014.

[13] Prostateconditions.org, 2019. [Online]. Available: https://www.prostateconditions.org/about-prostate-conditions/prostate-cancer/newly-diagnosed/gleason-score. [Accessed: 30- May- 2019]

[14] M. Litwin and H. Tan, "The Diagnosis and Treatment of Prostate Cancer (Review)", JAMA, vol. 317, nº 24, pp. 2532-2542, 2017.

[15] E. Rakha *et al.*, "Breast cancer prognostic classification in the molecular era: the role of histological grade", Breast Cancer Research, vol. 12, nº 4, p. 207, 2010.

[16] S. Petushi et al., "Large-scale computations on histology images reveal grade-differentiating parameters for breast cancer", BMC Medical Imaging, vol. 6, nº 1, p. 14, 2006.

[17] J. I. Chacón López-Muñiz et al., "SEOM clinical guidelines in advanced and recurrent breast cancer", Clinical & Translational Oncology, vol. 21, nº 1, pp. 31-45, 2019.

[18] P. Costa-Pinheiro , D. Montezuma , R. Henrique and C. Jerónimo , "Diagnostic and prognostic epigenetic biomarkers in cancer.", Future Medicine, vol. 7, nº 6, pp. 1003-1015, 2015.

[19] K. Banaudha and M. Verma, "Epigenetic biomarkers in liver cancer", Cancer Epigenetics, vol. 1238, pp. 65-76, 2015.

[20] W. Tan et al., "Incorporating MicroRNA into Molecular Phenotypes of Circulating Tumor Cells Enhances the Prognostic Accuracy for Patients with Metastatic Breast Cancer.", The Oncologist, 2019.

[21] R. Moradi Marjaneh et al., "MicroRNAs as potential therapeutic targets to predict responses to oxaliplatin in colorectal cancer: From basic evidence to therapeutic implication", IUBMB life, 2019.

[22] L. Chang, M. Chang, H. Chang and F. Chang, "Microsatellite Instability: A Predictive Biomarker for Cancer Immunotherapy", Applied Immunohistochemistry & Molecular Morphology, vol. 26, nº 2, pp. e15-e21, 2018.

[23] F. Teng, J. Yu, X. Meng and L. Kong, "Progress and challenges of predictive biomarkers of anti PD-1/PD-L1 immunotherapy: A systematic review", Cancer letters, vol. 414, pp. 166-173, 2018.

9. Acknowledgments

This report has been financed by Interreg Sudoe through the project ONCONET SUDOE, which seeks to connect researchers, care providers, biobanks and companies to set long-term relationships that foster innovation in the oncology sector.

The information shared in the document has been gathered through different surveys that were sent to the Biobanks and Diagnostic Platforms in the SUDOE area, as well as additional information that has been collected via phone calls or face-to-face meetings with experts. Despite the great efforts put in obtaining feedback from several players of the SUDOE area, only 5 entities filled in the whole survey regarding Biomarkers. These entities are: Centro Hospitalar e Universitário de Coimbra, Hospital de São José, Hospital de Braga, Hospital de Santo António and Hospital de São João. Their collaboration, as well as the collaboration of several experts and managers of the Onconet project has been crucial for the preparation of this report.