From Molecular Medicine and Surgery
Karolinska Institutet, Stockholm, Sweden

# HEREDITARY PREDISPOSITION AND PROGNOSTIC PREDICTION IN CANCER

Karin Wallander


Karolinska Institutet

Stockholm 2021

Published by Karolinska Institutet.
Printed by Universitetsservice US-AB, 2021

ISBN 978-91-8016-426-9
Cover illustration:
Crayfish from the lake Övingen, Småland, Sweden. Adapted by the author, from a photograph taken by Katarina Gyllensten. The Swedish word for crayfish, “kräfta”, is an old term for cancer.

# Hereditary predisposition and prognostic prediction in cancer

## THESIS FOR DOCTORAL DEGREE (Ph.D.)

By

**Karin Wallander**

*Principal Supervisor:*
Emma Tham
Karolinska Institutet
Department of Molecular Medicine and Surgery

*Co-supervisor(s):*
Kristina Lagerstedt Robinson
Karolinska Institutet
Department of Molecular Medicine and Surgery

Annika Lindblom
Karolinska Institutet
Department of Molecular Medicine and Surgery

Magnus Nordenskjöld
Karolinska Institutet
Department of Molecular Medicine and Surgery

Mats Lindblad
Karolinska Institutet
Department of Clinical Science, Intervention and Technology

*Opponent:*
Professor Fátima Carneiro
University of Porto
Department of Pathology

*Examination Board:*
Lucia Cavelier
Uppsala University
Department of Immunology, Genetics and Pathology

Maria Gustafsson Liljefors
Karolinska Institutet
Department of Clinical Science, Intervention and Technology

Wilhelm Graf
Uppsala University
Department of Surgical Sciences

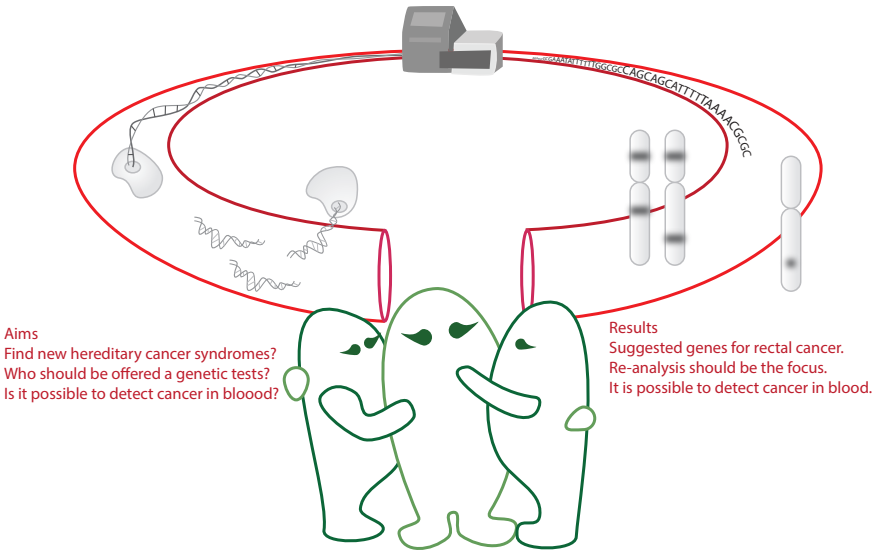
# POPULAR SCIENCE SUMMARY OF THE THESIS

Cancer is caused by changes in the genetic code, the DNA. Cells with cancer-causing genetic changes get characteristics such as more frequent division and efficient escape from the immune system defense. Hereditary cancer syndromes are caused by inherited genetic changes that increase the likelihood that cells will develop into a malignancy. We do not yet know the genetic mechanisms for all hereditary cancer syndromes. Further, we do not have a good enough biomarker for detecting and surveying cancer, whether it is inherited or not. Cell-free DNA is ~~DNA~~ released into the blood, and it might be useful as a biomarker for cancer.

Study I and Study II aimed at increasing the knowledge about hereditary cancer syndromes. We studied a family wherein many members had rectal cancer, and found six potential genetic risk factors. Also, we studied individuals with three or more different tumours, to see if they would benefit from the most thorough genetic test we can offer in clinical practice (whole-genome and whole-exome sequencing). Our conclusion was that it may be of value to re-evaluate the genetic information already available for these patients.

Study III aimed at increasing the knowledge about liquid biopsies (analysis of cell-free DNA) as a test for cancer. The lab workflow was based on the clinical non-invasive prenatal tests, and the data analyses were a further development from that. We analysed the genetic information with two different programs. We found that it was possible to detect DNA from stomach and oesophageal cancer in blood samples. One of the programs enabled detection of cancer-associated genetic aberrations in more samples than the other. This program is today used in prenatal diagnostics at our centre. Since we already have the clinical pipeline up and running, it is likely this specific cell-free DNA test can be of value in the future.

With these results, we believe our studies help improving the care we offer at clinical genetics departments.



# POPULÄRVETENSKAPLIG SAMMANFATTNING

Cancer beror på förändringar i tumörcellers arvsmassa (DNA). DNA-förändringar kan göra att cellerna till exempel delar sig oftare eller överlever även vid syrebrist. Ärftliga cancersyndrom orsakas av DNA-förändringar som ärvs inom familjer. Vi känner inte till alla ärftliga cancersyndrom än och vi vet inte alla genetiska orsaker till de vi känner till. Det finns idag inte någon tillräckligt bra biomarkör för att upptäcka, karaktärisera eller övervaka cancer, oavsett om den är nedärvd eller inte. Cell-fritt DNA (DNA som cirkulerar i blodkärlen) skulle kunna vara en lättillgänglig biomarkör för cancer.

Målet i Delstudie I och II var att öka kunskapen om ärftliga cancersyndrom. I en familj med många personer som fått ändtarmscancer, som vi undersökte i Delstudie I, hittade vi sex genetiska förändringar som skulle kunna ge ökad risk för den cancerformen. Det har föreslagits att personer som fått tre eller fler olika tumörer borde erbjudas den mest heltäckande genetiska analysen vi har på kliniken (hel-genom/hel-exom-sekvensering). Resultatet från Delstudie II visade att man som alternativ kan överväga att om-analysera resultaten från genetiska test som redan gjorts.

Målet i Delstudie III var att öka kunskapen om hur man kan använda cell-fritt DNA som en biomarkör för cancer. Samma rutiner för blodprovstagning och bearbetning av blodet som vi redan använder vid en viss typ av fosterdiagnostik kunde utnyttjas. Data-analysen var en vidareutveckling av den som användes för fosterdiagnostiken. Två olika program testades för att få fram listor på potentiellt cancer-associerade genetiska varianter i blodproven. Vi visade att det var möjligt att hitta cancer-förändringar med vår metod och att ett av programmen gjorde det mer effektivt än det andra. Eftersom vi redan har en klinisk rutin som fungerar kan vi i framtiden erbjuda denna analys av cell-fritt DNA som test också till cancerpatienter.

Med dessa resultat kan våra studier hjälpa till att förbättra vården vid sjukhusens kliniskt genetiska mottagningar och laboratorier.

# ABSTRACT

Cancer is a genetic disease. The cancer cells evolve clonally through changes of their genome, gaining malignant characteristics. In hereditary cancer syndromes, a genetic aberration which puts the cells one step closer to being malignant is inherited and present throughout whole body. Not all genes associated with hereditary cancer syndromes have been discovered, and the spectrum of cancers connected to each cancer-associated gene is not yet known. When it is feasible, the genetic changes within the tumour can be analysed by performing a tissue biopsy. Alternatively, analysis of cell-free DNA from a blood sample might give the same information. Cell-free DNA is released into the blood stream from both malignant and nonmalignant cells, and it carries the same genetic aberrations as the cells it comes from.

The aims of this thesis were to improve oncogenetic counselling, by increasing the knowledge about who should be offered genetic tests and what genes to test, and to improve treatment decision-making, by developing a clinical pipeline for a new cancer biomarker, cell-free tumour DNA.

In Study I, we explored the genetic cause for a suspected highly penetrant, autosomal dominant rectal cancer syndrome in one family. We analysed the blood samples from six family members by massive parallel sequencing, and found six variants that could contribute to the increased risk for rectal cancer in the family. The variants were in the genes *CENPB, ZBTB20, CLINK, LRRC26, TRPM1,* and *NPEPL1*. None of the variants have a known connection to hereditary cancer syndromes and further studies are needed to investigate their potential involvement.

In Study II, we evaluated the efficacy of performing massive parallel sequencing in patients with three or more primary tumours. We found two likely pathogenic variants in two individuals. In a woman with a clinical diagnosis of MEN1, we found a synonymous variant in the *MEN1* gene. By RNA analysis, we could show that the variant activates a cryptic splice site, which leads to a deletion of 14 nucleotides, and a frameshift in the RNA transcript. The variant segregated with disease in the family, and it was classified as likely disease-causing. In a woman with multiple Lynch-associated tumours and microsatellite instability, we found a synonymous variant in the *MLH1* gene. This variant has recently been described as an epimutation, leading to methylation of the *MLH1* promoter, and Lynch syndrome. Both these genes had previously been analysed in blood samples from the two women, but the variants had not been recognized as pathogenic. We therefore conclude that genetic re-analysis may be beneficial, especially for patients with a clear clinical diagnosis of a specific hereditary cancer syndrome.

In Study III, we developed a clinical method for detecting cancer-associated CNAs (copy number alterations) in cell-free tumour DNA from patients with gastro-oesophageal cancer. The same clinical laboratory protocol as for non-invasive prenatal testing at Karolinska University Hospital, Stockholm, was used, and for the last bioinformatic steps WISECONDOR and ichorCNA were used. This generated a list of potentially cancer-associated CNAs, which we manually annotated. We compared the variants to the tissue sample from the same participant, and also with a control set of samples from pregnant women. We found cancer-associated CNAs in 14/26 (54%) of the individuals with detectable

cancer-associated CNAs in the tissue samples. Potentially clinically actionable amplifications were detected in the genes *VEGFA, EGFR*, and *FGFR2*. This study showed that the clinical pipeline we have set up for the non-invasive prenatal testing workflow can be used in the oncogenetic field, and CNA analysis in patients with gastro-oesophageal cancer might be used as a biomarker in the future.

In conclusion, we report six new genes with a potential involvement in a hereditary rectal cancer syndrome, we suggest re-analysis might be beneficial for patients without a molecular diagnosis but who fulfil testing criteria for specific cancer syndromes, and we show that the non-invasive prenatal testing clinical pipeline can also be used in the oncogenetic setting.

# LIST OF SCIENTIFIC PAPERS

I. Wallander K, Thutkawkorapin J, Sahlin E, Lindblom A, Lagerstedt-Robinson K. Massive parallel sequencing in a family with rectal cancer. Hered Cancer Clin Pract. 2021 Apr 7;19(1):23. doi: 10.1186/s13053-021-00181-2. PMID: 33827643; PMCID: PMC8028209.

II. Wallander K, Thonberg H, Nilsson D, Tham E. Massive parallel sequencing in individuals with multiple primary tumours reveals the benefit of re-analysis. Hered Cancer Clin Pract. 2021 Oct 28; 19(1):46. doi: https://doi.org/10.1186/s13053-021-00203-z. PMID: 34711244; PMCID: PMC8555269.

III. Wallander K, Eisfeldt J, Lindblad M, Nilsson D, Billiau K, Foroughi H, Nordenskjöld M, Liedén A, Tham E. Cell-free tumour DNA analysis detects copy number alterations in gastro-oesophageal cancer patients. PLoS One. 2021 Feb 4;16(2):e0245488. doi: 10.1371/journal.pone.0245488. PMID: 33539436; PMCID: PMC7861431.

# LIST OF RELATED SCIENTIFIC PAPERS

I. Wallander K, Liu W, von Holst S, Thutkawkorapin J, Kontham V, Forsberg A, Lindblom A, Lagerstedt-Robinson K. Genetic analyses supporting colorectal, gastric, and prostate cancer syndromes. Genes Chromosomes Cancer. 2019 Nov;58(11):775-782. doi: 10.1002/gcc.22786. Epub 2019 Aug 7. PMID: 31334572; PMCID: PMC6771512

II. Barbany G, Arthur C, Liedén A, Nordenskjöld M, Rosenquist R, Tesi B, Wallander K, Tham E. Cell-free tumour DNA testing for early detection of cancer - a potential future tool. J Intern Med. 2019 Aug;286(2):118-136. doi: 10.1111/joim.12897. Epub 2019 Apr 8. PMID: 30861222.

# TABLE OF CONTENTS

# LIST OF ABBREVIATIONS

| | |
|---|---|
| CIMP | CpG island methylator phenotype |
| CIN | chromosomal instability |
| CNA | copy number alteration |
| ctDNA | cell-free tumour DNA |
| EBV | Epstein-Barr virus |
| GEAC | gastro-oesophageal adenocarcinoma |
| GS | genomically stable |
| MEN1 | multiple endocrine neoplasia type 1 |
| MPS | massive parallel sequencing |
| MSI | microsatellite instable/instability |
| NIPT | non-invasive prenatal test |
| SNV | single nucleotide variant |
| sWGS | shallow WGS |
| TMB | tumour mutational burden |
| WGS | whole-genome sequencing |
| WES | whole-exome sequencing |

# 1 BACKGROUND

## 1.1 WHAT IS CANCER?

### 1.1.1 Cancer hallmarks are acquired capabilities in malignant cells

Cancer is a result of natural selection. Normal cells can progressively evolve into cancer cells, acquiring capabilities enabling them to survive and multiply. Cancer cells are characterised by their unlimited proliferative potential. Their capabilities have been categorised comprehensively into “cancer hallmarks” and enabling capabilities, as described by Hanahan and Weinberg. They include for instance resisting cell death and genome instability (1). There are cellular intrinsic mechanisms working against unlimited growth. Therefore, multiple genetic changes are needed, in order for the many different defence mechanisms to be put out and cancerous growth to be achieved (2).

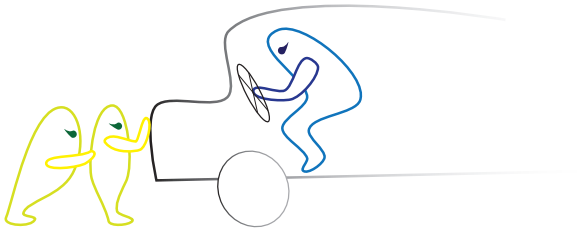
### 1.1.2 There are different kinds of genetic changes

Genetic changes increasing the risk for a cell to become malignant can be copy number alterations (CNAs), single nucleotide variants (SNVs) and indels, structural rearrangements, and epigenetic changes. CNAs result in less or more genetic material. An amplification, a type of CNA, is a segment of DNA that has been duplicated many times. Even though SNVs only include a change of one single base pair, they can occur at critical positions and change one or more proteins. Missense variants are SNVs predicted to change the amino acid, and SNVs not predicted to affect the amino acid sequence are called synonymous. Nonsense variants introduce a premature stop signal. Frameshift variants generate an unnatural reading frame in the translation from RNA to protein. Indels (short insertions and/or deletions) are similar to SNVs but include a change of not one but a few or a stretch of base pairs. Structural rearrangements, such as translocations, can lead to changes in the regulation of gene transcription or changed amino acid sequences in proteins (2). Epigenetic changes do not occur within the genetic code but can nevertheless alter the expression of genes (3). These different types of changes can have similar effects but may require different analyses to detect.

### 1.1.3 Tumour suppressor genes and oncogenes drive the evolvement towards cancer

Within a cancer cell, there are both genetic changes in genes involved in the cancer formation, called driver genes, and changes in genes not important for the cancer evolvement, called passenger genes (4). In general, the driver genes can be divided into tumour suppressor genes and oncogenes. Tumour suppressor genes, for instance the *CDH1* gene, code for proteins involved in blocking cancer development, through suppressing proliferation or through control of genomic damage. Usually, some tumour suppressor genes need to be blocked, by deletion of the gene, silencing of gene expression, or by SNVs that inactivate the protein product (loss-of function mutations), for cancer to occur. Both copies of the specific tumour suppressor gene need to be deactivated for the cell to gain a malignancy potential. Oncogenes, for instance the *EGFR* gene, code for proteins promoting cell proliferation, and they are activated rather than deactivated, for instance by amplification of the gene or an SNV that gives the protein an enhanced function.

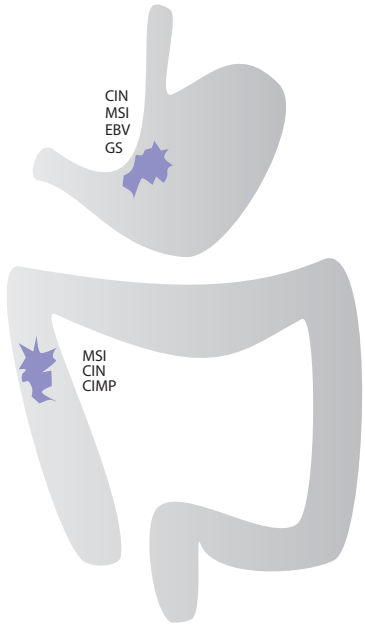
In general, only one of the paired oncogene copies needs to be changed for the cell to gain survival benefits.



*Tumour suppressor genes blocking cancer progression, an oncogene hitting the gas pedal.*

### 1.1.4 Gastrointestinal cancer can be classified based on genetic changes

Colorectal cancer formation is driven by three main molecular pathways: microsatellite instability (MSI), chromosomal instability (CIN), and CpG island methylator phenotype (CIMP) (5). A tumour with different microsatellite lengths compared to the normal tissue is classified as MSI. The microsatellites are short repetitive DNA sequences spread across the genome. When the mismatch repair system in a cell is out of order, frameshift SNVs or indels occur spontaneously more frequently than in normal cells. This is detectable as MSI, since the microsatellites gain varying size, which can be measured. The frameshift SNVs and indels can drive oncogenesis, by occurring in cancer driver genes. The DNA mismatch repair genes *MLH1, MSH2, MSH6, (EPCAM),* and *PMS2* are often classified as tumour suppressor genes (6). CIN tumours are characterized by CNAs spread throughout the genome, including amplifications of oncogenes (7). The mechanisms behind CIN often include changes of proteins involved in chromosome segregation in cell division, or stress response (8). Aneuploidy (abnormal chromosome number) is present in most cancers. The precise relationship between aneuploidy and cancer formation is not known (8) (9). CIMP drive tumourigenesis through transcription inactivation (obstruction of the process when DNA codes for RNA) in tumour suppressor genes. By DNA hypermethylation at promoters rich on C and G bases (CpG islands), the genes are turned off (5).




*Molecular subgroups of gastric and colon cancer.*

Gastric and oesophageal adenocarcinomas can be divided into four different molecular subtypes, according to the classification system suggested by the Cancer Genome Atlas Research Network. This is partly based on the genetic changes, molecular cancer-driving mechanisms, that have been activated in the tumour cells. The molecular subtypes are CIN, MSI, Epstein-Barr virus (EBV)-positive, and genomically stable (GS). The CIN subtype is commonly associated to pathogenic variants in the *TP53* gene. The MSI subtype is associated to hypermutation (many SNPs and indels in the tumour cell genomes) and a silenced *MLH1* gene. The EBV subtype often harbours pathogenic variants in the *PIK3CA* gene. The genomically stable subtype has recurrent (occurring in many different tumours) pathogenic variants in the *CDH1* and *RHOA* genes (10) (7). Other subclassifications, and cost-effective strategies for molecular classification, have been published (11). But in clinical practice in Sweden and many other countries, the histology (diffuse, intestinal or mixed), differentiation, and TNM stage (individual staging of the tumour, and whether it is spread to nodes and/or as distant metastases) are still the most common classifications and prognostic stratification systems (12).

### 1.1.5 The cancer-driving events differ according to tissue and cancer type

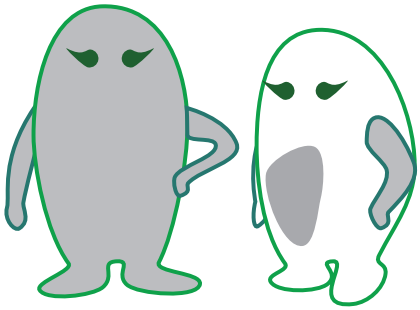
In some cancer types, the genome changes are well known and non-variable. This is true for chronic myeloid leukaemia, defined by the translocation between chromosome 9 and 22 (13). The *TP53* gene, contrarily, is commonly altered in most cancer types, albeit in many different ways (14). In gastro-oesophageal adenocarcinoma (GEAC), however, each specific SNV only occurs in a small subset of the cases, suggesting that there are limited possibilities to find targeted therapies applicable to a reasonably large cohort in clinical trials (15) (10). CNAs are more common than SNVs in GEAC and recurrent CNAs are seen in more than 50% of the cases (10) (16) (17). This is important to know when deciding which analysis method to use, to be able to detect genetic changes associated to the specific cancer.

## 1.2 IS CANCER INHERITED?

### 1.2.1 Somatic genetic changes occur within the organism

Most genetic aberrations within a tumour have occurred as the cells develop into cancer cells, within the organism. A normal colorectal epithelium can develop into adenomas, and acquire pathogenic variants in the WNT signalling pathway (most often inactivation of the *APC* gene), evolving to advanced adenomas. The epithelium cells typically then acquire activating SNVs at specific positions (hotspots) within the *KRAS* gene, and deletions of the part of chromosome 17 containing the *TP53* gene. Thereby, finally, they form an adenocarcinoma (5). Those genetic variants are called somatic. Less is known about the early genetic events in cancer formation than in advanced cancer, and there are ongoing initiatives for increasing the knowledge about cancer formation (18).

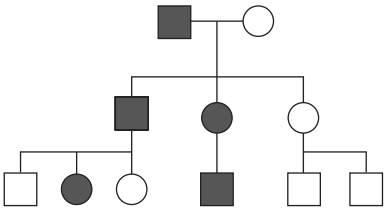
Cancer is a clonal disease, and not all cancer-associated pathogenic somatic variants occur in all cancer cells. For instance, in gastric cancer, *HER2* (*ERBB2*) gene amplification can sometimes be detected in some tissue biopsies but not others from the same patient. This has implications when it comes to treatment options, since there are targeted therapies for patients with HER2 overexpression (19).



*Germline genetic variants present in all cells (grey), and somatic genetic variants present in a subset of cells.*

### 1.2.2 Germline genetic changes are inherited

Genetic variants which exist in all cells of the body, and can be inherited, are called germline. If germline genetic variants are pathogenic, they can give rise to disease, and increase the risk for cancers. There are known inherited cancer syndromes, often caused by germline SNVs causing a high risk to develop cancer. The genetic change is commonly situated in a tumour suppressor gene and when the other gene copy is lost by a cellular mechanism, the cell acquires a growth advantage (20).



*Generic autosomal dominant pedigree. Transmission occurs from both fathers and mothers, and both men and women are affected.*

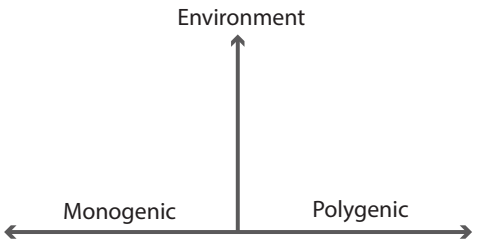
Cancer syndromes are almost always autosomal (not sex-linked) dominant or recessive. In dominant syndromes, a genetic variant in one of the two copies that exist for each gene increases the risk (gives predisposition) for cancer. In recessive inheritance, both copies are altered in the germline when the syndrome occurs. Most often, tumour suppressor genes, including genes maintaining genome stability, are responsible for hereditary cancer. Even though both copies of the tumour suppressor gene need to be inactivated on the cellular level to have impact, the cancer syndrome can be inherited in a dominant fashion. This is because the “second hit” (the mechanism inactivating the gene copy with no inherited pathogenic genetic variant) occurs somatically (2). In autosomal dominant cancer syndromes, the penetrance (the fraction of individuals affected by cancer compared to all individuals carrying pathogenic variants) is seldom 100%. The explanation may be either that no second hit has occurred in the critical tissue at the critical time, and a second hit may not in itself be cancer-causing. Often, more genetic changes than two are needed for the cell to be transformed into a malignant state. Also, the expressivity (the manifestations of the

disease) may vary within a family. The same pathogenic variant can cause one type of cancer in one individual and another in a relative carrying the exact same variant. Aspects to take into account when predicting cancer risks from specific sequence variants are environmental factors and modifying variants within the rest of the genome (21) (20).

### 1.2.3 Cancer is a complex disease

The heritability of cancer is overall approximately 30% (22). Heritability can be defined as the proportion of the variance in a specific trait that is genetic (2). Most of the genetic factors causing hereditary cancer are not known (22). Apart from monogenic cancer syndromes, such as Lynch syndrome and hereditary diffuse gastric cancer, most cancers are caused by both multiple genetic changes and the environment. Cancer can therefore be considered a complex disease (2).

Polygenic reasons for cancer (many genetic loci together increasing the susceptibility of cancer formation) have not been as extensively studied as monogenic. There are thousands of single-nucleotide polymorphisms (base pair variations in the germline genome) per cancer type, which have a known connection (23). These genomic variations might account for more than 10% of the heritability (24). The genetic variations do not in themselves have clinical relevance, but they can be combined into a polygenic risk score, which has the potential to discriminate those with a higher risk of developing cancer from those with lower risk. There is an ongoing debate if this score should be implemented in the clinical setting or not (25) (26). Example of environmental factors also influencing the cancer incidence are cigarette smoke, viral infections, and obesity (27).




*Cancer occurrence is usually both dependent on genetics and different degrees of environmental influence.*

### 1.2.4 Some monogenic gastrointestinal cancer syndromes are known

Among all individuals with colorectal cancer, 12-35% are suspected to be due to heritable factors (28) (22) (29). Around 5-10% of individuals with colorectal cancer carry pathogenic germline genetic variants in genes that have a known association to a high or a moderate risk for colorectal cancer (30) (31) (32). Like colorectal cancer, most gastric cancer does not seem to have a genetic cause. It is estimated that around 10% (1-20% depending on study) is due to genetic factors (28) (22) (33), and 1-3% of the total heritability is estimated to be caused by known genetic pathogenic variants (34).

Colorectal cancer syndromes can be divided into polyposis and non-polyposis, whereas hereditary gastric cancer syndromes include hereditary diffuse gastric cancer, familial intestinal gastric cancer and gastric adenocarcinoma and proximal polyposis of the stomach (35) (36). Colorectal polyposis syndromes, with multiple polyps of different types, are both dominant and recessive. Familial adenomatous polyposis, caused by pathogenic variants in

the tumour suppressor *APC* gene, is inherited in an autosomal dominant fashion, whereas for instance *MUTYH*-associated polyposis is autosomal recessive (35). Gastric adenocarcinoma and proximal polyposis of the stomach is inherited in a dominant fashion (37), and has been shown to be caused by pathogenic variants in a promoter (part of the regulatory region) for the *APC* gene. It is a variant in the familial adenomatous polyposis syndrome spectrum, but only includes polyps in the stomach and not in the bowel (38) (39) (40). New genes associated to recessive hereditary colorectal polyposis have been discovered by whole-exome sequencing, in for instance the *NTHL1* (41) and *MSH3* (42) gene.

Hereditary non-polyposis colorectal cancer syndromes can be divided into those with DNA deficient and proficient mismatch repair tumours (35). Pathogenic germline variants in the mismatch repair genes cause Lynch syndrome, which includes an increased risk for mismatch repair-deficient tumours in the colorectal tract and uterus, among others (43). In most families with apparent dominant non-polyposis colorectal cancer syndrome with a mismatch repair-proficient tumour, no causative gene is found. Causative genes have been suggested, and the only one with consistent association so far is the *RPS20* gene (44) (45). Among other candidates are genes involved in cell-cycle checkpoints and chromosome-associated proteins (46). An example of this is the *CENPE* gene (47) (48).

Gastric cancer is traditionally subdivided histologically into diffuse and intestinal, with the diffuse subtype typically containing signet ring cells. Hereditary diffuse gastric cancer syndrome is associated with pathogenic variants in the *CDH1* gene (49), and, in a minority of the cases, the *CTNNA1* gene (50) (51). It gives an increased risk for diffuse gastric cancer and lobular breast cancer, inherited in a dominant fashion. The *CDH1* gene codes for E-cadherin, a cell adhesion protein important for tissue architecture and homeostasis. It is considered a tumour suppressor gene. In more than 60% of stomach tissue biopsies from asymptomatic carriers of a pathogenic *CDH1* variant, signet ring cell carcinoma is detected. The detection rate depends on pathology protocol used (52) (53) (54). In a carrier of a pathogenic *CDH1* variant, the lifetime risk of diffuse gastric cancer is 30-70%, and the risk for lobular breast cancer in a woman might be around 50% (55) (56). The manifestations of cancers associated to *CDH1* pathogenic variants differ between different families, and in some families there seems to be an increased risk for colorectal cancer (57). Also, with increased genetic testing, it has become evident that not all families with a pathogenic variant in the *CDH1* gene display cases of gastric cancer. In some families, only lobular breast cancer occurs, and some families do not fulfil any of the clinical criteria for the syndromes. The risk for cancer in those families is not known (58).

### 1.2.5 Lynch syndrome can also be caused by epimutations

Constitutional (germline) epimutations should also be considered as a hereditary cause for colorectal cancer. Constant methylation of the promoter of the *MLH1* and *MSH2* gene has been shown to silence that copy of the gene, and lead to Lynch syndrome. The methylation can be secondary, caused by a genetic alteration, or primary, without any detectable change of the genome connected to the methylation. The secondary epimutations are autosomal dominant, but the inheritance for the primary epimutations does not follow that pattern and often occur without any family history of colorectal cancer. For the *MLH1* gene, both

secondary and primary epimutations have been discovered, and for the *MSH2* gene, the epimutations are secondary to deletions in the *EPCAM* gene (59).

### 1.2.6 MEN1 syndrome is monogenic but not all pathogenic variants are detected

Multiple endocrine neoplasia type 1 (MEN1) is an autosomal dominant cancer syndrome caused by pathogenic variants in the *MEN1* gene. The *MEN1* gene is a tumour suppressor, and most pathogenic germline variants described inactivate one copy of the gene. Parathyroid hyperplasia occurs in almost all individuals with MEN1. Pancreatic islet and anterior pituitary tumours are also frequent. Many other manifestations, such as lipoma and angiofibroma, have been described (60). Pathogenic variants in the *MEN1* gene can be found in 80-90% of individuals with a clinical diagnosis and also a family history of MEN1, or less if there is no family history (61) (62). When no pathogenic *MEN1* variant is found, there might be somatic predisposing variants not present in blood, or germline variants in the regulatory regions, or the tumours might have occurred by chance (60) (63).

### 1.2.7 Multiple primary tumours might have a genetic cause

Multiple primary tumours are usually defined as two or more tumours in one individual, not caused by metastasis, recurrence or local spread (64). Most publications concerning multiple primary tumour genetics include individuals with two primary malignancies, and reports on cohorts with three or more are rare. Approximately 0.1-0.2% of all patients with cancer have at least three primary malignancies (65) (66). Multiple primary tumours can be caused by inherited genetic variants, sometimes associated to an underlying developmental abnormality (mono- or polygenic), common exposure to carcinogenic agents or hormones (environmental), or a combination of all. Whitworth et al have performed whole-genome sequencing in a cohort of 460 individuals with at least two primary tumours before 60 years or three before 70 years of age (n=182), and they found pathogenic or likely pathogenic variants in 15% of the participants (67) (and personal communication). That indicates it might be clinically beneficial to offer genetic screening to all patients with three or more primary tumours, whether they fulfil testing criteria for a specific hereditary cancer syndrome or not.

## 1.3 WHY ARE GENETIC ANALYSES OF HEREDITARY CANCER PERFORMED?

### 1.3.1 Finding a pathogenic variant enables surveillance programs

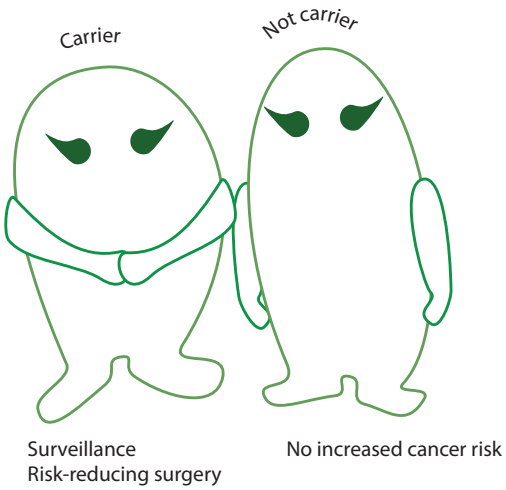
Hereditary cancer is generally suspected when there are cases of early onset cancer, or many cancers of the same spectrum within one family. In the clinical setting, families with a suspected hereditary cancer syndrome can often be offered genetic testing, such as a screening for genes with known association to a hereditary cancer syndrome. A prerequisite for this is that there are genes with a known connection to the suspected syndrome, and preferentially a family member with cancer is tested. There are criteria for when each known specific hereditary cancer syndrome should be suspected. If a pathogenic variant is found, predictive testing can be offered in the family. Healthy family members carrying the pathogenic variant have an increased risk for cancer (if the inheritance is dominant), and those not carrying the variant do not have an increased risk. Carriers of dominant cancer syndrome pathogenic variants are as a rule offered surveillance programs for the cancers

associated to the syndrome, and sometimes prenatal testing is an option (68). The surveillance program depends on the specific cancer syndrome, and some examples are listed below.

For suspected hereditary colorectal cancer, one or two different pre-set gene panels, including genes known to harbour pathogenic variants for adenomatous and non-polyposis hereditary colorectal cancer, is often used in the clinic (69). All individuals with colorectal cancer are recommended screening for Lynch syndrome, for instance by checking for MSI in the tumour, by the National Institute for Health and Care Excellence in Great Britain, and it has been shown to be cost effective (70) (71). Predictive testing is offered worldwide in families with a pathogenic variant in a DNA mismatch repair gene and surveillance programs with colonoscopy and different kinds of gynaecological assessment is recommended to carriers (72) (73) (74) (75). There are international colonoscopy surveillance recommendations for carriers of genetic variants that only moderately increase the risk for colorectal cancer, and more studies are needed to improve those guidelines (76) (72).

Concerning gastric cancer, genetic testing is only possible for hereditary diffuse gastric cancer and gastric adenocarcinoma and proximal polyposis of the stomach, as familial intestinal gastric cancer has no known connection to a specific gene (77). If a pathogenic variant is found in the *CDH1* gene, all family members carrying the variant are recommended prophylactic gastrectomy, and women are also recommended breast cancer surveillance and in some cases risk-reducing mastectomy. In families where no pathogenic *CDH1* variant is found, regular gastroscopy surveillance is preferred to gastrectomy. In the most recent international guidelines for hereditary gastric cancer, much emphasis is put on genetic testing (78).

Patients with MEN1 and their families should be offered care by a multi-disciplinary team, and all individuals with *MEN1* germline pathogenic variants should be recommended regular tumour screening (63).

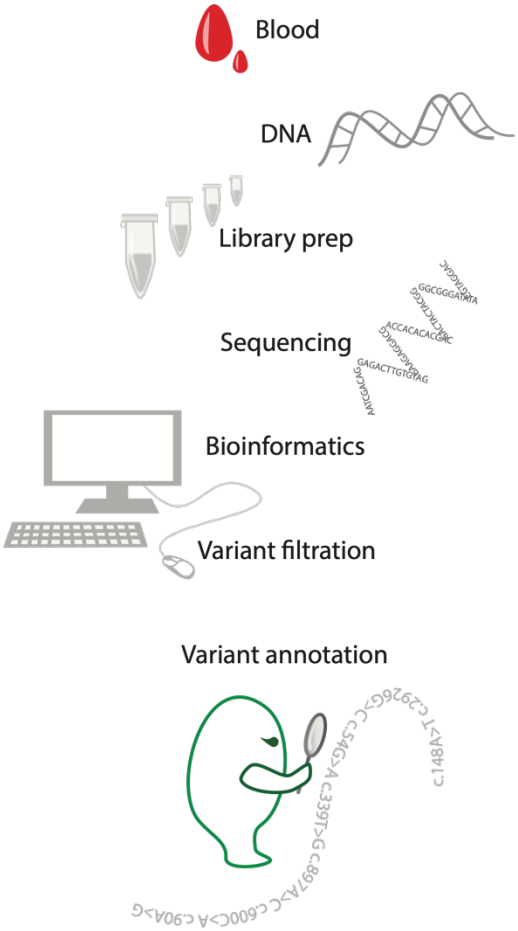



*A carrier of a pathogenic variant for an autosomal dominant cancer syndrome is generally offered surveillance for the associated cancers, and in some cases risk-reducing surgery. Family members not carrying the variant are not at increased risk for the inherited cancer syndrome.*

## 1.4 HOW DO GENETIC ANALYSES FOR HEREDITARY CANCER WORK?

### 1.4.1 WGS and WES enable fast sequencing of the whole genome

To analyse a genetic sequence, traditionally Sanger sequencing has been used. Its basic principles are in vitro DNA replication and random incorporation of chain-terminating bases, followed by size-dependent assortment and detection of synthesized bases (79). This is a feasible analysis method for specific genes, and frequently used still, but not when the entire genome must be sequenced. For this, massive parallel sequencing (MPS), with billions of sequencing reactions carried out in parallel, can be used. That is sometimes referred to as "next generation sequencing", as it is an improvement from the first sequencing techniques, such as Sanger, when it comes to data output and speed.




*Overview of a typical MPS workflow.*

After DNA has been isolated, the first step in MPS is library preparation, in which nucleic acids representing the input genome are created. Most often, this includes genome fragmentation, sometimes also amplification by PCR (polymerase chain reaction), and then attachment to specific adapters and identifying index tags (80). The second step is sequencing and imaging, in which the DNA sequence of the nucleic acids is registered. Depending on the company, different sequencing techniques are used. IonTorrent (Guilford, CT, USA) measures pH changes (81), and both Illumina (San Diego, CA, USA) and Roche (Basel, Switzerland) use optical detection (82) (83), to read the bases incorporated to the template strand. In optical detection, fluorescent-labelled bases are added and read as they pair with the bases of the DNA strand, the "target sequence". This generates reads (sequences of bases corresponding to the genome). In whole-exome sequencing (WES),

the genetic sequences are enriched for the exome (protein coding) parts of the genes and in whole-genome sequencing (WGS), all regions are mapped and analysed (80). The third step is bioinformatics, where the reads are processed so that specific genetic variants can be presented. The bioinformatics include genome alignment (in which the overlapping reads are aligned to the known reference sequence), variant analysis, variant calling, variant annotation, prioritization of variants, and variant quality control. At the Karolinska University Hospital Laboratory, Stockholm, Sweden, a collaboration with the Science for Life Laboratory offers a robust and efficient handling of clinical and research blood samples in the MPS pipeline. MIP, Mutation Identification Pipeline framework, is used for the bioinformatics steps. MIP runs the first output from the sequencing data through a pipeline of software tools, and generates a list of variants which can be visualized in Scout (a graphical user interface). This workflow has been thoroughly described by Stranneheim et al (84).

### 1.4.2 Variant filtration is necessary to find the genetic variants that could be of interest

There are usually too many variants detected by MPS to know which ones might be causative for a specific condition. Filtering of the variants is needed. This filtering process depends on what the aim of the analysis was. A frequently used approach is to remove variants common in the general population, and only include exonic variants or splice site variants (see below), which are predicted to affect the protein sequence. This process can be performed according to specific guidelines, and tools are developed to ease and standardize the process (85). Variants passing the filtering steps are classified according to criteria stipulated by the American College of Medical Genetics (ACMG). In this approach, the variants are sorted into benign/likely benign, of unknown significance, and pathogenic/likely pathogenic (86). A pathogenic variant is not necessarily disease-causing, since it may for instance occur in a gene causing a recessive disease, without another pathogenic variant on the other gene copy. The ACMG classification system is one-dimensional, and disease- and gene-specific systems have been proposed. An example of this is a proposition of adjustments for variant classification of the *MEN1* gene (87). The Clinical Genome Resource ClinGen has a sequence variant interpretation group, aiming at evolving the ACMG criteria (88). Recently, a classification system with independent tiers for assessment of functional consequences of the variant and the clinical importance, has been suggested. This would be especially important for variants classified as of unknown significance by the ACMG system (89).

### 1.4.3 RNA-sequencing can be used to detect splicing variants

Splicing is part of the transcription and translation process (when the DNA strand codes for a protein). Non-coding sequences (introns) are removed from the RNA copy of the DNA and the exons are ligated (2). DNA variants can disturb this process and give rise to RNA with aberrant splicing, for instance without the usual coding parts or with non-coding parts incorporated. These variants are often called splicing variants. This can lead to aberrant proteins and diseases. Analyses of RNA profiles in blood have been used to detect pathogenic DNA variants in patients with cancer syndromes, and could be considered an important supplement to DNA analyses (90).

### 1.4.4 Array-CGH finds CNAs

To look for CNAs, array-CGH (comparative genomic hybridization) can be used. Fluorescence-labelled test DNA and control DNA are simultaneously hybridized to DNA targets on an array, and a difference in signaling intensity between them indicates gain or loss of genetic material in the test DNA (91). This can be applied both to blood sample analysis, to look for inherited CNAs, and to analyses of tumour tissue, to look for somatic, potentially cancer-associated CNAs. The resolution of the array-CGH depends on the molecules that were used to construct the microarray, and they could either be spaced across the genome or concentrated at the locations of interest (2).

## 1.5 WHAT IS A LIQUID BIOPSY AND CELL-FREE DNA?

### 1.5.1 Cell-free DNA is released into the blood stream and detected in a blood sample

Human cells can release DNA into body fluids, such as blood and urine, wherein it can be detected and analysed. This cell-free DNA is released either by apoptosis or necrosis of the cells or by active release, and the way it is released affects the length of the DNA fragments. The fragments generated by apoptosis are about 160-170 basepairs long, or multiples thereof. This is because DNA is protected when it is wrapped around proteins, in what is called a chromatosome (the DNA strand and the protective histone proteins), and the cuts generally occur in between these entities (92). Cell-free DNA originating from necrotic cells is generally longer than from apoptosis, and the DNA actively released is usually associated to other proteins or present within vesicles (93) (94). There are indications that the cell-free DNA from malignant cells might be shorter than that originating from healthy cells. Studying the length of the cell-free DNA fragments is considered important when it comes to cancer detection, but there is a need for more research in that area (95) (96). Most of the cell-free DNA in plasma stems from haematopoietic cells (97) and in pregnant women, some of the cell-free DNA stems from the placenta (92).

Individuals with cancer have a higher mean cell-free DNA concentration than controls. Some of this cell-free DNA comes from the tumour cells, which is called cell-free tumour DNA (ctDNA) (98). The cell-free DNA is degraded by nuclease activity, by the liver and spleen or by urinary excretion, and its halftime in the circulation is estimated to 2 hours (99). This is under debate, though. The elimination process might be two-phased, and not enough in-vivo studies in individuals with malignant tumours have been performed (100).

### 1.5.2 Cell-free DNA can be used as a biomarker

Cell-free DNA is a potentially useful biomarker, since ctDNA often harbours the same genetic changes as the tumour cells. Analysis of cell-free DNA, a sub-category of “liquid biopsy”, has been shown to have great potential as a measure of tumour burden, a prognostic marker, a disease monitor, and a detection method for resistance mechanism within the tumour (98). The golden standard diagnostic tool up until now, when it comes to solid (non-haematological) tumours, has been tissue biopsies. This does, however, have limitations. A tissue biopsy is not readily available in all patients and not all tumour cells carry the same genetic changes, due to tumour heterogeneity (different cells from the same tumour have different characteristics, such as genetic aberrations). A single tissue sample

may therefore not be representable of the genetic changes in the whole tumour, and it has been shown that cancer-associated genetic changes that are missed in tissue biopsies can be found in plasma (101).

## 1.6 HOW IS CELL-FREE DNA ANALYZED?

### 1.6.1 Only the DNA outside of blood cells is desired in the analysis

Starting with a regular blood draw from the arm vein, with the same procedures as any other peripheral blood test, cell-free DNA can be isolated from the blood. Within a blood sample, there are both blood cells and cell-free DNA. Ideally, the cell-free DNA is isolated from plasma, which is the cell-free part of blood including clotting factors (achieved by treating the sample with anti-coagulants), and not the serum (the liquid remaining after coagulation of the blood) (102). To avoid contamination from the genomic DNA (DNA from within the blood cell nucleus), either rapid plasma isolation (before the blood cells have started breaking) or specific test tubes are usually used. There are different collection tubes designed to preserve cell-free DNA and stabilize the blood cells (103), for instance Streck tubes (Cell-Free DNA BCT, Streck, La Vista, NE, USA). Biological factors within the individual affect the cell-free DNA yield from a blood sample. Smoking, medical treatments, kidney failure, exercise, autoimmune diseases etcetera might influence the cell-free DNA levels and composition. Too little is so far known about these effects (104). After the cell-free DNA is isolated, it is analysed by different techniques depending on which aberrations are sought.

Other body fluids than plasma can also potentially be used for cell-free DNA analysis. Urine (105) and gastric wash have for instance been studied for GEAC patients (106).

### 1.6.2 Customised analyses are used, depending on what is sought

The sought aberration decides the analysis technique. This is true for all DNA analyses, not only cell-free DNA.

The earliest plasma analyses of GEAC often measured total cell-free DNA concentration. This can be done by a targeted genetic analysis of Alu repeats (a DNA sequence recurring throughout the human genome) (107) (108). But the total cell-free DNA concentration varies physiologically with age and co-morbidities, and it is a non-specific test for cancer (109).

Methylation (a methyl group attached to a base, most often cytosine, in the DNA strand) analysis is a more specific approach for ctDNA analysis than total cell-free DNA concentration measurements. GEAC patients can be distinguished from control groups by methylation pattern analysis, such as methylation-sensitive PCR (amplification of both potential methylated and un-methylated target regions in the sample, to see which is present). Genomic regions containing significantly different methylation patterns in cancer tissue compared to non-cancer tissues can be targeted. A methylation signature analysis of 5hmC (partly degraded methylated cytosine) (110), and analysis of specific genomic promoter regions in genes such as *SEPTIN9* (111)*, CDKN2A* (112)*, ZIC1, HOXD10, RUNX3* (113), *SOX17* (114), *APC*, *RASSF1A* (115), *RNF18,* and *SFRP2* (116) have been reported.

For non-targeted CNAs, a whole-genome sequencing approach is needed (117). Whole-genome sequencing can be performed with high or low coverage (number of unique reads

that include a given base when the sequence has been re-constructed). In our experience, a low-coverage approach, shallow WGS (sWGS), with 0.2X (sequencing 1/500 basepairs) detect CNAs that are 2 megabasepairs or longer. Methods for CNA analysis in cell-free DNA, using WISECONDOR (118) or ichorCNA (119) (120) (121) (122) to detect ovarian, prostate and breast cancer, have been reported. Reports on CNA detection in cell-free DNA are few in number, and often small proof-of-concept studies (123). When it comes to studies on CNAs in GEAC, there are even fewer studies. One study included a cohort of 30 individuals with mostly metastatic GEAC. ichorCNA was used for identifying CNAs, which were detected in 23 (77%) of the individuals (117). In advanced gastric cancer, the ctDNA fraction (ctDNA amount in relation to the total cell-free DNA) has been reported to be approximately 0.3-8% (124), and most studies using CNA analysis in cell-free DNA require at least 5-10% ctDNA fraction in order to proceed successfully with low-coverage whole genome sequencing (125) (119).

It might be beneficial to first make an approximation of the ctDNA fraction within the sample, and then use a targeted analysis (knowing what genetic aberrations to look for) for those with low ctDNA fractions, and an untargeted analysis (looking for any cancer-associated CNA, for instance) for those with high enough ctDNA fractions. One such approach is named FAST-Seq. A general overview of aneuploidy is established by a genome-wide z score, and this can be used as a pre-screening in the decision of further analyses (125) (126). Suggested approaches for detecting CNAs in cell-free DNA are for instance massively multiplexed PCR and next-generation sequencing (mmPCR-NGS) and Repetitive Element AneupLoidy Sequencing System (RealSeqS). mmPCR-NGS is based on SNP analysis, and might be able to detect low ctDNA fractions (127) (128). RealSeqS uses a single primer pair (starting points for the copying of the genome), pinpointing repetitive regions spread throughout the genome. It can be considered a further development of the Fast-Seq method, and can detect CNAs despite extremely low ctDNA fraction. RealSeqS has been used in a cohort including nonmetastatic GEAC patients and could detect CNAs in 49% of them (129).

Specific amplifications (such as in the *HER2* gene) can be quantified in a cell-free DNA gene panel (only analysing pre-selected genes and variants within genes) (130). The ctDNA *HER2* gene can also be compared to a reference gene and analysed by qPCR (amplification of the gene and at the same time measurement of the amount) (131), or by digital droplet-PCR (an amplification method which is highly specific for a gene or even an exact SNV, which can detect extremely low copy number levels of the target sequence) (132).

To detect cancer-associated SNVs in cell-free DNA, a gene-panel is often used. The panel typically includes known cancer driver genes. Possibly pathogenic SNVs not present in white blood cells, or validated in a tumour tissue, are considered cancer-associated. By that approach, about half of study participants with GEAC usually have detectable cancer-associated SNVs in the cell-free DNA, if they have cancer-associated SNVs detectable in the tumour tissue (133) (134) (135). When only a few single SNVs are sought, for instance when a tissue analysis has been performed in the same individual, ddPCR can be used (136). Cancer-associated structural rearrangements can be detected in gastric tumour tissue by sequencing, and subsequently by targeted sequencing in cell-free DNA, similar to for SNVs.

### 1.6.3 The NIPT workflow can be used for cancer CNA detection

In the non-invasive prenatal test (NIPT), trisomy 21, 18, and 13 are looked for within the blood of a pregnant woman. After plasma has been isolated, sWGS and interpretation of aberrant chromosome counts are performed. A likelihood of foetal aneuploidy is calculated. If foetal aneuploidy is statistically likely, an invasive test (of placenta or amniotic fluid) is recommended for confirmation (137). The principle of CNA detection in cell-free DNA from individuals with cancer is the same as in NIPT (92). Even though not intended, NIPT incidentally detects cancer-associated CNAs from pregnant women. Most often haematological malignancies but also solid tumours have been discovered (109).

Bioinformatic tools are used to analyse the results from sWGS. They normalize the number of reads at each genomic position by statistical methods. There are a number of softwares freely available, and most of them normalize the reads in the sample with help from a reference set of samples (138). WISECONDOR (WIthin-SamplE COpy Number aberration DetectOR, https://github.com/VUmcCGP/wisecondor) identifies optimal reference regions for each part of the chromosome, based on a set of normal samples. It calculates a z-score based on comparison between the sample regions and the reference region, within the same sample (139). WISECONDOR is, after our proof-of-principle study (Study III in this thesis), used in the NIPT workflow at the Department of Clinical Genetics, Karolinska University Hospital, Stockholm, Sweden. It has recently been improved in WISECONDOR X, with shorter analysis time (138). ichorCNA (https://github.com/broadinstitute/ichorCNA), another bioinformatic CNA tool but with initial intended use in cancer patients, utilises a probabilistic model (Hidden Markov Model) to predict CNAs and estimate a tumour fraction. The reads are normalized based on GC-content (the bases in the DNA), mappability biases (reads that could come from different positions withing the genome), and with a reference sample set (119).

The software outputs are lists of potential CNAs found within the plasma sample, and they are assessed, for instance compared to a tissue sample, before their association to cancer can be established. Not all of them are truly cancer-associated. They could be germline, or technical artifacts in regions which are hard to sequence, for instance (119).

### 1.6.4 Too little DNA is the major obstacle

The amount of input DNA is usually the greatest challenge when it comes to ctDNA analysis. In a test tube containing 10 ml blood, there is usually not more than 4 ml plasma. From this, about 40 ng of cell-free DNA can be isolated. This corresponds to the amount of DNA within 6000 normal cells. The ctDNA fraction varies according to tumour type and stage. Given a ctDNA fraction of 0.1%, only 6 molecules per tube would contain the DNA segments that are the most interesting to analyse. The lower the input DNA and the lower the ctDNA fraction, the lesser are the chances that there even exists one cancer-associated genetic change in a blood sample (98).

Some tumours shed more ctDNA than others and the later the stage, the higher the tumour fraction in general (140). There is an increasing number of studies proving the clinical utility and capability of cell-free DNA analysis in late stage cancers, when the ctDNA levels are

generally higher, but in the early stage cancers or in the cancer screening setting, the results are not yet as promising (98) (109).

To overcome the obstacles associated with low ctDNA fraction, technical development is needed. Increasing the blood volume analysed potentially increases the detection rate. Also, more targeted SNVs can be included. The same blood volume is required to detect one single SNV with a variant allele frequency (VAF, number of reads with the specific SNV divided by the total number of reads at that position) of 0.1% as of detecting any of ten SNVs with variant allele frequencies of 0.01%. It seems that a wider gene-panel (targeting 10,000 SNPs, which is more than what is most often used), with a sequencing depth (reads per genomic position) of only 20-50X, is possible (103). Another suggested solution is the use of unique molecular identifiers, which could help minimizing the error associated to technical issues in MPS. The unique molecular identifiers are introduced as identity markers for each DNA molecule before the amplification step in the sequencing process, and they enable exclusion of sequencing artefacts (141).

### 1.6.5 Clonal haematopoiesis is another obstacle

Genetic aberrations found in cell-free DNA do not necessarily originate from cancer. It is known that blood cells undergo clonal haematopoiesis (multiplication of cells from one single cell, all carrying the same genetic changes). SNVs in a specific set of genes, for instance *TP53* and *DNMT3A*, have been described. Clonal haematopoiesis is uncommon in young individuals (<1% before 40 years of age) but occurs with increasing frequency in older individuals (10-15% after 70 years of age). There is an increased risk for the cells displaying clonal haemopoietic markers to give rise to haematological cancer, but most do not (142). Not only SNVs but also CNAs can be present in cell-free DNA from healthy individuals. In non-cancer populations, 0.5-1% have been shown to carry somatic CNAs in their blood, with increasing numbers in older cohorts (143) (144). The malignancy potential of these CNAs is not yet known and there are no clinical guidelines on how to follow individuals carrying detectable CNAs in their blood. To exclude clonal haematopoiesis as the reason for a possibly pathogenic variant found in cell-free DNA, the cell-free DNA variant can be compared to white blood cell DNA variants from the same individual. If the variant is not found in the blood cells, it could be cancer-associated (109). When no DNA analysis of cancer tissue is available as a validation of cancer-associated cell-free DNA variants, white blood cell analysis should be used (133). Most of the genetic alterations in cell-free DNA have now been shown to stem from clonal haematopoiesis of white blood cells, and not from the tumour (145). When it comes to ctDNA as a treatment predictive marker, clonal haematopoiesis has to be considered. Clonal haematopoiesis might occur in the exact same genes as those that are known to be cancer driver genes, and gene targets used in precision medicine (103).

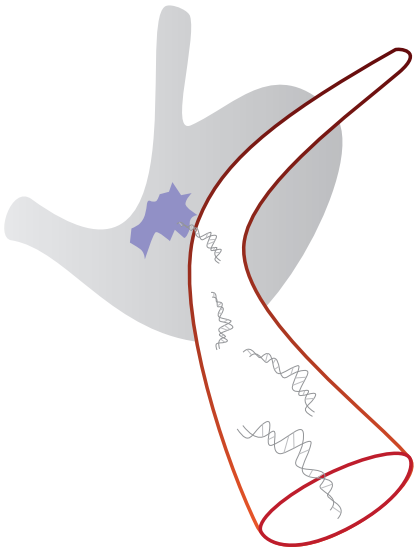
To evaluate the prognostic potential of ctDNA analysis, clonal haematopoietic SNVs need to be excluded from the filtered cell-free DNA variants, as has been described in individuals with oesophageal adenocarcinomas (146).

## 1.7 HOW CAN CELL-FREE DNA BE USED AS A GASTRO-OESOPHAGEAL CANCER BIOMARKER?

### 1.7.1 GEAC is deadly but there are treatment options

Gastric cancer is the fifth most common cancer worldwide, both when it comes to occurrence and cause of cancer death. Oesophageal cancer is the ninth most common cancer type, and the seventh most common cause for cancer death (147). The overall prognosis in GEAC patients depends on stage at discovery. The 5-year survival rate in localized oesophageal cancer is almost 50%, whereas it is only 5% in case of distant metastases (148). The corresponding numbers for gastric cancer are around 70% and 5% (149). Tumour stage at diagnosis is the main survival determinant, but also age, ethnicity, tumour localisation and treatment choice (12). Initially, symptoms are often subtle and unspecific. The diagnostic procedures include gastroscopy with tissue biopsies, computer tomography of the thorax and abdomen, and in some cases other forms of imaging (150). Barrett's oesophagus (changes of the squamous cell epithelium in the lower part of the oesophagus into an epithelium more like that in the stomach) increases the risk for oesopageal adenocarcinoma. There is no single genetic marker predicting which Barrett's oesophagus will develop into cancer, but CNAs are associated with an increased risk for malignant progression and are common in oesophageal adenocarcinomas (151). In Sweden, the curative treatment for GEAC typically includes surgery, chemotherapy, and radiotherapy (150).

Gastric tumours showing high expression of HER2 and PD-L1 might be more aggressive, and MSI seems to be correlated to a better prognosis (152). There are not yet any international consensus guidelines on molecular profiling and actionable genetic aberrations (12). It is, though, known that MSI tumours have worse treatment response to chemotherapy than microsatellite stable tumours (153), but a higher clinical benefit from immunotherapy (152). Also, if *HER2* (*ERBB2*) gene amplifications are detected, anti-HER2 treatment (monoclonal antibodies) can be considered (154). Other potentially actionable genetic aberrations (molecular findings indicating a targeted therapy might be possible) in GEAC include for instance amplification of the genes *VEGFR (KDR)* (155), *EGFR*, *MET* (156), and *FGFR2* (157).



*Gastric tumour releasing ctDNA into the blood stream.*

### 1.7.2 Cell-free DNA analyses are not good enough for GEAC patients yet…

In GEAC, there are not yet enough studies to justify implementation of cell-free DNA analysis as a clinical biomarker. We do not know enough of its physiology and how to handle the information we get from a test. Also, the amount of ctDNA released from the cancer naturally limits the detection rate (124), as discussed above. As GEACs do not harbour one specific genetic aberration but usually have CNAs spread all over the genome (15) (10) (16) (17), whole-genome sequencing technique analysing CNAs might be suitable for detection in cell-free DNA.

### 1.7.3 …but there are potential clinical applications

Potential clinical applications for cell-free DNA analysis in GEAC care are early detection, detection of minimal residual disease or recurrence, prognostic prediction, and guide for treatment.

Early GEAC detection by cell-free DNA analysis is not yet possible, as is true for most solid tumours (109). Though, analyses in GEAC patients show that the concentration of cell-free DNA is generally increased in GEAC patients compared to healthy controls (107). Methylation pattern analysis of 5hmC in cell-free DNA from a mixed patient cohort of any stage (but a majority high-grade dysplasia up to grade II) of oesophageal cancer, both squamous cell and adenocarcinomas, has a sensitivity of 94% and specificity of 86%. This suggests it might be used as a diagnostic biomarker in the future (110). Other methylation analysis approaches have also shown promising results for GEAC detection with high specificity, although the cohorts include predominantly gastric cancer of high stage (suggesting the tumour burden is larger than it would be in a screening setting) (112) (116). A combination of SNV, CNA and protein marker (proteins occurring in the blood stream with suggested different levels in cancer patients compared to control individuals) analysis has been suggested for detection of early stages of GEAC, with a potential detection rate of 80% for cancer-associated aberrations (129).

To distinguish Barrett's oesophagus from GEAC, methylation pattern (158) and deletion (159) analyses in cell-free DNA have been suggested. However, the exact correlation between early cancer-driving events, cancer onset, and cancer aggressiveness is not yet understood (151), and Barrett's oesophagus can harbour the same methylation profile as GEAC (160).

Minimal residual disease (cancer-associated cells or their DNA after cancer treatment) monitoring and recurrence detection in cell-free DNA are limited, just as early diagnostic detection, by low ctDNA levels. Minimal residual disease of GEAC can be detected in cell-free DNA by an individual broad gene panel, including SNVs and CNAs already identified in the tumour tissue (161). Recurrence of GEAC can in some cases be discovered earlier by cell-free DNA analysis than by computer tomography (162) (163), and methylation analysis of selected gene promoters seem to be able to detect early stages of GEAC (113).

Prognostic prediction in GEAC patients by cell-free DNA analysis is possible. In general, detectable cancer-associated SNVs or aberrant methylation in ctDNA is correlated to a worse prognosis (shorter time to relapse and shorter overall survival time) than when none are found (133) (134) (135) (146) (114) (115). Combined analyses of SNVs and

amplifications specific to each study participant with GEAC seem to indicate that detectable cancer-associated aberrations after surgery predicts a shorter overall survival than when no aberrations are found (progression-free survival of 190 days and 934 days, respectively). Also, the higher the VAF the shorter overall survival (136). Measurements of total cell-free DNA concentration can potentially also be used as a biomarker for tumour burden, as it is significantly higher in GEAC patients with advanced gastric cancers than in healthy controls (108). Cancer-associated structural rearrangements are detectable in gastric tumour tissue by sequencing. If they are also found in ctDNA from the same individual after surgery, the prognosis is worse than in the case of undetectable levels (164). Whole-genome CNA analysis using sWGS on cell-free DNA from individuals with late stage cancers may be a cost-effective measurement of disease burden. It is applicable across different cancer types when the ctDNA fractions are high enough (165) (166).

To guide GEAC treatment, *HER2* analysis in ctDNA might be applicable. It is suggested as a repeatable and less invasive detection method for *HER2* gene amplification, especially relevant for individuals with gastric and colorectal cancer (130). *HER2* gene amplifications can be detected in ctDNA after surgery by digital droplet PCR. Even though no *HER2* amplification is present in the diagnostic tumour tissue sample, it may occur during treatment and at recurrence (132). *EGFR* gene amplifications can also be detected in ctDNA, and their resistance mechanisms can be studied (167). Tissue and cell-free DNA *MET* gene amplification detection is described to correlate in 90% of individuals with GEAC (168).

### 1.7.4 Molecular classification of GEAC is possible in cell-free DNA

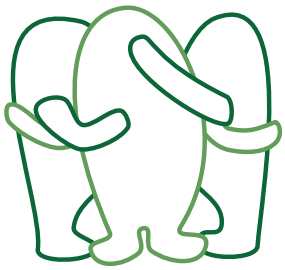
The Cancer Genome Atlas Research Network molecular sub-divisions of GEAC (7), see above, can so far partly be investigated in cell-free DNA. MSI detection by sequencing of cell-free DNA using a pre-set panel corresponds well to detection rates in tissue samples, both in gastric and oesophageal cancer patients (169) (170). EBV subtype detection in individuals with gastric cancer is sometimes possible, especially in case of active infection and large primary tumours. Indirect markers of the EBV subtype can also be used, such as genetic methylation markers recurrent in that specific subtype (152). A chromosomal instability score can be calculated from WGS data of CNAs in cell-free DNA from gastric cancer patients. This score would be high when CIN tumours are present, and could possibly predict and monitor therapeutic response (171). Assuming it is possible to feasibly detect the GEAC subtypes MSI, EBV and CIN, the genomically stable subtype would be annotated in the remaining samples, supported by *CDH1* and *RHOA* SNVs (7) (11).

Tumour mutational burden (TMB) can also potentially be investigated in cell-free DNA. In tissue, it is defined as the number of coding SNVs and indels per megabasepair of the targeted genome. High TMB, associated to higher response rate to cancer immune therapy, co-occurs with MSI in gastrointestinal cancers (172). TMB can also be measured in cell-free DNA, for instance by commercial kits, sometimes being a better representative of the total mutational load than a tissue biopsy (173) (174). TMB in combination with copy number load and *HER2* amplification status might detect GEAC recurrence before computer tomography screening (23).

# 2 RESEARCH AIMS

## 2.1 OVERALL AIM

To improve oncogenetic counselling, by increasing the knowledge about who should be offered genetic tests and what genes to test, and to improve oncology treatment decision-making, by developing a clinical pipeline for a new cancer biomarker, cell-free tumour DNA.



## 2.2 STUDY I

Explore the genetic cause for a suspected highly penetrant, autosomal dominant rectal cancer syndrome in one family.

## 2.3 STUDY II

Evaluate the yield of performing massive parallel sequencing in patients with three or more primary tumours.

## 2.4 STUDY III

Develop a clinical method for detecting cancer-associated CNAs in cell-free tumour DNA from patients with gastro-oesophageal cancer.

# 3 MATERIALS AND METHODS

## 3.1 PATIENT COHORTS

The participants in Study I and Study II were recruited at the Department of Clinical Genetics, Karolinska University Hospital, Stockholm, Sweden, after being referred for oncogenetic counselling, either personally or through a family member. The family in Study I included multiple cases of gastrointestinal cancer, and six family members with rectal cancer or large rectal tubulovillous adenomas with high grade dysplasia were included in the genetic analyses. All participants in Study II either had three primary tumours, with the first one before 60 years of age, or four primary tumours at any age.

The participants in Study III were recruited at the Department of Gastric Upper Abdominal Diseases at Karolinska University Hospital, Stockholm, Sweden, after being referred for gastric cancer treatment. All patients with a verified gastro-oesophageal cancer diagnosis were offered participation in the study.

## 3.2 SEQUENCING AND VARIANT FILTRATION

### 3.2.1 Study I

Germline DNA was isolated from peripheral blood in EDTA tubes. The DNA was sequenced using the Illumina platform and WES/WGS was performed at Clinical Genomics, SciLifeLab, Stockholm, Sweden. Only exonic or splicing variants occurring in all six family members were included. Variants with a minor allele frequency above 1% reported in publicly available cohorts, or more common in the public cohorts than in the local colorectal cancer cohort, were excluded. Also, variants occurring in the local breast cancer cohort were excluded.

### 3.2.2 Study II

The same pipeline for DNA isolation, sequencing and bioinformatics as in Study I was applied. An in silico gene panel containing genes with a known connection to tumour development was used. In the filtering steps, only exonic or splice site variants were included and synonymous variants were only included if they were predicted to affect splicing. Variants with a minor allele frequency above 0.1% reported in publicly available cohorts were excluded. Also, variants reported to be benign were excluded, and those with an allele fraction below 30% were considered somatic and not included.

For the expression analysis of the *MEN1* gene in participant E from Study II, RNA was isolated from a PAXgene tube. cDNA was prepared from the RNA, and sequenced with specific primers for the break between exon 3 and exon 4 of the *MEN1* gene.

### 3.2.3 Study III

Cell-free DNA was isolated from peripheral blood in Streck-tubes. 1 ml of plasma from each participant was included in library preparation and sequencing, which was performed at Clinical Genomics, SciLifeLab, Stockholm, Sweden, using the same clinical pipeline as the non-invasive prenatal diagnostics. The reference set was a locally collected cohort of pregnant women, analysed previously in the same pipeline as the plasma samples from the

participants. Single-end (sequencing of each fragment from just one end, not both) sWGS was used, aiming at a coverage of 0.2X. The cell-free DNA sequencing reads were converted and analyzed using WISECONDOR (https://github.com/VUmcCGP/wisecondor) and ichorCNA (https://github.com/broadinstitute/ichorCNA). In WISECONDOR, CNAs with a z score >4.95, and an effect size (difference in target bin sequencing coverage) ≥1.5% were called. In parallel, the sequencing data was analysed with ichorCNA. A bin size of 500kb and default settings without subclonal analysis were used, and, in accordance with the WISECONDOR analyses, only calls with an effect size ≥1.5% (converted from the reported segmental log2 ratio) were included. Both a targeted approach, only analysing 61 regions with recurrent CNAs in GEAC, and a whole-genome view was tried.

All calls were manually curated, with exclusion of calls from chromosome X, Y, and 19, recurrent calls from segmental duplication regions, or variable centromere regions, and likely germline variant, and calls present in the reference set. A CNA was classified as verified if it was also detected by array-CGH in the paired tissue, and unverified if it was not.

### 3.3 ARRAY-CGH

From the frozen tumour tissue in Study III, sampled as a gastroscopy biopsy or during surgery, DNA was isolated. After oligo array-CGH analysis, the calls were visualized and further analysed in Cytosure Interpret Software (Oxford Gene Technology, Begbroke, UK). CNAs not overlapping with previously described CNAs in our internal database (~8000 samples) or published data sets, and a log2 ratio not matching a germline variant, were considered cancer-associated.

# 4 RESULTS AND DISCUSSION

## 4.1 STUDY I

Six potentially disease-causing variants in the genes *CENPB, ZBTB20, CLINK, LRRC26, TRPM1,* and *NPEPL1* were found. They were all missense, and all of them occurred, although in low frequencies, in population databases of healthy individuals. The five variants were classified according to the ACMG criteria as of unknown significance, and further studies will be needed to validate or discard their potential involvement in a rectal cancer syndrome.

We believe the most interesting variant in the list is the one in the *CENPB* gene. That gene might be involved in the WNT signaling pathway, which is believed to be hyperactivated in almost all colorectal cancers.

Genetic counselling has been offered to all family members, and they are recommended endoscopic surveillance programs. It is not yet, though, possible to offer predictive testing in this family. From the pedigree, a monogenic cancer syndrome with an autosomal dominant inheritance pattern would be suspected, and we aimed at finding a pathogenic variant in a gene associated to a high-risk cancer syndrome. It is, however, possible that the potential pathogenic variant is situated within the non-coding parts of the genome, which we have not investigated. Also, there could be many different variants in different genes, together increasing the risk for cancer in the family but only causing a minor increased risk separately.

## 4.2 STUDY II

In two out of the ten participants, the pathogenic variants causing the tumour syndromes were found.

In a woman with a clinical diagnosis of MEN1, we found a synonymous variant in the *MEN1* gene. This variant had already been found during clinical testing before the start of the study, but it was discarded and not considered disease-causing. However, when using software prediction tools, we could see that the variant might affect splicing. In an RNA analysis, we showed that the variant resulted in activation of a new splice site, which led to a deletion of 14 nucleotides at the end of exon 3, and a frameshift in the RNA transcript. The protein, if the RNA would be translated, would therefore be truncated. In this participant, we also found two variants considered risk factors for cancer.

In a woman with colorectal cancer and microsatellite instability detected in the tumour tissue, we found another synonymous variant in the *MLH1* gene. This variant, too, had been detected in a clinical analysis before, but not acknowledged as disease-causing. When we performed our study, two reports showing this was an epimutation had been published. We could therefore classify it as likely pathogenic and disease-causing.

In the other participants, no known disease-causing variants were found, but we found three pathogenic incidental findings. These are connected to recessive cancer syndromes, and since we did not find another pathogenic variant in the genes, they were not considered disease-causing.

The two participants carrying disease-causing pathogenic variants got updated genetic counselling, and predictive testing has been offered in their families. We conclude that re-analysis of already performed genetic investigations should be considered in the clinic, especially when clinical criteria for a specific tumour syndrome are met. New studies are published, the prediction tools evolve, and the bioinformatic support improves over time. The chances of finding a disease-causing pathogenic variant might improve.

### 4.3 STUDY III

Of the 44 participants included, 26 (59%) had detectable cancer-associated CNAs in the tissue sample. In 14 individuals, amplifications could be seen, two of which can be considered potentially clinically actionable (in the genes *HER2* and *EGFR*). The targeted approach detected fewer cancer-associated CNAs than the genome-wide approach (in 11 versus 17 participants in total).

Of the individuals with cancer-associated CNAs in the tissue sample, 14 (54%) also had detectable cancer-associated CNAs in cell-free DNA. WISECONDOR had a higher detection rate than ichorCNA. Potentially clinically actionable amplifications were detected in the genes *VEGFA, EGFR*, and *FGFR2*.

CIN was detected in 15 individuals, one was based solely on the cell-free DNA analysis.

Three participants had cancer-associated CNAs detectable in the cell-free DNA but not in the tissue sample. However, all three had received neoadjuvant chemotherapy before the tissue sampling, and it is reasonable to believe these CNAs would have been detectable in the tissue at the time of plasma sampling.

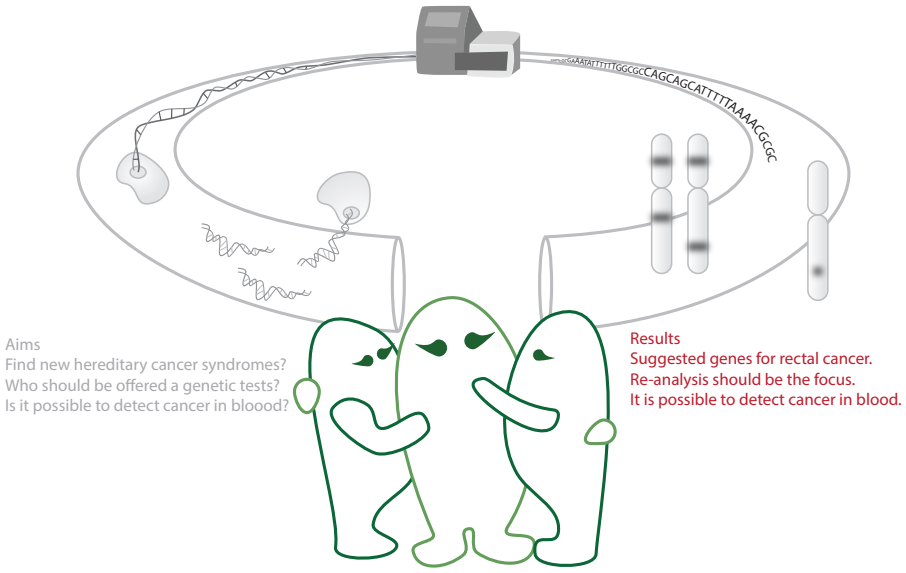
ichorCNA estimates a ctDNA fraction when it is at least 3%, and there is at least one amplification and one deletion event, both larger than 100Mb. In 12 samples, ichorCNA calculated a ctDNA fraction above 0. The highest fraction was 8%, but we saw no correlation between clinical characteristics and ctDNA fraction or effect size.

To find a threshold for cancer-associated CNA calls in cell-free DNA, when tissue status is unknown, we plotted the CNA length against the effect size. We could see that all WISECONDOR CNAs >30Mb long and with an effect size >3% (or 5-30Mb and effect size >5%) were tissue-verified cancer-associated CNAs. For ichorCNA, the threshold would be 30Mb in size and an effect size >3% (or size 2-30Mb and effect size 5%). After this definition, three CNA calls in cell-free DNA were classified as cancer-associated even though they were not found in the tissue by the stringent filtering process.

This study showed that we could detect cancer-associated CNAs in cell-free DNA for GEAC patients, with a detection rate matching other published studies for this cancer type. The clinical pipeline we have set up for the NIPT workflow can be used in the oncogenetic field. We will be able to detect cancer-associated CNAs in the patients even though there is no tumour tissue sample available. In those cases, clonal haematopoiesis needs to be considered, and an additional analysis of peripheral blood DNA would be recommended. In gastric cancer, as well as many other solid tumours, the ctDNA fraction is generally low, which decreases the detection rate regardless of which technical detection method is used (109).

# 5 CONCLUSIONS

In study I, we suggest pathogenic variants in six different genes could increase the risk for rectal cancer. In study II, we found that re-analysis of already performed genetic tests might be beneficial in the oncogenetic setting, especially if a patient fulfills the testing criteria for a specific gene. In study III, we show that it is possible to use our clinical pipeline to detect cancer-associated genetic changes in a blood test from patients with gastro-oesophageal cancer. These findings have the potential to improve the oncogenetic counselling, and to lay the groundwork for use of a new cancer biomarker at our clinic.



# 6 POINTS OF PERSPECTIVE

## 6.1 WHY IS THERE MISSING HERITABILITY IN CANCER SYNDROMES?

The entire heritability of cancer has not been explained by monogenic syndromes and polygenic susceptibility inheritance, as of now. Some argue we might still find monogenic causes for up to half of the missing heritability in colorectal cancer, for instance (175). Study I is an example of ways of trying to find this type of missing heritability. However, highly-penetrant pathogenic variants in cancer syndrome genes might not be more common than 9% (97/1040), when screening cancer patients in general (176). It is possible that mainly ultra-rare variants remain to be discovered, and they are not easily validated in other families than those already known to carry the variants (177) (178). We also lack knowledge about the spectrum of cancers connected to specific genes. With increasing use of WGS/WES both in the clinic and in the research setting, pathogenic variants in genes not traditionally connected to the cancers occurring in the family are discovered. This will expand and evolve our understanding of the molecular background to the cancers, and in some cases testing criteria for cancer syndromes will be adapted. An example of this is the *CDH1* gene, traditionally thought to be connected to diffuse gastric cancer, but the newest guidelines acknowledge its role in isolated hereditary lobular breast cancer (78).

Genome-wide association studies, identifying common genetic variants with a cumulative important increased risk for cancer, indicate that up to 15% of the colorectal cancer heritability might be attributed to these low-risk variants (179) (180) (181) (182) (177). However, not only SNPs but also larger structural rearrangements should perhaps be included in this approach of finding low-risk genetic variants (177).

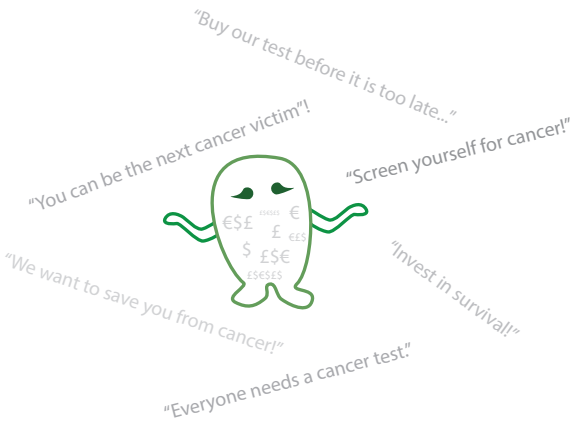
In case of somatic mosaicism, the pathogenic variant is only present in certain cells. Then, it might not be discovered in a peripheral blood sample. And epimutations are also important to consider. It is reasonable to believe that more epimutations than those already known might be discovered in the future (177). Indeed, in Study II, we found a disease-causing epimutation not discovered by the clinical analyses. Technical limitations of WGS/WES should also be considered potential reasons for the missing heritability, as well as non-genetic shared environmental factors.

Not surprisingly, more studies are needed. The validation of genetic variants already discovered is also important. WGS/WES generate variants that could be connected to an increased risk for cancer, and functional analyses are important tools for actually showing, or rule out, this. Functional analyses to validate genetic variants include for instance creating an organoid (3D cellular models of a tissue) gene-modified by CRISPR-Cas (the gene knife), and perform analyses of enzyme activity, RNA expression, and protein-protein interactions in this model (183). For a new gene to be implemented in a clinical gene panel, knowledge about how to act when a pathogenic variant is found is needed, just as much as a validation of the cancer-connection. A potential next step after Study I would be to perform functional analyses of the *CENPB* gene.

## 6.2 WHEN SHOULD WE USE CELL-FREE DNA ANALYSES?

### 6.2.1 Politics and commercial interests cannot be neglected

The European Medicines Agency and FDA (Food and Drug Administration in the US) have recommended some liquid biopsy tests as companion diagnostic tests (analyses which help matching the patient's specific tumour to a treatment), and for tumour profiling (characterization) or detection (184) (185). There are initiatives for collecting large datasets of plasma samples from cancer patients and also from healthy individuals, mostly with the focus on developing a biomarker for finding the cancer early. A promising cancer screening method is a cell-free DNA methylation pattern assay, which both seems to be sensitive and sometimes able to tell from which organ the cancer-associated aberrations stem (186). The efficacy of this approach is controversial. Even though critics argue the test methods are not validated for large asymptomatic populations (187), NHS (National Health Service in Great Britain) has started a clinical trial for assessing the blood test in 140,000 individuals (188). When it comes to treatment choices and resistance tracing, the efficacy of cell-free DNA analysis is usually higher. Specific aberrations in the *EGFR* gene, for instance, leading to higher response rate to *EGFR* targeted therapy in lung cancer patients, can be detected by a commercial, FDA approved, test (98). It is clear that liquid biopsy applications are not only a medical concern anymore, but also a commercial and political issue.




### 6.2.2 Clinical implementation is a complex process

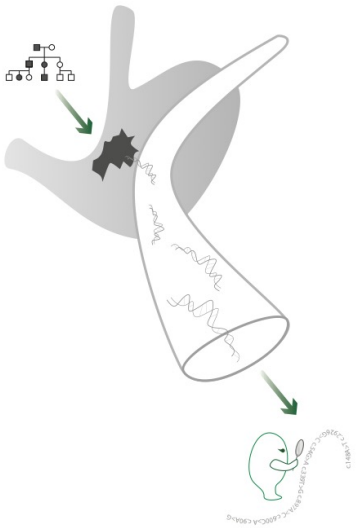
A potential cell-free DNA analysis with appropriate analytical and clinical validity might be considered for clinical implementation. Both the testing routine, the technical aspects of cell-free DNA analysis, the bioinformatic steps, and the interpretation of the results then need to be taken into account. This is a complex process, involving different personnel, and it is important to note that for many care centers a totally new routine is necessary (103). When considering the utility of a cell-free DNA analysis, the clinical context must be considered. Therefore, using already existing expertise and methods has its advantages. This is why we thought it important to perform Study III, as a proof-of principle and stress-testing of the current pipelines at our institutions.

### 6.2.3 Cell-free DNA analyses are likely to be clinical routine soon

ctDNA is an attractive candidate as GEAC biomarker. It is a minimally invasive and readily accessible way of receiving information on a potential tumour and its genetic characteristics. In ten years, it is likely the methylation analyses can be used clinically, both to detect GEAC early and to get prognostic information (111). Perhaps a combination of SNV, CNA, tumour mutational burden, and methylation analyses will make it possible to detect GEAC even when the ctDNA fraction is low, for screening, minimal residual disease surveillance, and early recurrence detection. It is already possible to detect amplifications in for instance the *HER2* gene, and evaluate microsatellite status in ctDNA from GEAC patients (132) (130) (167) (169) (170), and as this information can guide treatment choices, perhaps this is also two of the first applications which will be used world-wide clinically.

## 6.3 CAN THE SUB-STUDIES BE COMBINED INTO ONE CLINICAL APPLICATION?

The risk of cancer is generally high in carriers of a pathogenic variant in a gene associated to a cancer syndrome. Surveillance programs are most often recommended. In some cases, such as hereditary diffuse gastric cancer, carriers of pathogenic variants in the *CDH1* gene are usually recommended risk-reducing gastrectomy (78). But not all carriers will get a life-threatening cancer, and a gastrectomy is a taxing experience. There are ongoing studies, including individuals with hereditary cancer syndrome but no known cancer, who undergo regular collection of cell-free DNA, to see if this might detect the cancers at an early stage (189). Perhaps, in the future, analysis of cell-free DNA in carriers of a pathogenic variant in the *CDH1* gene could detect cancer occurrence early enough. Then, only those with an actual gastric cancer would have to undergo gastrectomy. For this to happen, we need to know more about hereditary cancer syndromes and their genetic causes, and we also need better cell-free DNA analysis techniques, and more knowledge of the nature of cell-free DNA. The sub-studies in this thesis address these different issues.



*Inherited CDH1 pathogenic variant → diffuse gastric cancer → detectable ctDNA?*

# 7 ACKNOWLEDGEMENTS

The research we perform would not have been possible if it were not for the **study participants**, actively taking part in the research and generously sharing their story.

**Emma Tham**, thank you for introducing me into the world of genetics! How you manage to be on top of all your projects, both at the clinic and on the research side, I do not know. But I do know I am grateful to you, for being a better role model than anyone could ever hope for. You are not only a fantastic doctor and researcher, but also a caring colleague and friend.

I have had the privilege to be supervised by a dream team. **Mats Lindblad**, with inspiring care for his patients. I am impressed by your clinical skills and how you are able to incorporate research into your everyday clinical practice. You see the whole picture, from the tiny molecules to the psychological aspects of medicine. **Magnus Nordenskjöld**, with unsurpassed experience and knowledge in the field, but still a way of keeping things simple. Thank you for showing trust in me right from the start, and for sharing horrible and educational anecdotes from the research community. **Kristina Lagerstedt-Robinson**, with great and broad knowledge and impressive speed. You are an excellent teacher, and you manage to make genetics seem almost logical, and definitely fun. **Annika Lindblom**, who refreshingly demands me to first think and then act. Your research carrier is impressive, and you are an inspiration to us all. **Inger Öhman**, my mentor, with toxic enthusiasm and love for science. Thank you for your professional advice and for encouraging me to try to be one step ahead at all times.

Many thanks to **all clinical laboratory geneticists** at the clinic, for hosting a warm and professional atmosphere. You explain complex genetic mechanisms in ways understandable for unexperienced new doctors at the clinic, and never ever show any tiredness of having to do this. A special thanks to you who have been part of the studies in this thesis: **Agne Liedén**, **Jesper Eisfeldt, Håkan Thonberg, Daniel Nilsson, Ellika Sahlin, Hafdís Helgadóttir**. I hope I will get the chance to do more research as well as clinical work with you.

I also want to thank all **biomedical analysist**, for actively and intelligently taking part in the research performed at out clinic. Thank you for not laughing at me when I ask stupid questions! **Isabel Neira** and the **sample reception team** at the lab, thank you for making this clinical research possible.

My fellow **residents**, thank you for making it fun and interesting to go to work. You are honest, friendly, and hard-working. I am proud to be part of our team and I look forward to continuing working together with you. A special thanks to **Cecilia Arthur**, my work-BFF. Ever since we first worked together as medical secretaries, you have been my idol.

All **specialist** at out clinic, thank you for your support and for showing how important and enjoyable it is to be a clinical geneticist. It is a privilege to be taught by you. Even on international conferences, I find the talks and comments you make the most well-put and

interesting. You make education part of the everyday clinical work, and I am grateful for the time you spend on helping me become a better doctor. **Maria Johansson Soller, Peter Gustafsson, Richard Rosenquist Brandell**, **Erik Iwarsson**, a special thanks to you for facilitating my role as both PhD student and MD.

**Isabelle Krakowski, Meis Omran, Alexandra Grahn**, and my other research school classmates, you do inspiring work that matters for the patients.

Thank you, **Fulya Taylan** and **Nina Jäntti**, for guiding me through my first days of wet-labbing. To my **research group Clinical genetics**, thank you for always providing interesting comments and discussions during seminars and journal clubs. **Berit Sunde**, thank you for being part of our team and making the cell-free DNA studies possible.

I am grateful for the support from the **administrative staff** at MMK, Karolinska Institutet and the financial support from **Region Stockholm**, through my employment "Forskar-ST".

**Ashwini Jeggari, Hassan Foroughi** and the other skilled and dedicated scientists at **Science for Life Laboratory**, I look forward to more collaborations in the future!

**Johan Wallander** and **Kerstin Lindholm Wallander**, my wise parents. Thank you for being loving, sensible and encouraging life coaches. **Hans Lindgren**, you always find a fresh perspective. **Peter Wallander** and **Malin Wallander**, thank you for being there for me always. Growing up with you has been an amazing adventure. Farmor **Ingeborg Wallander**, you are the coolest. **Katarina Gyllensten** and **Salustio Cuba Gamarra**, I am grateful and proud that I get to be part of your family. **Bönorna, Agenterna, Uppsala-people**, thank you for your friendship and support.

**Laura Cuba Wallander** and **Alf Cuba Wallander**, if it were not for you, I would be way more well rested and way way less happy. **Ollanta Cuba Gyllensten**, pick (almost) any romantic love song title and it will be true for how I feel about you.

# 8 REFERENCES


1. Hanahan D, Weinberg RA. Hallmarks of cancer: the next generation. Cell. 2011;144(5):646-74.
2. Strachan T, Read AP. Human Molecular Genetics, fifth edition. Boca Raton, FL, USA: CRC Press, Taylor & Francis Group, Garland Science Book; 2019.
3. Shen S, Chen X, Li H, Sun L, Yuan Y. MLH1 Promoter Methylation and Prediction/Prognosis of Gastric Cancer: A Systematic Review and Meta and Bioinformatic Analysis. J Cancer. 2018;9(11):1932-42.
4. Vogelstein B, Papadopoulos N, Velculescu VE, Zhou SB, Diaz LA, Kinzler KW. Cancer Genome Landscapes. Science. 2013;339(6127):1546-58.
5. Nguyen HT, Duong HQ. The molecular characteristics of colorectal cancer: Implications for diagnosis and therapy. Oncol Lett. 2018;16(1):9-18.
6. Boland CR, Koi M, Chang DK, Carethers JM. The biochemical basis of microsatellite instability and abnormal immunohistochemistry and clinical behavior in Lynch syndrome: from bench to bedside. Fam Cancer. 2008;7(1):41-52.
7. Bass AJ, Thorsson V, Shmulevich I, Reynolds SM, Miller M, Bernard B, et al. Comprehensive molecular characterization of gastric adenocarcinoma. Nature. 2014;513(7517):202-9.
8. Tijhuis AE, Johnson SC, McClelland SE. The emerging links between chromosomal instability (CIN), metastasis, inflammation and tumour immunity. Mol Cytogenet. 2019;12:17.
9. Bastians H. Causes of Chromosomal Instability. Recent Results Cancer Res. 2015;200:95-113.
10. Cancer genome atlas research network, Analysis working group, Integrated genomic characterization of oesophageal carcinoma. Nature. 2017;541(7636):169-75.
11. Garcia-Pelaez J, Barbosa-Matos R, Gullo I, Carneiro F, Oliveira C. Histological and mutational profile of diffuse gastric cancer: current knowledge and future challenges. Mol Oncol. 2021.
12. Cancercentrum. Swedish National Treatment Guidelines for oesophagus and gastric cancer. 2020.
13. Koretzky GA. The legacy of the Philadelphia chromosome. J Clin Invest. 2007;117(8):2030-2.
14. Sabapathy K, Lane DP. Therapeutic targeting of p53: all mutants are equal, but some mutants are more equal than others. Nat Rev Clin Oncol. 2018;15(1):13-30.
15. Rocken C. Molecular classification of gastric cancer. Expert Review of Molecular Diagnostics. 2017;17(3):293-301.
16. Pasello G, Agata S, Bonaldi L, Corradin A, Montagna M, Zamarchi R, et al. DNA copy number alterations correlate with survival of esophageal adenocarcinoma patients. Modern Pathology. 2009;22(1):58-65.
17. Schumacher SE, Shim BY, Corso G, Ryu MH, Kang YK, Roviello F, et al. Somatic copy number alterations in gastric adenocarcinomas among Asian and Western patients. Plos One. 2017;12(4).
18. Spira A, Yurgelun MB, Alexandrov L, Rao A, Bejar R, Polyak K, et al. Precancer Atlas to Drive Precision Prevention Trials. Cancer Res. 2017;77(7):1510-41.
19. Wakatsuki T, Yamamoto N, Sano T, Chin K, Kawachi H, Takahari D, et al. Clinical impact of intratumoral HER2 heterogeneity on trastuzumab efficacy in patients with HER2-positive gastric cancer. J Gastroenterol. 2018;53(11):1186-95.
20. Weinberg R. The Biology of Cancer. New York, NY, USA: Garland Science; 2014.
21. Debniak T, Lubinski J. Principles of genetic predisposition to malignancies. Hered Cancer Clin Pract. 2008;6(2):69-72.

22. Mucci LA, Hjelmborg JB, Harris JR, Czene K, Havelick DJ, Scheike T, et al. Familial Risk and Heritability of Cancer Among Twins in Nordic Countries. Jama. 2016;315(1):68-76.
23. Zhang YD, Hurson AN, Zhang H, Choudhury PP, Easton DF, Milne RL, et al. Assessment of polygenic architecture and risk prediction based on common variants across fourteen cancers. Nat Commun. 2020;11(1):3353.
24. Fanfani V, Citi L, Harris AL, Pezzella F, Stracquadanio G. The Landscape of the Heritable Cancer Genome. Cancer Res. 2021;81(10):2588-99.
25. Kachuri L, Graff RE, Smith-Byrne K, Meyers TJ, Rashkin SR, Ziv E, et al. Pan-cancer analysis demonstrates that integrating polygenic risk scores with modifiable risk factors improves risk prediction. Nat Commun. 2020;11(1):6084.
26. Sud A, Turnbull C, Houlston R. Will polygenic risk scores for cancer ever be clinically useful? NPJ Precis Oncol. 2021;5(1):40.
27. World Health Organisation (WHO) Cancer Prevention, available at https://www.who.int/health-topics/cancer#tab=tab_2. Accessed on 09-05-21. [
28. Czene K, Lichtenstein P, Hemminki K. Environmental and heritable causes of cancer among 9.6 million individuals in the Swedish Family-Cancer Database. Int J Cancer. 2002;99(2):260-6.
29. Lichtenstein P, Holm NV, Verkasalo PK, Iliadou A, Kaprio J, Koskenvuo M, et al. Environmental and heritable factors in the causation of cancer - Analyses of cohorts of twins from Sweden, Denmark, and Finland. New England Journal of Medicine. 2000;343(2):78-85.
30. DeRycke MS, Gunawardena S, Balcom JR, Pickart AM, Waltman LA, French AJ, et al. Targeted sequencing of 36 known or putative colorectal cancer susceptibility genes. Mol Genet Genomic Med. 2017;5(5):553-69.
31. AlDubayan SH, Giannakis M, Moore ND, Han GC, Reardon B, Hamada T, et al. Inherited DNA-Repair Defects in Colorectal Cancer. Am J Hum Genet. 2018;102(3):401-14.
32. Yurgelun MB, Kulke MH, Fuchs CS, Allen BA, Uno H, Hornick JL, et al. Cancer Susceptibility Gene Mutations in Individuals With Colorectal Cancer. J Clin Oncol. 2017;35(10):1086-95.
33. Ajani JA, D'Amico TA, Almhanna K, Bentrem DJ, Chao J, Das P, et al. Gastric Cancer, Version 3.2016, NCCN Clinical Practice Guidelines in Oncology. J Natl Compr Canc Netw. 2016;14(10):1286-312.
34. Oliveira C, Pinheiro H, Figueiredo J, Seruca R, Carneiro F. Familial gastric cancer: genetic susceptibility, pathology, and implications for management. Lancet Oncology. 2015;16(2):E60-E70.
35. Valle L, Vilar E, Tavtigian SV, Stoffel EM. Genetic predisposition to colorectal cancer: syndromes, genes, classification of genetic variants and implications for precision medicine. J Pathol. 2019;247(5):574-88.
36. Gullo I, van der Post RS, Carneiro F. Recent advances in the pathology of heritable gastric cancer syndromes. Histopathology. 2021;78(1):125-47.
37. Worthley DL, Phillips KD, Wayte N, Schrader KA, Healey S, Kaurah P, et al. Gastric adenocarcinoma and proximal polyposis of the stomach (GAPPS): a new autosomal dominant syndrome. Gut. 2012;61(5):774-9.
38. Li J, Woods SL, Healey S, Beesley J, Chen XQ, Lee JS, et al. Point Mutations in Exon 1B of APC Reveal Gastric Adenocarcinoma and Proximal Polyposis of the Stomach as a Familial Adenomatous Polyposis Variant. American Journal of Human Genetics. 2016;98(5):830-42.
39. Mitsui Y, Yokoyama R, Fujimoto S, Kagemoto K, Kitamura S, Okamoto K, et al. First report of an Asian family with gastric adenocarcinoma and proximal polyposis of

the stomach (GAPPS) revealed with the germline mutation of the APC exon 1B promoter region. Gastric Cancer. 2018;21(6):1058-63.
40. Repak R, Kohoutova D, Podhola M, Rejchrt S, Minarik M, Benesova L, et al. The first European family with gastric adenocarcinoma and proximal polyposis of the stomach: case report and review of the literature. Gastrointest Endosc. 2016;84(4):718-25.
41. Weren RD, Ligtenberg MJ, Kets CM, de Voer RM, Verwiel ET, Spruijt L, et al. A germline homozygous mutation in the base-excision repair gene NTHL1 causes adenomatous polyposis and colorectal cancer. Nat Genet. 2015;47(6):668-71.
42. Adam R, Spier I, Zhao B, Kloth M, Marquez J, Hinrichsen I, et al. Exome Sequencing Identifies Biallelic MSH3 Germline Mutations as a Recessive Subtype of Colorectal Adenomatous Polyposis. Am J Hum Genet. 2016;99(2):337-51.
43. Lynch HT, de la Chapelle A. Hereditary colorectal cancer. N Engl J Med. 2003;348(10):919-32.
44. Nieminen TT, O'Donohue MF, Wu Y, Lohi H, Scherer SW, Paterson AD, et al. Germline mutation of RPS20, encoding a ribosomal protein, causes predisposition to hereditary nonpolyposis colorectal carcinoma without DNA mismatch repair deficiency. Gastroenterology. 2014;147(3):595-8.e5.
45. Broderick P, Dobbins SE, Chubb D, Kinnersley B, Dunlop MG, Tomlinson I, et al. Validation of Recently Proposed Colorectal Cancer Susceptibility Gene Variants in an Analysis of Families and Patients-a Systematic Review. Gastroenterology. 2017;152(1):75-7.e4.
46. Terradas M, Capellá G, Valle L. Dominantly Inherited Hereditary Nonpolyposis Colorectal Cancer Not Caused by MMR Genes. J Clin Med. 2020;9(6).
47. DeRycke MS, Gunawardena SR, Middha S, Asmann YW, Schaid DJ, McDonnell SK, et al. Identification of novel variants in colorectal cancer families by high-throughput exome sequencing. Cancer Epidemiol Biomarkers Prev. 2013;22(7):1239-51.
48. Cicek MS, Cunningham JM, Fridley BL, Serie DJ, Bamlet WR, Diergaarde B, et al. Colorectal cancer linkage on chromosomes 4q21, 8q13, 12q24, and 15q22. PLoS One. 2012;7(5):e38175.
49. Guilford P, Hopkins J, Harraway J, McLeod M, McLeod N, Harawira P, et al. E-cadherin germline mutations in familial gastric cancer. Nature. 1998;392(6674):402-5.
50. Majewski IJ, Kluijt I, Cats A, Scerri TS, de Jong D, Kluin RJ, et al. An α-E-catenin (CTNNA1) mutation in hereditary diffuse gastric cancer. J Pathol. 2013;229(4):621-9.
51. Weren RDA, van der Post RS, Vogelaar IP, van Krieken JH, Spruijt L, Lubinski J, et al. Role of germline aberrations affecting CTNNA1, MAP3K6 and MYD88 in gastric cancer susceptibility. J Med Genet. 2018;55(10):669-74.
52. van Dieren JM, Kodach LL, den Hartog P, van der Kolk LE, Sikorska K, van Velthuysen MF, et al. Gastroscopic surveillance with targeted biopsies compared with random biopsies in CDH1 mutation carriers. Endoscopy. 2020;52(10):839-46.
53. Mi EZ, Mi EZ, di Pietro M, O'Donovan M, Hardwick RH, Richardson S, et al. Comparative study of endoscopic surveillance in hereditary diffuse gastric cancer according to CDH1 mutation status. Gastrointest Endosc. 2018;87(2):408-18.
54. Rocha JP, Gullo I, Wen XG, Devezas V, Baptista M, Oliveira C, et al. Pathological features of total gastrectomy specimens from asymptomatic hereditary diffuse gastric cancer patients and implications for clinical management. Histopathology. 2018;73(6):878-86.
55. Roberts ME, Ranola JMO, Marshall ML, Susswein LR, Graceffo S, Bohnert K, et al. Comparison of CDH1 Penetrance Estimates in Clinically Ascertained Families vs Families Ascertained for Multiple Gastric Cancers. Jama Oncology. 2019;5(9):1325-31.

56. Hansford S, Kaurah P, Li-Chang H, Woo M, Senz J, Pinheiro H, et al. Hereditary Diffuse Gastric Cancer Syndrome CDH1 Mutations and Beyond. Jama Oncology. 2015;1(1):23-32.
57. Figueiredo J, Melo S, Carneiro P, Moreira AM, Fernandes MS, Ribeiro AS, et al. Clinical spectrum and pleiotropic nature of CDH1 germline mutations. Journal of Medical Genetics. 2019;56(4):199-208.
58. Corso G, Montagna G, Figueiredo J, La Vecchia C, Fumagalli Romario U, Fernandes MS, et al. Hereditary Gastric and Breast Cancer Syndromes Related to CDH1 Germline Mutation: A Multidisciplinary Clinical Review. Cancers (Basel). 2020;12(6).
59. Peltomäki P, Olkinuora A, Nieminen TT. Updates in the field of hereditary nonpolyposis colorectal cancer. Expert Rev Gastroenterol Hepatol. 2020;14(8):707-20.
60. Thakker RV, Newey PJ, Walls GV, Bilezikian J, Dralle H, Ebeling PR, et al. Clinical practice guidelines for multiple endocrine neoplasia type 1 (MEN1). J Clin Endocrinol Metab. 2012;97(9):2990-3011.
61. Sakurai A, Suzuki S, Kosugi S, Okamoto T, Uchino S, Miya A, et al. Multiple endocrine neoplasia type 1 in Japan: establishment and analysis of a multicentre database. Clin Endocrinol (Oxf). 2012;76(4):533-9.
62. Tham E, Grandell U, Lindgren E, Toss G, Skogseid B, Nordenskjöld M. Clinical testing for mutations in the MEN1 gene in Sweden: a report on 200 unrelated cases. J Clin Endocrinol Metab. 2007;92(9):3389-95.
63. Brandi ML, Agarwal SK, Perrier ND, Lines KE, Valk GD, Thakker RV. Multiple Endocrine Neoplasia Type 1: Latest Insights. Endocr Rev. 2021;42(2):133-70.
64. SEER (Surveillance, Epidemiology, and End Results Program). Solid Tumour Rules 2018. Available at https://seer.cancer.gov/tools/solidtumor/. Accessed on 09-16-2021. [
65. Bittorf B, Kessler H, Merkel S, Bruckl W, Wein A, Ballhausen WG, et al. Multiple primary malignancies: An epidemiological and pedigree analysis of 57 patients with at least three tumours. European Journal of Surgical Oncology. 2001;27(3):302-13.
66. Salem A, Abu-Hijlih R, Abdelrahman F, Turfa R, Amarin R, Farah N, et al. Multiple Primary Malignancies: Analysis of 23 Patients with at Least Three Tumors. Journal of Gastrointestinal Cancer. 2012;43(3):437-43.
67. Whitworth J, Smith PS, Martin JE, West H, Luchetti A, Rodger F, et al. Comprehensive Cancer-Predisposition Gene Testing in an Adult Multiple Primary Tumor Series Shows a Broad Range of Deleterious Variants and Atypical Tumor Phenotypes. American Journal of Human Genetics. 2018;103(1):3-18.
68. ACOG Committee opinion 793, Hereditary cancer syndromes and risk assessment. The American College of Obstetricians and Gynecologists. Available at: https://www.acog.org/clinical/clinical-guidance/committee-opinion/articles/2019/12/hereditary-cancer-syndromes-and-risk-assessment#:~:text=ABSTRACT%3A%20A%20hereditary%20cancer%20syndrome,syndromes%20exhibit%20autosomal%20dominant%20inheritance. Accessed on 09-18-2021. [
69. Blueprint genetics, Genetic testing for hereditary colorectal cancer, https://blueprintgenetics.com/tests/panels/hereditary-cancer/hereditary-colorectal-cancer-panel/, accessed on 10-04-2021 https://blueprintgenetics.com/tests/panels/hereditary-cancer/hereditary-colorectal-cancer-panel/ [
70. 5 Committee discussion, Molecular testing strategies for Lynch syndrome in people with colorectal cancer, Guidance, NICE. Available at https://www.nice.org.uk/guidance/dg27/chapter/5-Committee-discussion. Accessed on 09-14-2021. [

71. Snowsill T, Coelho H, Huxley N, Jones-Hughes T, Briscoe S, Frayling IM, et al. Molecular testing for Lynch syndrome in people with colorectal cancer: systematic reviews and economic evaluation. Health Technol Assess. 2017;21(51):1-238.
72. Monahan KJ, Bradshaw N, Dolwani S, Desouza B, Dunlop MG, East JE, et al. Guidelines for the management of hereditary colorectal cancer from the British Society of Gastroenterology (BSG)/Association of Coloproctology of Great Britain and Ireland (ACPGBI)/United Kingdom Cancer Genetics Group (UKCGG). Gut. 2020;69(3):411-44.
73. van Leerdam ME, Roos VH, van Hooft JE, Balaguer F, Dekker E, Kaminski MF, et al. Endoscopic management of Lynch syndrome and of familial risk of colorectal cancer: European Society of Gastrointestinal Endoscopy (ESGE) Guideline. Endoscopy. 2019;51(11):1082-93.
74. Rubenstein JH, Enns R, Heidelbaugh J, Barkun A. American Gastroenterological Association Institute Guideline on the Diagnosis and Management of Lynch Syndrome. Gastroenterology. 2015;149(3):777-82; quiz e16-7.
75. Ryan N, Snowsill T, McKenzie E, Monahan KJ, Nebgen D. Should women with Lynch syndrome be offered gynaecological cancer surveillance? Bmj. 2021;374:n2020.
76. Katona BW, Yurgelun MB, Garber JE, Offit K, Domchek SM, Robson ME, et al. A counseling framework for moderate-penetrance colorectal cancer susceptibility genes. Genet Med. 2018;20(11):1324-7.
77. Carvalho J, Oliveira P, Senz J, São José C, Hansford S, Teles SP, et al. Redefinition of familial intestinal gastric cancer: clinical and genetic perspectives. Journal of Medical Genetics. 2020:jmedgenet-2019-106346.
78. Blair VR, McLeod M, Carneiro F, Coit DG, D'Addario JL, van Dieren JM, et al. Hereditary diffuse gastric cancer: updated clinical practice guidelines. Lancet Oncology. 2020;21(8):E386-E97.
79. Sanger F, Coulson AR. A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. J Mol Biol. 1975;94(3):441-8.
80. Metzker ML. Sequencing technologies - the next generation. Nat Rev Genet. 2010;11(1):31-46.
81. IonTorrent, ThermoFisher, sequencing technique. Available at https://www.thermofisher.com/se/en/home/life-science/sequencing/next-generation-sequencing/ion-torrent-next-generation-sequencing-technology.html. Accessed on 10-05-21. [
82. Illumina sequencing by synthesis, available at https://emea.illumina.com/science/technology/next-generation-sequencing/sequencing-technology.html, accessed on 10-05-21 [
83. Roche sequencing technique, available at https://sequencing.roche.com/en/research-application/research/human-genetics.html. Accessed on 10-05-21.
84. Stranneheim H, Lagerstedt-Robinson K, Magnusson M, Kvarnung M, Nilsson D, Lesko N, et al. Integration of whole genome sequencing into a healthcare setting: high diagnostic rates across multiple clinical entities in 3219 rare disease patients. Genome Med. 2021;13(1):40.
85. Pedersen BS, Brown JM, Dashnow H, Wallace AD, Velinder M, Tristani-Firouzi M, et al. Effective variant filtering and expected candidate variant yield in studies of rare human disease. NPJ Genom Med. 2021;6(1):60.
86. Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. Genet Med. 2015;17(5):405-24.

87. Romanet P, Odou MF, North MO, Saveanu A, Coppin L, Pasmant E, et al. Proposition of adjustments to the ACMG-AMP framework for the interpretation of MEN1 missense variants. Hum Mutat. 2019;40(6):661-74.
88. ClinGen, Clinical Genome Resource, Sequence Variant Interpretation. Available at https://clinicalgenome.org/working-groups/sequence-variant-interpretation/. Accessed on 09-12-2021. [
89. Houge G, Laner A, Cirak S, de Leeuw N, Scheffer H, den Dunnen JT. Stepwise ABC system for classification of any type of genetic variant. Eur J Hum Genet. 2021.
90. Landrith T, Li B, Cass AA, Conner BR, LaDuca H, McKenna DB, et al. Splicing profile by capture RNA-seq identifies pathogenic germline variants in tumor suppressor genes. NPJ Precis Oncol. 2020;4:4.
91. Bejjani BA, Shaffer LG. Application of array-based comparative genomic hybridization to clinical diagnostics. J Mol Diagn. 2006;8(5):528-33.
92. Lo YM, Chan KC, Sun H, Chen EZ, Jiang P, Lun FM, et al. Maternal plasma DNA sequencing reveals the genome-wide genetic and mutational profile of the fetus. Sci Transl Med. 2010;2(61):61ra91.
93. Jahr S, Hentze H, Englisch S, Hardt D, Fackelmayer FO, Hesch RD, et al. DNA fragments in the blood plasma of cancer patients: quantitations and evidence for their origin from apoptotic and necrotic cells. Cancer Res. 2001;61(4):1659-65.
94. Rykova EY, Morozkin ES, Ponomaryova AA, Loseva EM, Zaporozhchenko IA, Cherdyntseva NV, et al. Cell-free and cell-bound circulating nucleic acid complexes: mechanisms of generation, concentration and content. Expert Opin Biol Ther. 2012;12 Suppl 1:S141-53.
95. Mouliere F, Chandrananda D, Piskorz AM, Moore EK, Morris J, Ahlborn LB, et al. Enhanced detection of circulating tumor DNA by fragment size analysis. Sci Transl Med. 2018;10(466).
96. Chiu RWK, Heitzer E, Lo YMD, Mouliere F, Tsui DWY. Cell-Free DNA Fragmentomics: The New "Omics" on the Block. Clin Chem. 2020;66(12):1480-4.
97. Snyder MW, Kircher M, Hill AJ, Daza RM, Shendure J. Cell-free DNA Comprises an In Vivo Nucleosome Footprint that Informs Its Tissues-Of-Origin. Cell. 2016;164(1-2):57-68.
98. Heitzer E, Haque IS, Roberts CES, Speicher MR. Current and future perspectives of liquid biopsies in genomics-driven oncology. Nat Rev Genet. 2019;20(2):71-88.
99. Diehl F, Schmidt K, Choti MA, Romans K, Goodman S, Li M, et al. Circulating mutant DNA to assess tumor dynamics. Nat Med. 2008;14(9):985-90.
100. Khier S, Lohan L. Kinetics of circulating cell-free DNA for biomedical applications: critical appraisal of the literature. Future Sci OA. 2018;4(4):Fso295.
101. Parikh AR, Leshchiner I, Elagina L, Goyal L, Levovitz C, Siravegna G, et al. Liquid versus tissue biopsy for detecting acquired resistance and tumor heterogeneity in gastrointestinal cancers. Nature Medicine. 2019;25(9):1415-+.
102. Pittella-Silva F, Chin YM, Chan HT, Nagayama S, Miyauchi E, Low SK, et al. Plasma or Serum: Which Is Preferable for Mutation Detection in Liquid Biopsy? Clin Chem. 2020;66(7):946-57.
103. Ignatiadis M, Sledge GW, Jeffrey SS. Liquid biopsy enters the clinic - implementation issues and future challenges. Nat Rev Clin Oncol. 2021;18(5):297-312.
104. Merker JD, Oxnard GR, Compton C, Diehn M, Hurley P, Lazar AJ, et al. Circulating Tumor DNA Analysis in Patients With Cancer: American Society of Clinical Oncology and College of American Pathologists Joint Review. J Clin Oncol. 2018;36(16):1631-41.

105. Shi XQ, Xue WH, Zhao SF, Zhang XJ, Sun W. Dynamic tracing for epidermal growth factor receptor mutations in urinary circulating DNA in gastric cancer patients. Tumour Biol. 2017;39(2):1010428317691681.
106. Pizzi MP, Bartelli TF, Pelosof AG, Freitas HC, Begnami MD, de Abrantes LLS, et al. Identification of DNA mutations in gastric washes from gastric adenocarcinoma patients: Possible implications for liquid biopsies and patient follow-up. Int J Cancer. 2019;145(4):1090-8.
107. Qian C, Ju SQ, Qi J, Zhao JM, Shen XJ, Jing RR, et al. Alu-based cell-free DNA: a novel biomarker for screening of gastric cancer. Oncotarget. 2017;8(33):54037-45.
108. Pu WY, Zhang R, Xiao L, Wu YY, Gong W, Lv XD, et al. Prediction of cancer progression in a group of 73 gastric cancer patients by circulating cell-free DNA. Bmc Cancer. 2016;16.
109. Barbany G, Arthur C, Lieden A, Nordenskjold M, Rosenquist R, Tesi B, et al. Cell-free tumour DNA testing for early detection of cancer - a potential future tool. Journal of Internal Medicine. 2019;286(2):118-36.
110. Tian X, Sun B, Chen C, Gao C, Zhang J, Lu X, et al. Circulating tumor DNA 5-hydroxymethylcytosine as a novel diagnostic biomarker for esophageal cancer. Cell Res. 2018;28(5):597-600.
111. Huang ZB, Zhang HT, Yu B, Yu DH. Cell-free DNA as a liquid biopsy for early detection of gastric cancer. Oncol Lett. 2021;21(1):3.
112. Guo L, Huang C, Ji QJ. Aberrant promoter hypermethylation of p16, survivin, and retinoblastoma in gastric cancer. Bratisl Lek Listy. 2017;118(3):164-8.
113. Lin Z, Luo M, Chen X, He X, Qian Y, Lai S, et al. Combined Detection of Plasma ZIC1, HOXD10 and RUNX3 Methylation is a Promising Strategy for Early Detection of Gastric Cancer and Precancerous Lesions. J Cancer. 2017;8(6):1038-44.
114. Balgkouranidou I, Karayiannakis A, Matthaios D, Bolanaki H, Tripsianis G, Tentes AA, et al. Assessment of SOX17 DNA methylation in cell free DNA from patients with operable gastric cancer. Association with prognostic variables and survival. Clin Chem Lab Med. 2013;51(7):1505-10.
115. Balgkouranidou I, Matthaios D, Karayiannakis A, Bolanaki H, Michailidis P, Xenidis N, et al. Prognostic role of APC and RASSF1A promoter methylation status in cell free circulating DNA of operable gastric cancer patients. Mutat Res. 2015;778:46-51.
116. Zhang X, Zhang X, Sun B, Lu H, Wang D, Yuan X, et al. Detection of aberrant promoter methylation of RNF180, DAPK1 and SFRP2 in plasma DNA of patients with gastric cancer. Oncol Lett. 2014;8(4):1745-50.
117. Davidson M, Barber LJ, Woolston A, Cafferkey C, Mansukhani S, Griffiths B, et al. Detecting and Tracking Circulating Tumour DNA Copy Number Profiles during First Line Chemotherapy in Oesophagogastric Adenocarcinoma. Cancers. 2019;11(5).
118. Cohen PA, Flowers N, Tong S, Hannan N, Pertile MD, Hui L. Abnormal plasma DNA profiles in early ovarian cancer using a non-invasive prenatal testing platform: implications for cancer screening. Bmc Medicine. 2016;14.
119. Adalsteinsson VA, Ha G, Freeman SS, Choudhury AD, Stover DG, Parsons HA, et al. Scalable whole-exome sequencing of cell-free DNA reveals high concordance with metastatic tumors. Nature Communications. 2017;8.
120. Stover DG, Parsons HA, Ha G, Freeman SS, Barry WT, Guo H, et al. Association of Cell-Free DNA Tumor Fraction and Somatic Copy Number Alterations With Survival in Metastatic Triple-Negative Breast Cancer. J Clin Oncol. 2018;36(6):543-53.
121. Choudhury AD, Werner L, Francini E, Wei XX, Ha G, Freeman SS, et al. Tumor fraction in cell-free DNA as a biomarker in prostate cancer. JCI Insight. 2018;3(21).

122. Francini E, Ou FS, Rhoades J, Wolfe EG, O'Connor EP, Ha G, et al. Circulating Cell-Free DNA as Biomarker of Taxane Resistance in Metastatic Castration-Resistant Prostate Cancer. Cancers (Basel). 2021;13(16).
123. Molparia B, Nichani E, Torkamani A. Assessment of circulating copy number variant detection for cancer screening. Plos One. 2017;12(7).
124. Zill OA, Banks KC, Fairclough SR, Mortimer SA, Vowles JV, Mokhtari R, et al. The Landscape of Actionable Genomic Alterations in Cell-Free Circulating Tumor DNA from 21,807 Advanced Cancer Patients. Clinical Cancer Research. 2018;24(15):3528-38.
125. Belic J, Koch M, Ulz P, Auer M, Gerhalter T, Mohan S, et al. Rapid Identification of Plasma DNA Samples with Increased ctDNA Levels by a Modified FAST-SeqS Approach. Clinical Chemistry. 2015;61(6):838-49.
126. Tsui DWY, Cheng ML, Shady M, Yang JL, Stephens D, Won H, et al. Tumor fraction-guided cell-free DNA profiling in metastatic solid tumor patients. Genome Med. 2021;13(1):96.
127. Kirkizlar E, Zimmermann B, Constantin T, Swenerton R, Hoang B, Wayham N, et al. Detection of Clonal and Subclonal Copy-Number Variants in Cell-Free DNA from Patients with Breast Cancer Using a Massively Multiplexed PCR Methodology. Transl Oncol. 2015;8(5):407-16.
128. Jamal-Hanjani M, Wilson GA, Horswell S, Mitter R, Sakarya O, Constantin T, et al. Detection of ubiquitous and heterogeneous mutations in cell-free DNA from patients with early-stage non-small-cell lung cancer. Ann Oncol. 2016;27(5):862-7.
129. Douville C, Cohen JD, Ptak J, Popoli M, Schaefer J, Silliman N, et al. Assessing aneuploidy with repetitive element sequencing. Proc Natl Acad Sci U S A. 2020;117(9):4858-63.
130. Lee J, Franovic A, Shiotsu Y, Kim ST, Kim KM, Banks KC, et al. Detection of ERBB2 (HER2) Gene Amplification Events in Cell-Free DNA and Response to Anti-HER2 Agents in a Large Asian Cancer Patient Cohort. Front Oncol. 2019;9:212.
131. Grenda A, Wojas-Krawczyk K, Skoczylas T, Krawczyk P, Sierocińska-Sawa J, Wallner G, et al. HER2 gene assessment in liquid biopsy of gastric and esophagogastric junction cancer patients qualified for surgery. BMC Gastroenterol. 2020;20(1):382.
132. Shoda K, Ichikawa D, Fujita Y, Masuda K, Hiramoto H, Hamada J, et al. Monitoring the HER2 copy number status in circulating tumor DNA by droplet digital PCR in patients with gastric cancer. Gastric Cancer. 2017;20(1):126-35.
133. Leal A, van Grieken NCT, Palsgrove DN, Phallen J, Medina JE, Hruban C, et al. White blood cell and cell-free DNA analyses for detection of residual disease in gastric cancer. Nat Commun. 2020;11(1):525.
134. Yang J, Gong Y, Lam VK, Shi Y, Guan Y, Zhang Y, et al. Deep sequencing of circulating tumor DNA detects molecular residual disease and predicts recurrence in gastric cancer. Cell Death Dis. 2020;11(5):346.
135. Azad TD, Chaudhuri AA, Fang P, Qiao Y, Esfahani MS, Chabon JJ, et al. Circulating Tumor DNA Analysis for Detection of Minimal Residual Disease After Chemoradiotherapy for Localized Esophageal Cancer. Gastroenterology. 2020;158(3):494-505.e6.
136. Openshaw MR, Mohamed AA, Ottolini B, Fernandez-Garcia D, Richards CJ, Page K, et al. Longitudinal monitoring of circulating tumour DNA improves prognostication and relapse detection in gastroesophageal adenocarcinoma. Br J Cancer. 2020;123(8):1271-9.
137. Bianchi DW, Platt LD, Goldberg JD, Abuhamad AZ, Sehnert AJ, Rava RP. Genome-wide fetal aneuploidy detection by maternal plasma DNA sequencing. Obstet Gynecol. 2012;119(5):890-901.

138. Raman L, Dheedene A, De Smet M, Van Dorpe J, Menten B. WisecondorX: improved copy number detection for routine shallow whole-genome sequencing. Nucleic Acids Research. 2019;47(4):1605-14.
139. Straver R, Sistermans EA, Holstege H, Visser A, Oudejans CBM, Reinders MJT. WISECONDOR: detection of fetal aberrations from shallow sequencing maternal plasma based on a within-sample comparison scheme. Nucleic Acids Research. 2014;42(5).
140. Bettegowda C, Sausen M, Leary RJ, Kinde I, Wang YX, Agrawal N, et al. Detection of Circulating Tumor DNA in Early- and Late-Stage Human Malignancies. Science Translational Medicine. 2014;6(224).
141. Salk JJ, Schmitt MW, Loeb LA. Enhancing the accuracy of next-generation sequencing for detecting rare and subclonal mutations. Nat Rev Genet. 2018;19(5):269-85.
142. Gibson CJ, Steensma DP. New Insights from Studies of Clonal Hematopoiesis. Clin Cancer Res. 2018;24(19):4633-42.
143. Jacobs KB, Yeager M, Zhou WY, Wacholder S, Wang ZM, Rodriguez-Santiago B, et al. Detectable clonal mosaicism and its relationship to aging and cancer. Nature Genetics. 2012;44(6):651-U68.
144. Laurie CC, Laurie CA, Rice K, Doheny KF, Zelnick LR, McHugh CP, et al. Detectable clonal mosaicism from birth to old age and its relationship to cancer. Nature Genetics. 2012;44(6):642-U58.
145. Razavi P, Li BT, Brown DN, Jung B, Hubbell E, Shen RL, et al. High-intensity sequencing reveals the sources of plasma circulating cell-free DNA variants. Nature Medicine. 2019;25(12):1928-+.
146. Ococks E, Frankell AM, Masque Soler N, Grehan N, Northrop A, Coles H, et al. Longitudinal tracking of 97 esophageal adenocarcinomas using liquid biopsy sampling. Ann Oncol. 2021;32(4):522-32.
147. International Agency for Reasearch on Cancer, fact sheet, World Health Organization. Available at https://gco.iarc.fr/today/fact-sheets-cancers, accessed on 10-19-2021. [
148. Survival rates for Esophageal cancer, American Cancer Society. Available at https://www.cancer.org/cancer/esophagus-cancer/detection-diagnosis-staging/survival-rates.html, accessed on 11-03-2021. [
149. Survival rates stomach cancer, American Cancer Society. Available at https://www.cancer.org/cancer/stomach-cancer/detection-diagnosis-staging/survival-rates.html, accessed on 11-03-2021. [
150. Tomasello G, Petrelli F, Ghidini M, Pezzica E, Passalacqua R, Steccanella F, et al. Tumor regression grade and survival after neoadjuvant treatment in gastro-esophageal cancer: A meta-analysis of 17 published studies. Eur J Surg Oncol. 2017;43(9):1607-16.
151. Contino G, Vaughan TL, Whiteman D, Fitzgerald RC. The Evolving Genomic Landscape of Barrett's Esophagus and Esophageal Adenocarcinoma. Gastroenterology. 2017;153(3):657-73.e1.
152. Lengyel CG, Hussain S, Trapani D, El Bairi K, Altuna SC, Seeber A, et al. The Emerging Role of Liquid Biopsy in Gastric Cancer. J Clin Med. 2021;10(10).
153. Pietrantonio F, Miceli R, Raimondi A, Kim YW, Kang WK, Langley RE, et al. Individual Patient Data Meta-Analysis of the Value of Microsatellite Instability As a Biomarker in Gastric Cancer. J Clin Oncol. 2019;37(35):3392-400.
154. Zhao D, Klempner SJ, Chao J. Progress and challenges in HER2-positive gastroesophageal adenocarcinoma. J Hematol Oncol. 2019;12(1):50.
155. Chivu-Economescu M, Matei L, Necula LG, Dragu DL, Bleotu C, Diaconu CC. New therapeutic options opened by the molecular classification of gastric cancer. World Journal of Gastroenterology. 2018;24(18):1942-61.
156. Bonelli P, Borrelli A, Tuccillo FM, Silvestro L, Palaia R, Buonaguro FM. Precision medicine in gastric cancer. World J Gastrointest Oncol. 2019;11(10):804-29.

157. Liu YJ, Shen D, Yin X, Gavine P, Zhang T, Su X, et al. HER2, MET and FGFR2 oncogenic driver alterations define distinct molecular segments for targeted therapies in gastric carcinoma. British Journal of Cancer. 2014;110(5):1169-78.
158. Zhai R, Zhao Y, Su L, Cassidy L, Liu G, Christiani DC. Genome-wide DNA methylation profiling of cell-free serum DNA in esophageal adenocarcinoma and Barrett esophagus. Neoplasia. 2012;14(1):29-33.
159. Rumiato E, Boldrin E, Malacrida S, Realdon S, Fassan M, Morbin T, et al. Detection of genetic alterations in cfDNA as a possible strategy to monitor the neoplastic progression of Barrett's esophagus. Transl Res. 2017;190:16-24.e1.
160. Xu E, Gu J, Hawk ET, Wang KK, Lai M, Huang M, et al. Genome-wide methylation analysis shows similar patterns in Barrett's esophagus and esophageal adenocarcinoma. Carcinogenesis. 2013;34(12):2750-6.
161. Zviran A, Schulman RC, Shah M, Hill STK, Deochand S, Khamnei CC, et al. Genome-wide cell-free DNA mutational integration enables ultra-sensitive cancer monitoring. Nat Med. 2020;26(7):1114-24.
162. Egyud M, Tejani M, Pennathur A, Luketich J, Sridhar P, Yamada E, et al. Detection of Circulating Tumor DNA in Plasma: A Potential Biomarker for Esophageal Adenocarcinoma. Annals of Thoracic Surgery. 2019;108(2):343-9.
163. Zhang C, Chen Z, Chong X, Chen Y, Wang Z, Yu R, et al. Clinical implications of plasma ctDNA features and dynamics in gastric cancer treated with HER2-targeted therapies. Clin Transl Med. 2020;10(8):e254.
164. Kim YW, Kim YH, Song Y, Kim HS, Sim HW, Poojan S, et al. Monitoring circulating tumor DNA by analyzing personalized cancer-specific rearrangements to detect recurrence in gastric cancer. Experimental and Molecular Medicine. 2019;51.
165. Wan JCM, Massie C, Garcia-Corbacho J, Mouliere F, Brenton JD, Caldas C, et al. Liquid biopsies come of age: towards implementation of circulating tumour DNA. Nature Reviews Cancer. 2017;17(4):223-38.
166. Creemers A, Krausz S, Strijker M, van der Wel MJ, Soer EC, Reinten RJ, et al. Clinical value of ctDNA in upper-GI cancers: A systematic review and meta-analysis. Biochimica Et Biophysica Acta-Reviews on Cancer. 2017;1868(2):394-403.
167. Maron SB, Alpert L, Kwak HA, Lomnicki S, Chase L, Xu D, et al. Targeted Therapies for Targeted Populations: Anti-EGFR Treatment for EGFR-Amplified Gastroesophageal Adenocarcinoma. Cancer Discov. 2018;8(6):696-713.
168. Lee J, Kim ST, Kim K, Lee H, Kozarewa I, Mortimer PGS, et al. Tumor Genomic Profiling Guides Patients with Metastatic Gastric Cancer to Targeted Treatment: The VIKTORY Umbrella Trial. Cancer Discov. 2019;9(10):1388-405.
169. Willis J, Lefterova MI, Artyomenko A, Kasi PM, Nakamura Y, Mody K, et al. Validation of Microsatellite Instability Detection Using a Comprehensive Plasma-Based Genotyping Panel. Clin Cancer Res. 2019;25(23):7035-45.
170. Eisenberger CF, Stoecklein NH, Jazra S, Hosch SB, Peiper M, Scheunemann P, et al. The detection of oesophageal adenocarcinoma by serum microsatellite analysis. Eur J Surg Oncol. 2006;32(9):954-60.
171. Chen ZH, Zhang C, Zhang MQ, Li BF, Niu YY, Chen LM, et al. Chromosomal instability of circulating tumor DNA reflect therapeutic responses in advanced gastric cancer. Cell Death & Disease. 2019;10.
172. Chalmers ZR, Connelly CF, Fabrizio D, Gay L, Ali SM, Ennis R, et al. Analysis of 100,000 human cancer genomes reveals the landscape of tumor mutational burden. Genome Med. 2017;9(1):34.
173. Qiu P, Poehlein CH, Marton MJ, Laterza OF, Levitan D. Measuring Tumor Mutational Burden (TMB) in Plasma from mCRPC Patients Using Two Commercial NGS Assays. Sci Rep. 2019;9(1):114.

174. Si H, Kuziora M, Quinn KJ, Helman E, Ye J, Liu F, et al. A Blood-based Assay for Assessment of Tumor Mutational Burden in First-line Metastatic NSCLC Treatment: Results from the MYSTIC Study. Clin Cancer Res. 2021;27(6):1631-40.
175. Win AK, Jenkins MA, Dowty JG, Antoniou AC, Lee A, Giles GG, et al. Prevalence and Penetrance of Major Genes and Polygenes for Colorectal Cancer. Cancer Epidemiol Biomarkers Prev. 2017;26(3):404-12.
176. Mandelker D, Zhang L, Kemel Y, Stadler ZK, Joseph V, Zehir A, et al. Mutation Detection in Patients With Advanced Cancer by Universal Sequencing of Cancer-Related Genes in Tumor and Normal DNA vs Guideline-Based Germline Testing. Jama. 2017;318(9):825-35.
177. Schubert SA, Morreau H, de Miranda N, van Wezel T. The missing heritability of familial colorectal cancer. Mutagenesis. 2020;35(3):221-31.
178. Rasnic R, Linial N, Linial M. Expanding cancer predisposition genes with ultra-rare cancer-exclusive human variations. Sci Rep. 2020;10(1):13462.
179. Schmit SL, Edlund CK, Schumacher FR, Gong J, Harrison TA, Huyghe JR, et al. Novel Common Genetic Susceptibility Loci for Colorectal Cancer. J Natl Cancer Inst. 2019;111(2):146-57.
180. Law PJ, Timofeeva M, Fernandez-Rozadilla C, Broderick P, Studd J, Fernandez-Tajes J, et al. Association analyses identify 31 new risk loci for colorectal cancer susceptibility. Nat Commun. 2019;10(1):2154.
181. Dunlop MG, Tenesa A, Farrington SM, Ballereau S, Brewster DH, Koessler T, et al. Cumulative impact of common genetic variants and other risk factors on colorectal cancer risk in 42,103 individuals. Gut. 2013;62(6):871-81.
182. Jiao S, Peters U, Berndt S, Brenner H, Butterbach K, Caan BJ, et al. Estimating the heritability of colorectal cancer. Hum Mol Genet. 2014;23(14):3898-905.
183. Bonjoch L, Franch-Expósito S, Garre P, Belhadj S, Muñoz J, Arnau-Collell C, et al. Germline Mutations in FAF1 Are Associated With Hereditary Colorectal Cancer. Gastroenterology. 2020;159(1):227-40.e7.
184. EpiproColon, available at https://www.epiprocolon.com/us/medical-professionals/question-answers/. Accessed on 10-13-2021. [
185. National Cancer Institute, article "FDA approves blood test that can help guide cancer treatment". Available at https://www.cancer.gov/news-events/cancer-currents-blog/2020/fda-guardant-360-foundation-one-cancer-liquid-biopsy Accessed on 10-27-2021 [
186. Liu MC, Oxnard GR, Klein EA, Swanton C, Seiden MV. Sensitive and specific multi-cancer detection and localization using methylation signatures in cell-free DNA. Ann Oncol. 2020;31(6):745-59.
187. Fiala C, Diamandis EP. A multi-cancer detection test: focus on the positive predictive value. Ann Oncol. 2020;31(9):1267-8.
188. NHS-Galleri Trial Clinical, available at https://grail.com/clinical-studies/nhs-galleri-trial-clinical/. Accessed on 10-19-2021 [
189. Clinical Trials, US National Library of Medicin, National Institute of Health. "Cell-free DNA in Hereditary And High-Risk Malignancies (CHARM)". Available at https://clinicaltrials.gov/ct2/show/NCT04261972?term=cell+free&cond=hereditary&draw=2&rank=1 Accessed on 10-20-2021 [