Diploma Thesis

Probability Prediction of Homologous Recombination Deficiency in Serous Ovarian Cancer Through a Neural Network Based on Histopathological Slides from TCGA

Submitted by

Christina Bömcke

In Fulfilment of the Requirements for the Degree of

Doctor medicinae universae

(Dr. med. univ.)

At the

Medical University of Graz

Conducted at the

Diagnostic and Research Institute of Pathology

Under the Supervision of

Ao.Univ.-Prof. Dr.med.univ. Peter Regitnig and

Univ.-Doz. Ing. Mag.phil. Mag.rer.nat. Dr.phil. Andreas Holzinger

Graz, 05.07.2021

Eidesstattliche Erklärung

Ich erkläre ehrenwörtlich, dass ich die vorliegende Arbeit selbstständig und ohne fremde Hilfe verfasst habe, andere als die angegebenen Quellen nicht verwendet habe und die den benutzten Quellen wörtlich oder inhaltlich entnommenen Stellen als solche kenntlich gemacht habe.

Graz, 05.07.2021 *Christina Bömcke, eh*

Acknowledgments

I would like to express my sincere gratitude to my supervisor, Ao. Univ.-Prof. Dr.med.univ Peter Regitnig for his excellent guidance and continuous support over the course of this thesis. In addition, I would like to thank Marcus Bloice, MSc, for his interdisciplinary collaboration and specifically for sharing his work and in-depth knowledge on neural networks. I would also like to thank Ao.Univ.-Prof. Dr. Andreas Holzinger for the co-supervision of my thesis.

List of content

List of abbreviations

AI	Artificial intelligence
BER	Base excision repair
BMI	Body mass index
CDK	cyclin-dependent kinase
DSB	Double strand break
FN	False negative
FP	False positive
HGSOC	High grade serous ovarian cancer
HNPCC	Hereditary Non-Polyposis Colorectal Cancer
HR	Homologous Recombination
HRD	Homologous Recombination Deficiency
LGSOC	Low grade serous ovarian cancer
ML	Machine Learning
OC	Ovarian Cancer
PARP	Poly (ADP-ribose) polymerase
STIC	Serous Tubal Intraepithelial Carcinoma
TCGA	The Cancer Genome Atlas
TN	True negative
TP	True positive

List of figures

List of tables

Zusammenfassung

Einleitung: Das Ovarialkarzinom zählt zu den häufigsten Todesursachen weltweit. Aufgrund eines langen asymptomatischen Krankheitsverlaufs wird die Diagnose häufig erst in sehr fortgeschrittenen Stadien gestellt und die Prognose ist meist ungünstig. Das seröse Ovarialkarzinom welches als low-grade und high-grade vorliegen kann, ist der häufigste morphologische Typ. Eine der Ursachen für das Entstehen dieses bösartigen Tumors, ist eine Störung in einem der Reparaturmechanismen der DNA, nämlich der homologen Rekombination. Fällt dieser Reparaturmechanismus der Zelle weg, kommt es zu einer ineffizienten DNA Reparatur, die die Krebsentstehung begünstigt. Patientinnen mit *BRCA1* oder *BRCA2* Mutationen weisen genetisch bedingt eine Homologe Rekombinations-Defizienz auf und leiden daher gehäuft unter Brust- oder Eierstockkrebs. Die Homologe Rekombinations-Defizienz kann aber auch durch andere Ursachen bedingt sein, wie zum Beispiel andere genetische Mutationen, DNA-Methylierungen oder andere uns bisher unbekannte Faktoren. Therapeutisch gesehen, kann man sich diese Defizienz zu Nutze machen und mit sogenannten Poly (ADP-ribose)-Polymerase-Inhibitoren (PARP-Inhibitoren) gezielt dort eingreifen. Die PARP-Inhibitoren verhindern die DNA Reparatur über den Weg der Basen-Exzisionsreparatur (BER), und zwingen die Zelle, auf die homologe Rekombination, als zweiten Reparaturmechanismus, zurückzugreifen. Wenn dieser aber, wie beispielsweise bei BRCA-Mutierten, auch nicht funktioniert, geht die DNA der Zelle zugrunde und das Tumorwachstum wird gehemmt. Voraussetzung für die effektive Wirkung der PARP-Inhibitoren ist also eine Homologe Rekombinations-Defizienz (HRD). Histologisch könnte sich die HRD durch eine unterschiedliche Morphologie der Zellen bemerkbar machen. Es gibt bis dato aber keine Anhaltspunkte, anhand welcher Merkmale sich die Defizienz erkennen lässt. Um dieser Frage nachzugehen, bietet die künstliche Intelligenz ein leistungsfähiges Werkzeug zur automatischen Analyse und Klassifizierung von Bilddaten. In dieser Arbeit wird untersucht, ob es mittels künstlicher Intelligenz, also neuronaler Netzwerke, möglich ist, eine solche Defizienz anhand mikroskopischer Aufnahmen vorherzusagen.

Methodik: Mikroskopische gescannte Präparate von 436 Patientinnen mit Ovarialkarzinom wurden, zusammen mit den Mutationsdaten vom GDC Portal TCGA, heruntergeladen. Der HRD Status wurde definiert als Mutation in *BRCA1*, *BRCA2*, *CHEK1* oder *PTEN*. Ein neuronales Netzwerk wurde mit diesen Bildern auf das Vorhersagen über das Vorliegen von HRD trainiert.

Ergebnisse: Eine Vorhersage des HRD Status durch ein neuronales Netzwerk konnte in dieser Arbeit nicht bewiesen werden. Die Treffergenauigkeit liegt bei 45,3%. Die Klassifizierung zwischen Tumor und normalem Gewebe gelingt jedoch mit dieser Methode mit einer Treffergenauigkeit von 83,9%.

Conclusio: In dieser Arbeit wurde ansatzweise gezeigt, dass es prinzipiell möglich ist mithilfe von künstlicher Intelligenz in histologischen Präparaten morphologische Korrelate zu finden, die eine HRD identifizieren.

Abstract

Introduction: Ovarian cancer is one of the most frequent causes of death worldwide. Due to a long asymptomatic course of the disease, the diagnosis is often only made in very advanced stages and the prognosis is mostly unfavourable. Serous ovarian carcinoma, which can occur as low-grade and high-grade, is the most common morphological type. One of the causes for the development of this malignant tumour is a disturbance in one of the repair mechanisms of the DNA, namely homologous recombination. If this repair mechanism of the cell fails, inefficient DNA repair occurs, which promotes the development of cancer. Patients with BRCA1 or BRCA2 mutations are genetically deficient in homologous recombination and therefore often suffer from breast or ovarian cancer. Homologous recombination deficiency can also be caused by additional factors, such as other genetic mutations, DNA methylation or factors that are still currently unknown. This deficiency and specifically intervene with so-called PARP inhibitors therapeutically can be used. Those prevent DNA repair by inhibiting the base excision repair pathway (BER) and force the cell to resort to homologous recombination as a second repair mechanism. However, if this also does not work, as in case of BRCA mutants, the DNA of the cell perishes and hence tumour growth is inhibited. The prerequisite for the effectiveness of PARP inhibitors is therefore a homologous recombination deficiency (HRD). Histologically, this deficiency may be seen in the different morphology of the cells. Up to date, however, only limited research has been undertaken on a potential correlation of the visual characteristics of cells and its HRD status. In order to address this, artificial intelligence provides a powerful tool to automatically analyse and classify image data. This work investigates whether it is possible to automatically correlate and predict HRD status based on microscopic images by means of neuronal networks.

Methods: Microscopic whole slide images (WSI) of 436 patients with ovarian cancer were analysed, together with mutation data from the GDC portal TCGA. HRD status was defined based on a mutation in *BRCA1*, *BRCA2*, *CHEK1* or *PTEN*. A neural network was trained to predict the presence of HRD using these images.

Results: With an accuracy of only 45.3%, the neural network was not able to predict whether an HRD is present or not. However, the classification between tumour and normal tissue succeeded with a hit accuracy of 83.9%.

Conclusion: In this work, it was shown in principle that it is possible to use artificial intelligence to find morphological correlates in histological specimens. The classification of HRD in this project was not successful.

1 Introduction

1.1 Ovarian cancer

1.1.1 Epidemiology

In 2018, nearly 185 000 deaths were caused by ovarian cancer (OC) and almost 300 000 new cases occurred. With an incidence rate of 6.6 per 100 000, it is the tenth most common type of cancer worldwide at a mortality rate of 3.9 per 100 000. In Europe, the rates are even higher, with an incidence of 9.5 and a mortality rate of 5.1 per 100 000. (1)
Figure 1 and 2 display age-standardized rates of incidence and mortality from different countries over time, indicating a downward trend over the last decades.

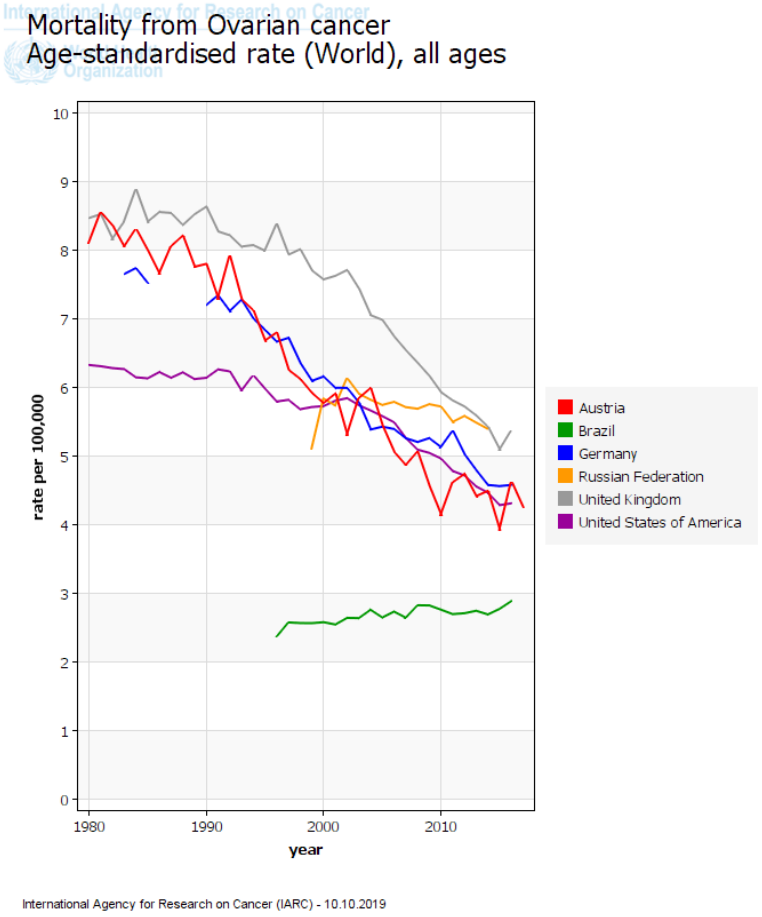

Figure 1: Mortality rates of OC from 1980 to 2010 in selected countries (2)

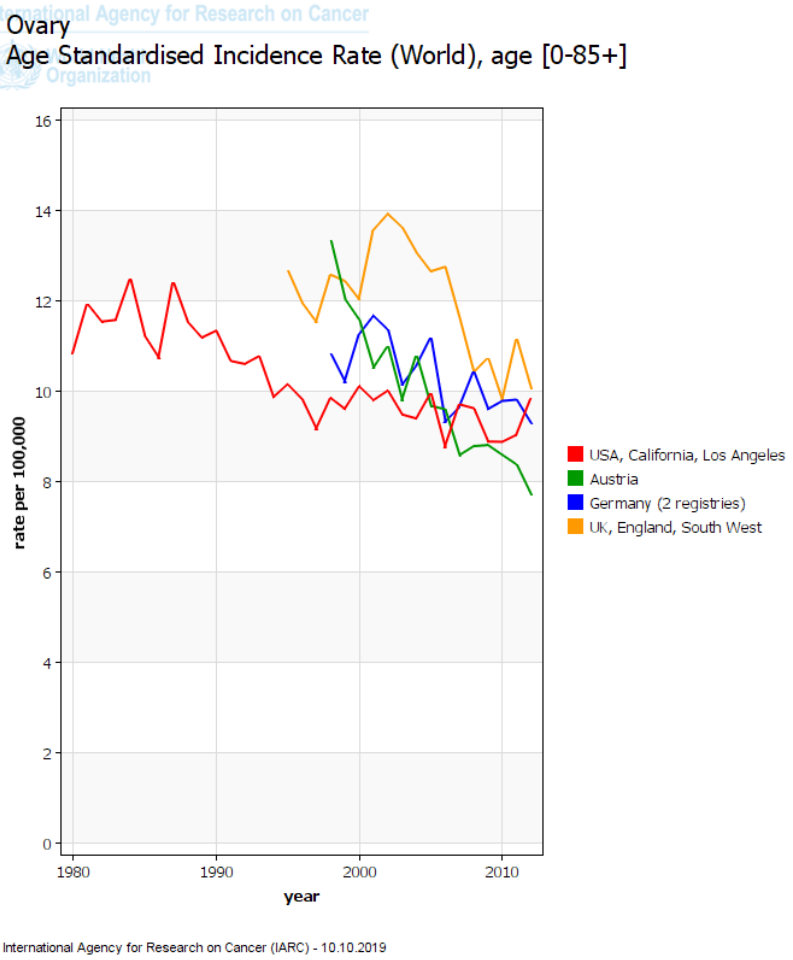

Figure 2: Incidence rates of OC between 1980 and 2010 in selected countries (2)

Concerning the occurrence of OC, there are big geographic differences. Figure 3 shows a distinct north-south gradient (2).

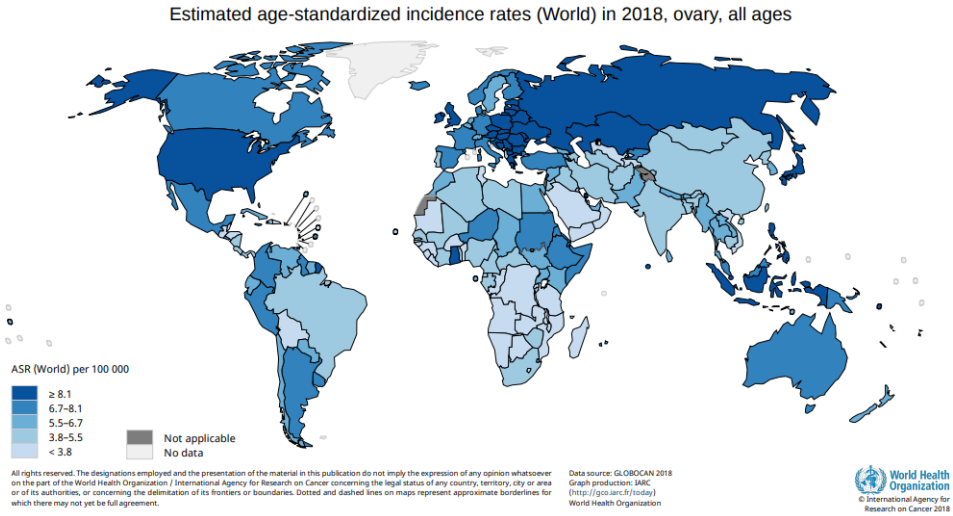

Figure 3: OC incidence in the world (2)

In Austria, ovarian cancer is the sixth most common cancer in the female population-counting about 761 new cases and 520 deaths in 2018; the cumulative risk of disease up to the age of 75 is 0.9 %, the cumulative risk of death up to the age of 75 is 0.5 % (3).
An age-standardised mortality of 10.4/100,000 women in 2018 and a 5-year survival across all stages of approximately 45 % indicates the lethality of the disease. The introduction of platinum-containing chemotherapy has had the most significant positive effect on overall survival to date (3).

The risk factors of OC depend strongly on its different subtypes. In general, the risk of developing ovarian carcinoma increases when women suffer from hereditary breast and ovarian cancer (HBOC) or from lynch syndrome (HNPCC) (4). Mutations in DNA mismatch repair genes (especially *MLH1*, *MSH2*, *MSH6* and *PMS2*) are associated with 10-15% of hereditary EOCs (3).
Most of ovarian cancers appear between the ages of 40 and 65 (5). Apart from the mentioned genetic susceptibility, one of the main risk factors of the epithelial carcinoma is the "incessant ovulation" (6). An increased number of ovulations during life leads to "a hormone induced injury involving processes of trauma and repair, with possibilities for DNA damage" (7). Less ovulation (late menarche, early menopause, hormonal contraception, breastfeeding, multiparity) reduces the risk of disease (3). However, this merely explains the development of low-grade carcinomas (8).
An elevated BMI in adulthood and hormone therapy in the peri- and post-menopause favour the development of OC (9).

1.1.2 Morphology and genetics

1.1.2.1 Histological subtypes

The World Health Organisation divides ovary carcinomas according to cell type into three main groups:

1. Epithelial tumours
2. Germ cell tumours
3. Sex cord-stroma tumours (5).

Epithelial tumours are further divided in serous, mucinous, endometrioid, clear cell, Brenner, and seromucinous tumours. The ovarian surface epithelium, named Müllerian epithelium, is the basis of these tumours. An exception is high-grade serous ovarian cancer, where the source is mostly the fallopian tube. Both affect mainly women around an average age of 50 years. The second group of ovary carcinomas describes germ cell tumours that consist of teratoma, dysgerminoma, yolk sac tumour and choriocarcinoma and concern mainly young women until an age of 25. The last group, sex cord-stroma tumours, contains fibroma, granulosa cell tumours, thecoma, and Sertoli-Leydig cell tumours. Contrary to the first two groups, this form of OC derives from the stromal component of the ovary and can occur at any age. (5)

All these tumours differ in their morphological features. The most common malignant one is the serous carcinoma (5). As it is the main focus of this thesis, the description of the morphology is largely confined to this tumour subgroup.

The serous carcinoma can emerge as low (LGSOC) or high grade carcinoma (HGSOC). Due to their clinical behaviour, low-grade carcinomas are classified as type-1 and grow slowly, while high-grade carcinomas belong to type-2, which show rapid aggressive growth (5, 8, 10).

LGSOC originates from benign cystadenoma through an adenoma-carcinoma sequence. The above-mentioned “incessant ovulation” is one reason for developing LGSOC because repetitive micro lesions in the ovarian surface epithelium (OSE) lead to neoplastic transformations.
Macroscopically, these tumours are big and consist of both solid and cystic parts.
The histology shows either a papillary or solid tumour growth and comparably few nuclear atypia and mitoses.

HGSOC occur much more frequently (90-96% of all serous OC) and are associated with a high mortality. This is due to the fact that it often proceeds asymptomatically and is therefore only detected in advanced stages (11). In addition, due to anatomical conditions, the tumour spreads rapidly in the abdominal cavity, as the physiological peritoneal fluid

distributes malignant cells to other organs (12). Once peritoneal metastases are present, the prognosis remains very poor (13).

The macroscopic distinction between LGSOC and HGSOC is not always clear. Nevertheless, HGSOC are mostly chambered, multicytic and very irregular and inhomogeneous (14).

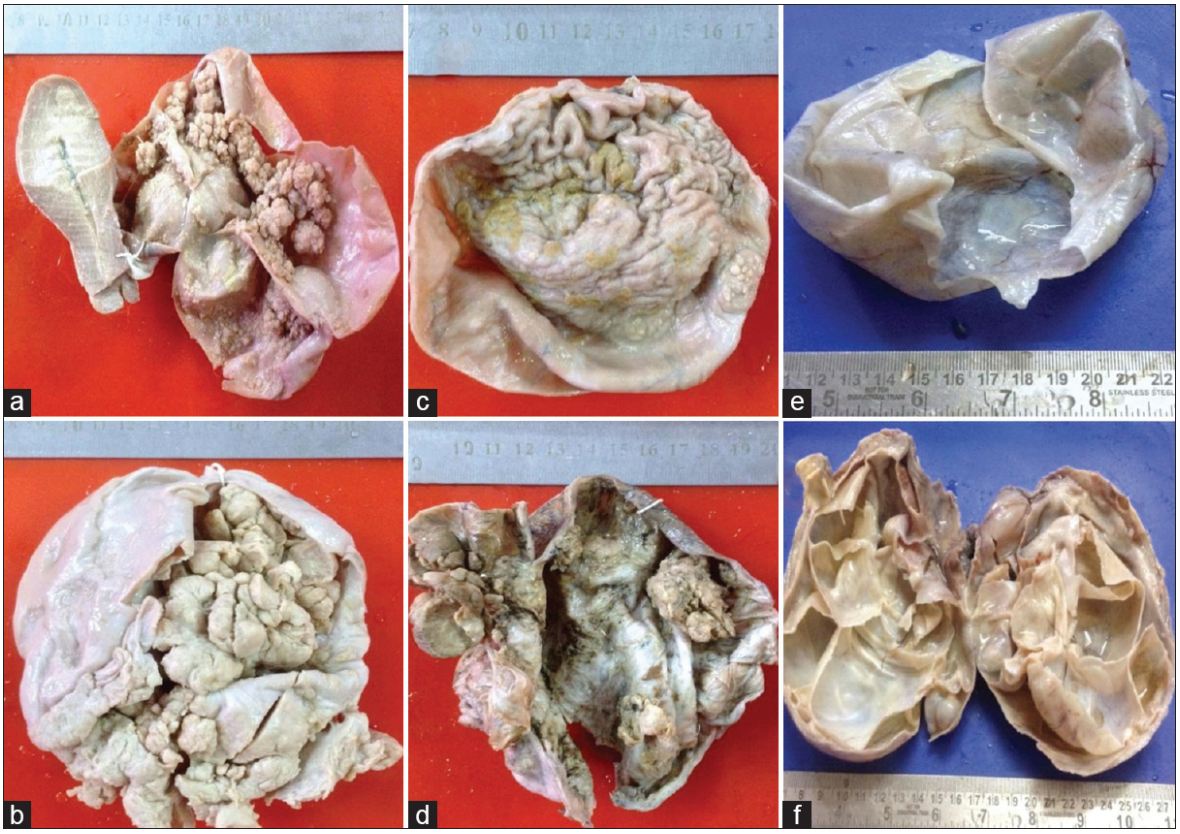

Figure 4: Examples of some epithelial ovarian tumours (14).

In microscopy, HGSOC features many atypical tumour cells next to slit-like interspaces with a larger quantity of mitoses and nuclear atypia compared to LGSOC. Nuclei are large and hyperchromatic, tend to have bizarre forms and the nucleoli are often prominent and eosinophilic. One can also find conspicuous psammoma bodies in combination with necrosis (24). The architecture is similar to normal tube tissue, presenting papillary, glandular and cribriform areas (22). If the histology shows the typical pattern of papillary architecture and psammoma bodies, it can almost certainly be assumed to be HGSOC. The solid pattern, on the other hand, makes it more difficult to distinguish HGSOC from other EC. However, squamous differentiation, adenofibromatous background, and endometriosis are more typical of ECs (15).

Various immunohistochemical markers confirm the suspicion of HGSOC: first and foremost the Wilms tumour 1 (WT1) protein. Also the markers p16 and p53 are more likely to be positive in HGSOC than in EC (15). WT1 is a tumour suppressor gene located at chromosome 11p13; a high expression had been found in ovarian serous carcinomas (16). The p16 tumour suppressor protein inhibits the cyclin-dependent kinase (CDK) 4 and 6 and thus stops the cells from entering the S phase during cell cycle (17). p53, known as "guardian of the genome" is activated when damage or changes to the DNA are detected and interrupts the cell cycle so that a malignant cell cannot proliferate (18).

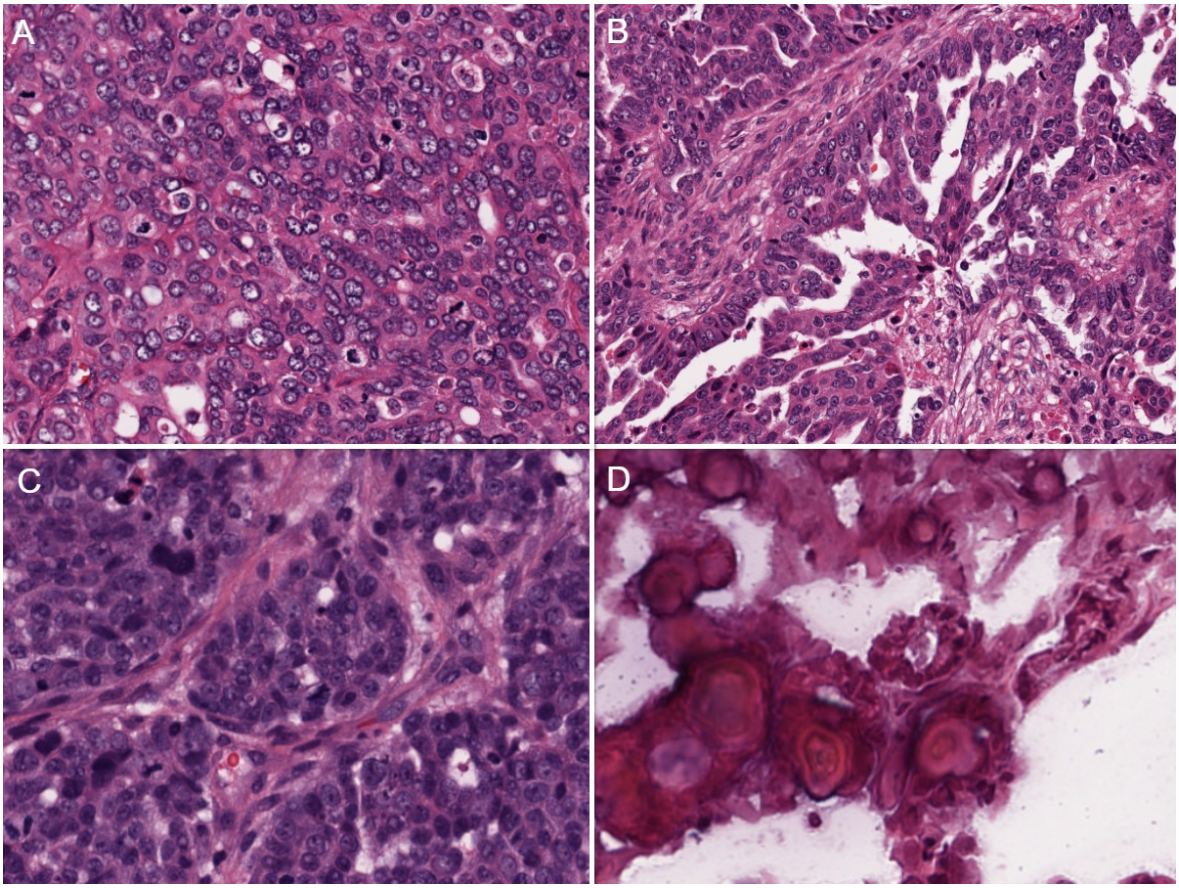

Figure 5: Histology of serous ovarian cancer (A) High-grade serous carcinoma with solid tumour architecture (TCGA-23-1122). **(B)** High-grade serous carcinoma with papillary architecture next to slit-like interspaces (TCGA-23-1119). **(C)** Cytologic atypia, mitoses and tumour stroma (TCGA-42-2582). **(D)** Psammombodys (TCGA-13-1485).

There are various theories to the origin of HGSOC.

At first it was assumed that HGSOC originate from the ovary itself; more precisely from the ovarian surface epithelium (OSE). During ovulation, cells from the OSE would spread into the stroma and form cysts, called cortical inclusion cysts, from which the carcinomas then develop (11, 19). But this theory has its limitations in the following respects. Ovarian cancer tissue resembles in its histology tumours from organs that develop from Müllerian

duct, such as the fallopian tube or the endometrium. Thus the serous subtype of OC looks like the tumours in the fallopian tube what makes an origin there likely (20).
Furthermore, examinations to find precursor lesions of HGSOC in the OSE had no success (8). In contrast, many studies confirm the existence of precursor lesions in the fallopian tube, especially in the fimbrial end (21-23). Typically these were identified as serous tubal intraepithelial carcinoma (STIC). STIC is a non-invasive lesion, confined to the epithelium of the fallopian tube. They present a malignant cell population, a disorderly pattern of growth and a strong p53 expression, so-called "p53 signature" (24, 25). Various characteristics indicate that STICs are related to HGSOC. The *TP53* mutation is identical to those in HGSCOs and the same oncogene products are detectable (25, 26).
STIC lesions have been observed in 50 to 60% of all pelvic high grade serous carcinoma cases (24, 27). But STIC lesions are not found in all cases of HGSOC: 45% of all serous ovarian cancers were not associated with STIC. This suggests that a considerable number of HGSOC might develop in other ways (24). But nowadays, it is upcoming evidence that most of HGSOC arise from the epithelium of the fimbrial end of the fallopian tube, so women who have completed family planning should be informed about the possibility of selective salpingectomy as part of benign gynaecological procedures. (3, 11, 26).

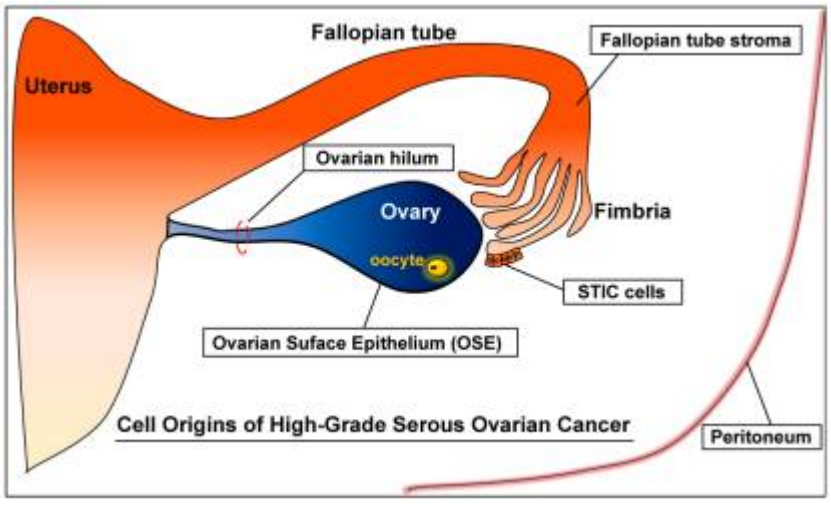

Figure 6: Cells of origin – HGSOC (11).

.

The question still remains, how the tumour, emerging in the tube, reaches the ovary. There may be two possible answers: Either STICs in the fimbriated end of the fallopian tube

spread to the ovary or, during ovulation, cells from the fallopian tube implant in the ovary epithelium and build cysts which subsequently lead to ovarian cancer (28, 29).

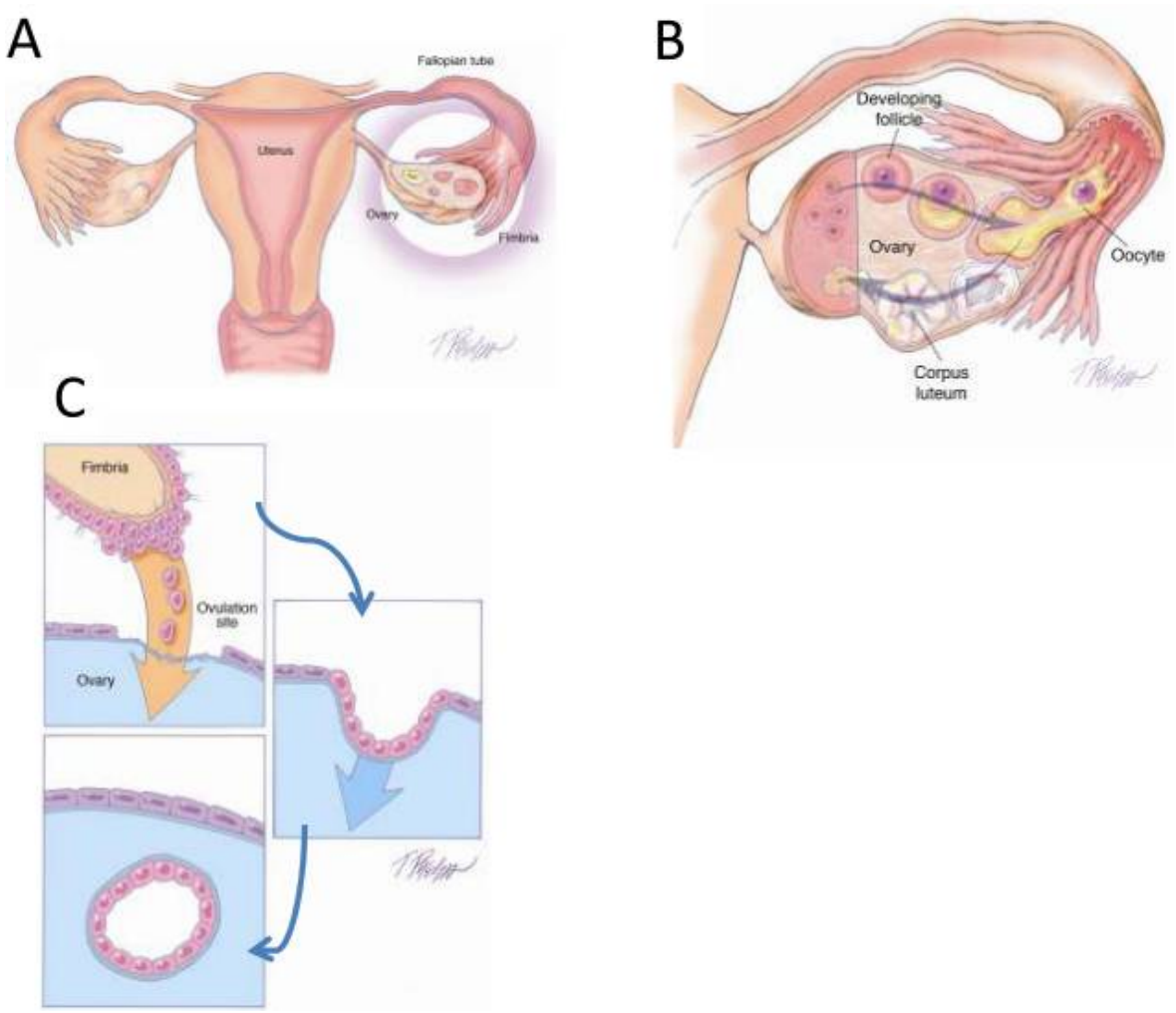

Figure 7: Possible transfer of tubal epithelium to the ovary (29).

1.1.2.2 Gene mutations

LGSOC is frequently associated with genetic mutations in the following genes: *KRAS, BRAF, PTEN*. The prognosis is favourable (4).

HGSCO tend to be genetically unstable. Molecular differences can be found in almost all HGSOC cases in the *TP53* gene (30). This gene acts as a tumour suppressor, from which the protein p53 is produced by transcription and translation. Since it possesses the ability to induce apoptosis by regulating pro-apoptotic genes through various mechanisms, the development and growth of tumour cells can be prevented (31). Somatic mutations in the

TP53 gene lead to a dysfunction of the p53 protein by building p53 isoforms with gain-of-function properties like enhanced cell survival, proliferation, adhesion, and invasion. These properties favour the development of tumours. For this reason, dedicated research aims at developing therapies targeted on the *TP53* mutation (8).

TP53 dysfunction is also used in immunohistochemistry as either complete staining loss or overexpression. McAlpine et al. revealed, that an overexpression of p53 in HGSOC is significantly associated with mutations in the BRCA genes (32). It can therefore be assumed that a BRCA mutation causes molecular changes in the cell that lead to these TP53 abnormalities (32). However, it seems clear that the simultaneous presence of TP53 overexpression favours the prognosis (32).

BRCA germ line mutations have been identified as other common mutations which are responsible for most of the genetic epithelial ovarian cancer diseases and are especially frequent in HGSOC (33). Both *BRCA1* and *BRCA2* act as tumour suppressor genes, because the proteins they are encoding play a central role in DNA repair. For example, they repair double-strand breaks in DNA by homologous recombination. A broken strand invades an undamaged template and can be repaired in this way. If these genes are mutated or inactivated, this repair function may be lost or degraded. As a consequence, humans become more susceptible to cancers, particularly to breast and ovarian cancer. The lifetime risk of developing the latter for carriers of germline *BRCA1* mutation varies from 40%-60% and for carriers of germline *BRCA2* mutation from 11%-30% (33). OC associated with germline mutations in the *BRCA* genes show a characteristic pattern: Firstly, the diagnostic age often correlates with a very young patient age; secondly, the overall survival rates are higher compared to non-BRCA cases, despite of frequent metastases in organs; thirdly, patients tend to respond well to platinum and certain non-platinum chemotherapy agents as well as to PARP- inhibitors due to their associated homologous recombination deficiency. (33)
TP53 and *BRCA* mutations are by far the most common mutations in HGSOC diseases, but other mutations such as *FAT3, CSND3, NF1, CDK12, RB1* and *GABRA6A* also occur repeatedly (34).

1.1.2.3 DNA repair mechanisms

During the cell cycle, replication errors occur constantly, and further induce breaks in the DNA. For the genetic stability and a normal cell function it is essential that mechanisms in the cell recognise and correct these errors in the DNA (34). Usually these single or double-strand breaks are continuously repaired by means of various biochemical mechanisms and almost any damage is thus eliminated (35).

Base excision repair

In the case of repair of single-strand breaks, the auto-ADP-ribosylation of poly-(ADP-Ribose) Polymerases (PARPs) is important for the dissociation of the DNA to enable the repair enzymes to access the defect. Especially PARP-1, PARP-2 and PARP-3 play a key role in this mechanism. (35)

Homologous recombination pathway

For the repair of double-strand breaks two options are available: the homologous recombination (HR) pathway and the non-homologous end joining pathway (NHEJ). The HR pathway takes a more important position in this respect as it works more efficiently and carefully and uses homologous gene sections as a basis. The NHEJ pathway, on the other hand, is very prone to errors, as it performs a random recombination. If these pathways do not function properly, mutations accumulate and trigger to genetic instabilities and elevated tumour formation as it is seen in BRCA mutated patients. (34)
The FANC protein, encoded on the *BRIP1* gene, is also able to repair double-strand breaks in DNA by homologous recombination. If this gene is mutated, the HR pathway can be defective. This is why Fanconi anaemia is mentioned in Figure 8; if both copies of the gene are mutated, the so-called Fanconi anaemia occurs, in which too few erythrocytes and leucocytes are synthesised (36).

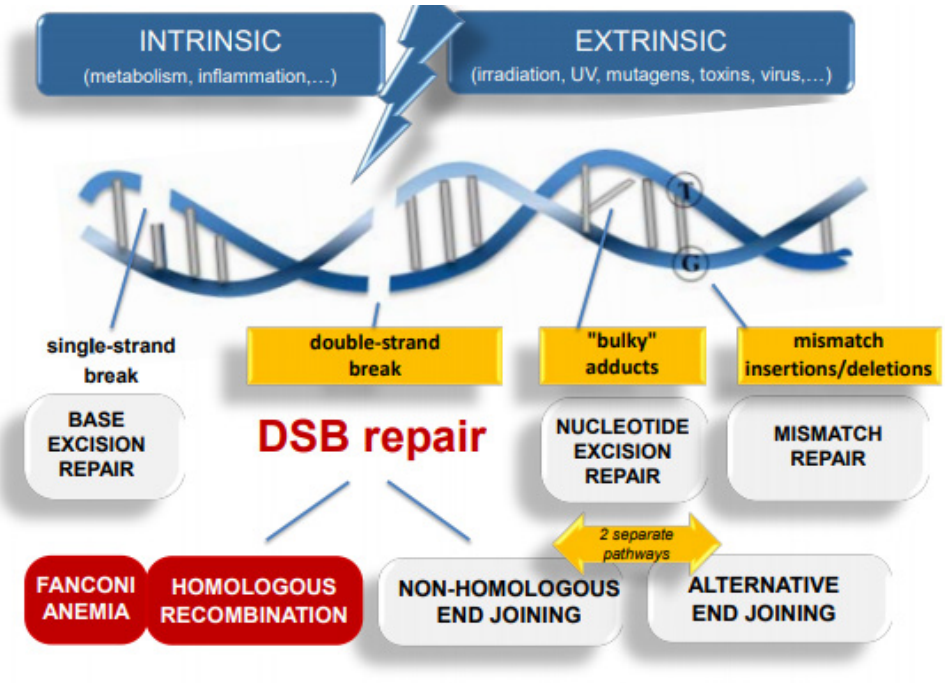

Figure 8: Possibilities of DNA repair (37)

The HR pathway showed in Figure 9 works as follows:

If the DNA presents double strand breaks, the homologous chromosome will pair during meiosis and on the defective chromosome, the area around the defect will be trimmed to create 3' overhangs, induced by the protein Rad51. It then also initiates the strand invasion by establishing a connection between the defective single strand and the homologous strand of the sister chromatid. This leads to base pairing of both strands and the formation of a D-loop. The strand exchange and the creation of Holliday junctions begin. Branch migration then means the migration of the junctions along the strand. Dissolving these Holliday junctions is the final step. Enzymes detect these sites and introduce symmetrical single strands. Then only one DNA ligase needs to catalyze the linkage of the DNA backbone and the double strand is restored. (38, 39)

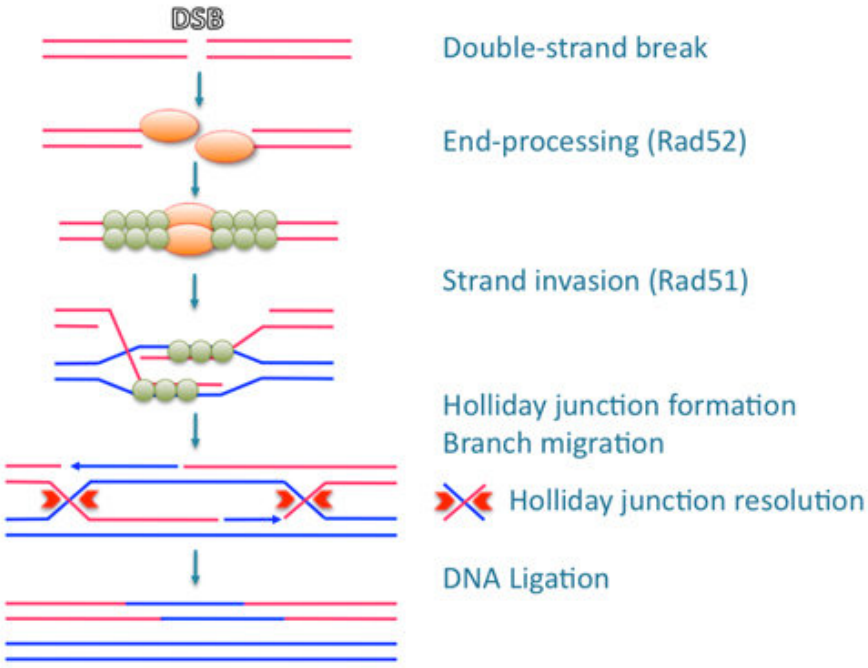

Figure 9: Homologous Recombination Pathway(38)

Some gene mutations are known to be responsible for a deficiency in the homologous recombination pathway. One of the more common genes affected is the *BRCA1*, *BRCA2* and *ARID1A* gene. Other genes are *ATM, ATRX, CHEK2, BAP1, BARD1, BLM, BRIP1, FANCA/C, MRE11A, NBN, PALB2, RAD50, RAD51, RAD51B, WRN, CHEK1, FANCL, RAD1C, RAD54L, RAD51D, CDK12* (40, 41).

It is known that if there are mutations in these genes, the cells often present a homologous recombination deficiency (HRD). But also epigenetic traits like methylation can contribute to missing homologous recombination (HR) functions, see figure 10 (40).

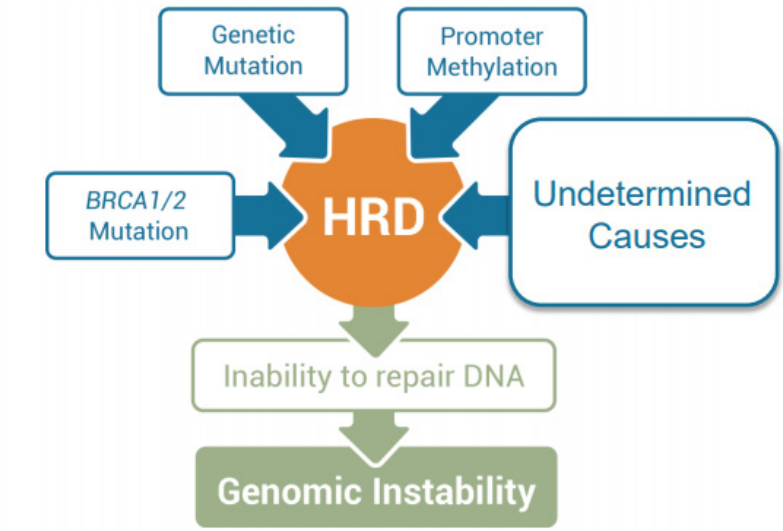

Figure 10: Causes of homologous recombination deficiency (42).

To what extent these so-called HRD genes have an effect on tumorigenesis, progression and sensitivity to PARP-Inhibitors, was investigated by Takaya et al. They concluded that only the following four genes actually have a significant effect on the HR pathway, namely *BRCA1, BRCA2, PTEN, CHEK1* (43).

Myriad, a molecular diagnostic company, developed a score to better identify the genomic effect of HRD. For this, they defined three biomarkers: loss of heterozygosity (LOH), telomeric allelic imbalance (TAI), and large-scale state transitions (LST). LOH means the absence of an allele. Normally, the chromosomes with the different genes consist of two copies. If the remaining allele is destroyed or does not function properly, the gene cannot work properly. This may often be problem-free and does not lead to consequences. But if this affects a tumour suppressor gene, the cell begins to divide abnormally and becomes cancerous. TAI refers to the number of subchromosomal regions with allelic mismatch extending to the telomere, and LST was defined as the number of breaks between regions longer than 10 Mb (43).

The sum of these three values yields a score which indicates whether or not HRD is prevalent (42). The value 42 was set as the cut off (44).

1.1.2.4 PARP Inhibitors

In recent years the HRD has been taken into advantage of a therapy. A drug for *BRCA* mutated ovarian cancer has been developed which aims to inhibit the above-mentioned PARPs in the BER repair, so that the repair of single strand breaks is no longer possible and double strand breaks are caused instead. These would normally be repaired by HR but since this pathway is switched off in the mutant tumour cell, the tumour cell undergoes apoptotis (45).

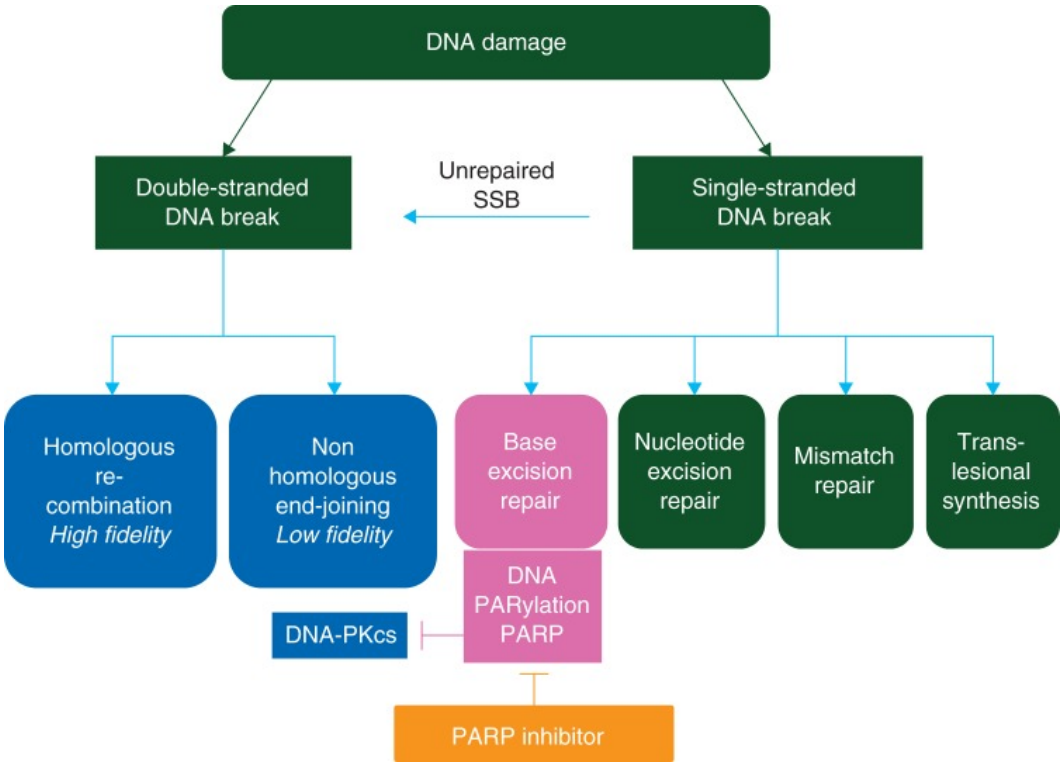

Figure 11: DSB and PARP Inhibitors (46)

So far, studies have shown that these PARP inhibitors are 1000 -times more effective in patients with *BRCA1* or *BRCA2* mutations than in *BRCA* wild type patients. Thus, PARP inhibitors were initially used exclusively for *BRCA*-mutated patients (47). One study, however, re-examined the efficacy of PARP inhibitors also in non-*BRCA* mutated patients and showed that an effect could also be achieved in 24% of these patients (48). The reason for this is that in such cases there is a *BRCA*-mutation-like behaviour. This means that even without a *BRCA1/2* mutation the HR pathway is disturbed. This may be due to other gene mutations that cause HR deficiency (49). Furthermore, it has been found that PARP inhibitors show good efficacy in patients who have little or no response to platinum-based chemotherapy. In these patients, the cancer cells are insensitive to platinum. These cells have been studied and found to have a very high PARP enzyme activity. This is apparently essential for the survival of these cells, as the treatment with PARP inhibitors has led to considerably increased apoptosis (50).

Methylation of HR related genes also plays a role. For example, if the *BRCA* promoter is methylated, this leads to inactivation of the gene and therefore to loss of function (46).

Thus in 2017, the U. S. Food and Drug Administration approved the use of PARP inhibitors as a second line therapy following platinum-based chemotherapy independent of *BRCA* mutations (47).

The Nova study investigated the efficacy of niraparib (a PARP inhibitor) in BRCA patients and in HRD positive patients. In both cases, niraparib was shown to prolong progression-free survival. While niraparib resulted in a progression-free interval of almost 13 months, the placebo group had a progression-free interval of just under 4 months. In BRCA patients, the difference was even more pronounced: 21 months in the niraparib group to 5.5 months in the placebo group. (51)

1.1.3 Diagnostic

To date, there is no established method of early detection of OC. The tumour marker CA-125 can be also elevated in many benign diseases (endometriosis, early pregnancy, liver disease, colitis, etc.). CA-125 determinations alone are therefore unsuitable in the early detection of ovarian carcinoma. Early stages are usually an incidental finding. Advanced stages are recognised by symptoms such as increase in abdominal girth, ascites, pain, constipation, nausea, weight loss. (3)

Examinations:

- Medical history including family history of tumour disease, BMI, all comorbidities.
- Clinical gynaecological examination including vaginal sonography.
- Complete laboratory including tumour markers (CA-125, CEA, possibly CA 19.9).
- CT abdomen/thorax, MRI for special questions (e.g. suspicion of liver metastases).
- In the case of elevated CEA and a family history of exposure, colonoscopy and gastroscopy; mammography if the last one is older than 12 months.
- If stage FIGO IV is suspected, histological confirmation should be sought (3).

Sonography:

Probably the most important imaging procedure in the clarification of adnexal processes is gynaecological ultrasound. However, it is a challenge to differentiate between benign and malignant, and only a few ultrasound experts have the necessary skills to recognise malignancy on ultrasound. (3)

The IOTA group (International Ovarian Tumour Analysis) has set itself the goal of introducing a standardised terminology of morphology and establishing simple characteristics for malignant and benign adnexal tumours, thus creating the "Simple Rules" with their M and B criteria seen in figure 12. They can be applied to 77 % of adnexal tumours (sensitivity and specificity of 92 % and 96 %, respectively). The remaining 23 % of adnexal tumours are inconclusive, because either both B and M or no criteria apply to the description of the tumour. (52)

Classification	Malignant Characteristics
M1	Irregular solid tumor
M2	Presence of ascites
M3	At least 4 papillary structures
M4	Irregular multilocular solid tumor with largest diameter ≥100 mm
M5	High Doppler blood flow (color score 4)
	Benign Characteristics
B1	Unilocular
B2	Presence of solid components with largest component < 7 mm
B3	Presence of acoustic shadows
B4	Smooth multilocular tumor with largest diameter < 100 mm
B5	No Doppler blood flow (color score 1)

Figure 12: IOTA criteria adapted by Moon et al. (53)

IOTA published also various models that could be used to calculate the risk of malignancy. The ADNEX model was established as the most informative. By means of nine parameters (e.g. age of the patient, size of the tumour, number of papillary projections, CA-125), it can indicate the risk for adnexal tumours. (52)

1.1.4 Therapy options

The risk of disease for women with verified *BRCA1* or *BRCA2* germ-line mutation can be reduced by 80-90% by carrying out the bilateral salpingo-oophorectomy (4). Prophylactic removal of both adnexa is therefore recommended for these women after completion of family planning (3).

The most important prognostic factor is the stage. Due to a long asymptomatic phase the OC are mostly detected in advanced stages. The therapy depends strongly on the stage (30).

Stage I: Limited to Ovaries (10%)	
IA	One ovary, no ascites, intact capsule, no tumor on external surface
IB	Both ovaries, no ascites, intact capsule, no tumor on external surface
IC	One or both ovaries with capsular involvement, ruptured capsule, ascites, or positive peritoneal washings
Stage II: Pelvic Extension (5%)	
IIA	To uterus or fallopian tubes
IIB	To other pelvic organs (e.g., bladder, rectum, or vagina)
IIC	Pelvic extension with factors as in IC
Stage III: Upper Abdominal Involvement and/or Positive Lymph Nodes (70%)	
IIIA	Microscopic seeding outside of abdominal peritoneum with negative lymph nodes
IIIB	Gross deposits less than 2 cm with negative lymph nodes
IIIC	Gross deposits greater than 2 cm and/or positive lymph nodes
Stage IV: Distant Metastases (pleural effusion, liver parenchyma, etc.) (15%)	

Figure 13: Stages of ovarian cancer (54).

OC in early stages (I to IIA at FIGO Classification) need to undergo inoperative staging, in order to detect possibly further tumour manifestations in the abdominal region. This could entail a change in therapy. The intraoperative staging is at the same time the operative therapy; it includes the inspection and palpation of the whole abdominal cavity, a peritoneal cytology, biopsies from all conspicuous places, peritoneal biopsies from inconspicuous places, bilateral adnexal extirpation, hysterectomy, infracolic omentectomy, and pelvine/paraaortal lymphadenectomy. For patients staged IA Grade 1 after operative

staging there is no indication for adjuvant chemotherapy. Patients staged IA or IB Grade 3 or IC receive platinum-based chemotherapy over six cycles. The evidence-based recommendation is a monotherapy with carboplatin for six cycles (4).

In case of OC in an advanced stage (FIGO IIB to IV), the therapy is the macroscopically complete resection, followed by a cytotoxic chemotherapy. The better the primary surgery, the better the survival. There is no indication for a second-look operation. The firstline chemotherapy is a combination therapy, which consists of carboplatin and paclitaxel for six cycles every three weeks (4).

The German guidelines of 2019 list for the first time PARP inhibitors as maintenance therapy for BRCA1- or BRCA2-mutant ovarian carcinomas. Patients in stage III/IV with a proven mutation in the BRCA gene and after response to platinum-containing therapy should receive a PARP inhibitor. The same applies to HGSOC patients after relapses. A precondition is always that the tumour is sensitive to platinum-containing therapy (4).

1.2 The Cancer Genome Atlas (TCGA)

The Cancer Genome Atlas is a cancer genomic program, containing 20,000 samples of primary cancers with their matching normal counterparts. Thus, 33 types of cancer could be characterized with regard to genetic modifications. By having collected and analyzed data of over 11,000 individuals, TCGA can provide information about genomic sequence, expression, methylation, and copy number variation for research purposes. (55, 56)
To use this program, it is only necessary to enter the cancer type as keyword and the list of all registered cases appears, consisting among others of the genetic mutation, clinical data and scanned histological slides of the specific samples. All data in this thesis were obtained from this database.

1.2.1 From DNA Sequence to Mutation data

The TCGA pipelines extract the mutation data by comparing whole exome sequencing and whole genome sequencing data from tumour samples with a reference genome. The

reference genome is the GRCh38. Thus, the pipeline can identify somatic mutations and variations. This process happens in six steps (57):

- Genome Alignment
- Alignment Co-Cleaning
- Somatic Variant Calling
- Variant Annotation
- Mutation Aggregation
- Aggregated Mutation Masking

Genome Alignment:
For sequencing, the DNA is broken down into pieces and the correct positions in the genome are not longer known. During the alignment, a program arranges the genome pieces where they fit with the highest probability and sequenced segments are brought to the correct position in the genome. BAM Files are created and represent the aligned sequence (57).

Alignment Co-Cleaning:
In this step, a software tries to improve the quality of the alignment by filtering errors caused by insertion or deletion and by using the Base Quality Score Recalibration to detect systematic errors (57).

Somatic Variant Calling:
Different pipelines like MuSE, MuTect2, VarScan2 or SomaticSniper now compare the alignment with the reference genome GRCh38. For this purpose, tumour-normal pairs are formed and somatic variants can be detected. Each pipeline has its own filter mechanisms and settings and creates at the end a VCF formatted file with raw genotypic information related to genomic positions. Due to the different procedures of the pipelines, also the content is slightly different (57).

Variant Annotation:
This term stands for the assignment of functional information to the DNA variants. By comparing with databases, containing multiple data of mutations, it is possible to identify

which of the changes in the genome are known mutations and therefore relevant and which are not. In this way, variants of interest can be identified among all mutations. This will create Annotated VCF Files from the raw VCF files (58).

Mutation Aggregation:
The annotated VCF Files are combined and a MAF file is produced for each variant caller pipeline. This MAF File contains the aggregated mutations from the different VCF Files (57).

Aggregated Mutation Masking:
Since the MAF files contain potentially germ line variants/mutations that could identify individuals, the program has to adapt those files by removing parts that contain that private information. Thus, masked MAF files are created.
While these criteria cause the pipeline to over-filter some of the true positive somatic variants in open-access MAF files, they prevent personally identifiable germline mutation information from becoming publicly available (57).

1.3 Machine Learning

Research in this field began already about 70 years ago. In 1955, John McCarthy, Professor of Computer Science at Stanford University, coined the term “artificial intelligence” (AI). He was interested in creating systems that mimic human intelligence and was one of the first to deal intensively with formalizing common sense. Among other things, this involved the automation of intelligent behaviour and the independent problem solving by machines. In his paper “Artificial Intelligence, Logic and Formalizing Common Sense”, published in 1989 he put forward the thesis that “common-sense problems can be solved by logical reasoning”. Today he is known as the “father of AI”. (38, 56)

Nowadays, there are two theses about AI (59):

1) Strong AI: Computers are intelligent and can make independent decisions.
2) Weak AI: Computers only can simulate intelligent human behaviour.

While the strong thesis is highly controversial, numerous applications in everyday life already prove the weak thesis. Some examples are Spam filtering, Face recognition, Machine translation, Speech recognition, Data mining and Robot motion. (60)

Machine Learning is a subset of the Artificial Intelligence technique (2). The term stands for generating artificial knowledge from experience. IT-systems are able to learn data sets and find patterns and regularities by using algorithms. Gained knowledge can be generalized and is useful for the analysis of previously unknown data (61).

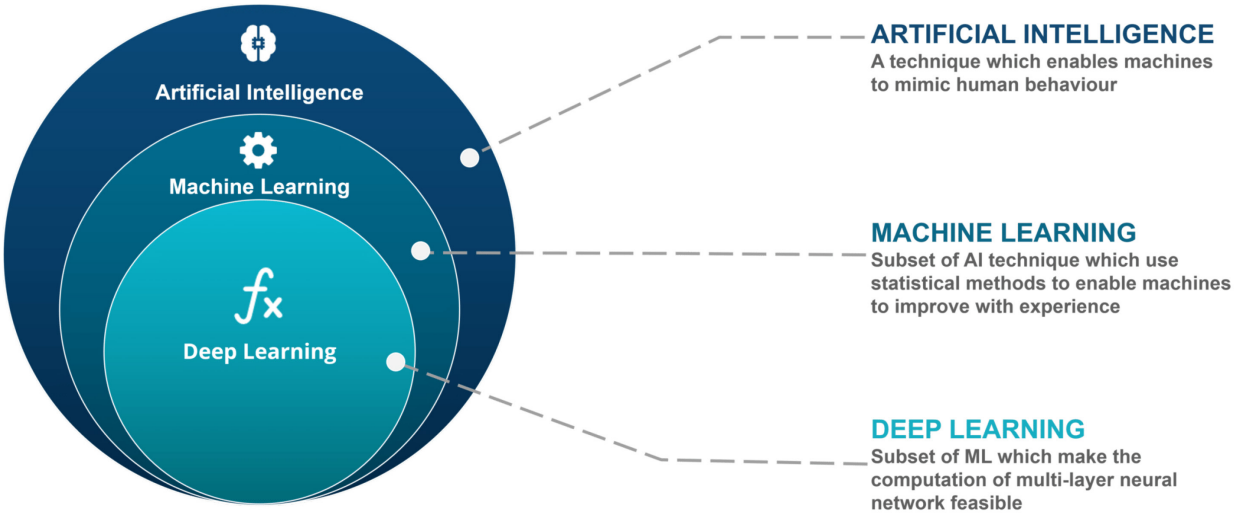

Figure 14: Overview of AI components (2).

Those learning models consist of neuronal networks. In neuronal networks, several neurons work together, similar to human brains, get input signals and generate output signals. The signals pass through some layers and lead in the end to a result. In order to use those networks for unknown data, it is necessary to train the software at first. The machine learning model needs for this input data and output data. Therefore, a big training-set with many samples is necessary. After each training-set, a validation-set can indicate the error rate. At the end, a test set with unknown results has to be solved by the neuronal network. Depending on the error rate it is decided, whether it can be used or not to a greater extend. (62)

This ability of neuronal networks is also used in medicine. Especially in image analysis, deep learning models in form of convolutional neuronal networks (CNN) become more and more popular. In contrast to machine learning, deep leaning models use several layers and it is not obvious to the user which methods the network uses to arrive at the result.
Convolutional means that filters are used that identify different patterns, such as edges or circles. In deeper layers, depending on the CNN, eyes, ears or other more sophisticated objects can then also be detected. After the image has passed through all the layers with the various filters, objects depicted in the image can be predicted with a certain degree of accuracy (63). Studies exist that demonstrate that CNNs outperform humans in certain areas (64, 65).

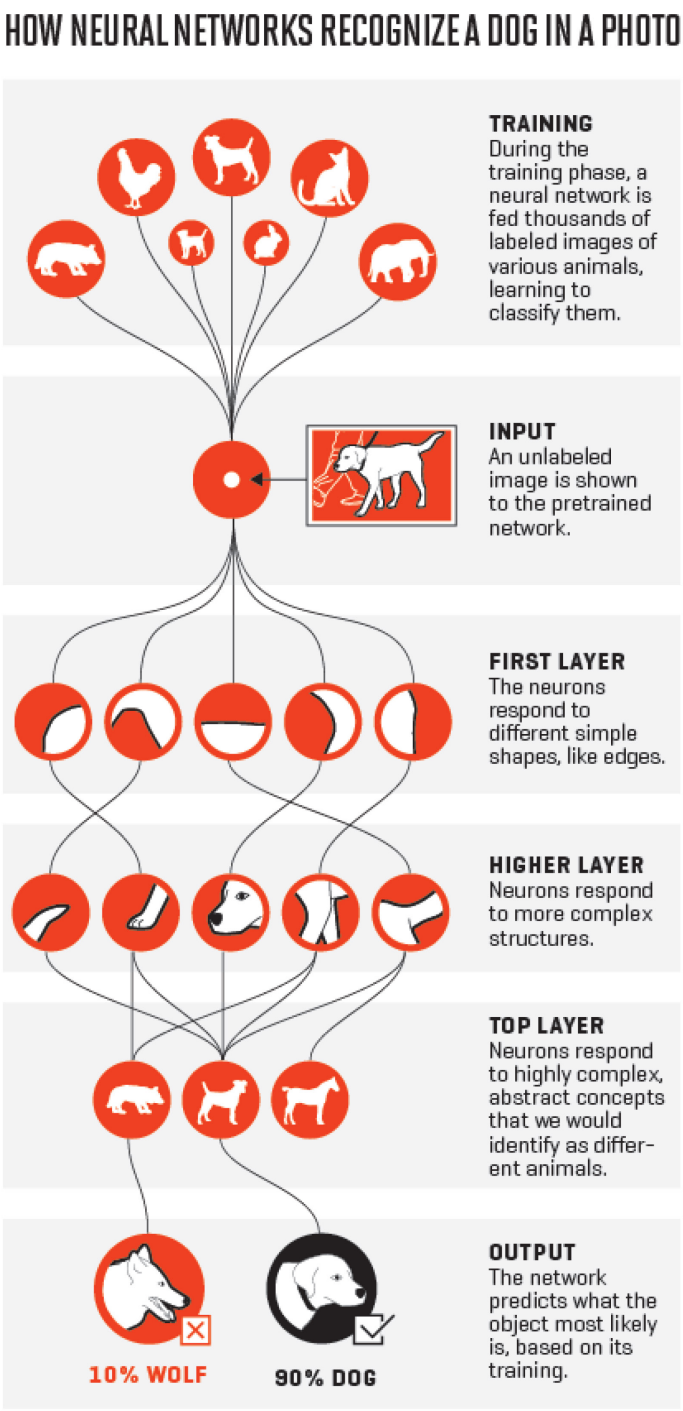

Figure 15: A simplified example how a CNN gets to its result (64).

In 2016, Liang, Z., et al. already used automatic image recognition technologies like a CNN to differ malaria-infected cells from uninfected cells. With an accuracy of 97.37%, the machine learning model could classify the individual cells (57).

Another example is the use of a deep learning model for the histopathology detection of non-small cell lung cancer. Coudray, N., et al. published in 2018, that their CNN worked with a performance with an average area under the curve of 0.97 (65).
In the classification of different grades of cervical lesions in cervical cell images (normal, mild, moderate, severe and carcinoma) using a CNN, an average accuracy of 95% could be achieved.

1.4 Hypothesis

It was previously published that genetic instability affects the morphology of tumours. The type and number of mutations changes the morphology, so that HRD positive tumours are six times more likely to exhibit non-classical histological patterns such as solid, pseudo-endometrioid architecture or transitional growth patterns, called “SET”. In addition, in BRCA1 mutations, for example, prominent tumour-infiltrating lymphocytes, nuclear atypia with bizarre nuclei, and a high mitotic index are particularly apparent (66).
This was published in 2016 by Ritterhouse et al., also underlining that “non-classic histology is strongly associated with mutations in homologous recombination genes”(66).
Based on this thesis, the following hypothesis was formulated:
A trained neural network can predict the homologous recombination deficiency in serous ovarian cancer cases by means of microscopic morphological changes on histopathological whole slide images.
As a second hypothesis, we assumed that a homologous recombination disorder in ovarian carcinomas has a positive influence on overall survival.

2 Methods

First of all, it must be mentioned that all computational tasks concerning the use of the neuronal network, which needs expert computational knowledge, were performed by **Marcus Bloice,** Institute for Medical Informatics, Statistics, and Documentation, Medical University of Graz.

2.1 Survival analysis

An univariate data analysis with SPSS was accomplished. Frequency and distribution of the gene mutations was determined. Subsequently, the overall survival in the different groups (HRD negative/ HRD positive) was examined in detail. A Kaplan-Meier curve was plotted in SPSS and a p-value was determined by log-rank test.

2.2 Image data

2.2.1 Data gathering

Our data source is the TCGA. We examined 436 cases from the TCGA-OV, containing patient cases with serous cystadenocarcinoma. From this database, we could extract mutation data, clinical data and one to six histological whole slide images (WSI) in a specific large scale image format, called SVS format, for each case. Most of the images were slides from frozen tissues, which have undergone morphological changes during preparation.

Only 106 out of 436 cases had diagnostic slides available. Diagnostic slides consist of formalin-fixed paraffin-embedded tissues and consequently prove high quality. In total we had 1091 whole slide images at our disposal, see diagram below. Each whole slide image has a size of 1-2GB.

2.2.2 Data pre-processing

Some steps of pre-processing were necessary in order to train the neuronal network. At first, the images were tiled to provide smaller segments. At a certain magnification, cells become better recognizable and the classification into tumour or non-tumour is more accurate, hence the tiling. Around 14 million tiles were created; each tile with a size of 224x224 pixels.

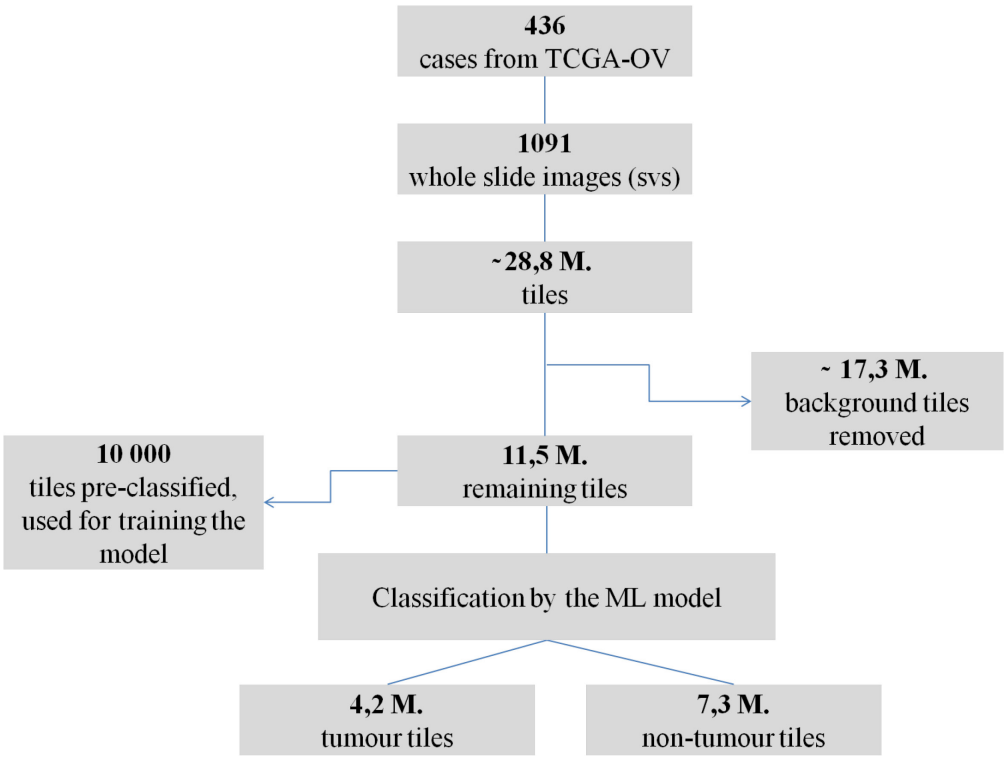

Figure 16: Number of tiles for the classification into tumour and non-tumour.

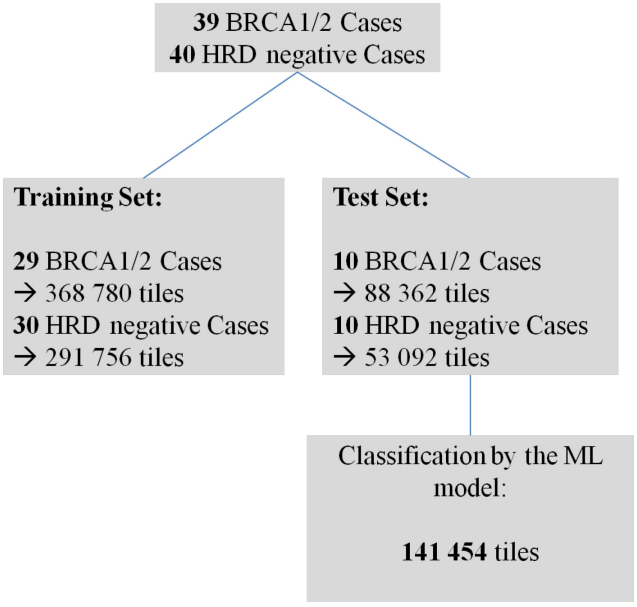

Figure 17: Number of tiles for classification between BRCA1/2 and HRD negative.

To get the best possible data set of images, it was necessary to sort out empty tiles. “Empty tiles” mean tiles, which are not meaningful, because they provide no or useless

information. These tiles often come from the edge of the whole slides where only background exists. The following figure serves as example for a background tile.

Figure 18: Background tile with a small cellular contamination, but without further information

Sorting was done by converting the tiles to greyscale and calculating their average pixel colour. The average pixel value is calculated by dividing the sum of the colour value of all pixels by the number of pixels in the image. The result is a value between 0 = black and 255 = white. A threshold of 220 was defined. All images with an average pixel value above this threshold were thrown out. Images below this threshold were considered to be tissue.

2.2.3 Classification into tumour or non-tumour

After removing the useless background tiles, 11 524 672 tiles were left. About 10 000 of them were classified by Peter Regitnig, Diagnostic- and Research Institute of Pathology, into tumour, no tumour or inconclusive (out of focus, edge artefacts, colour marks), by the help of the tile tagger web application, developed for this purpose by Marcus Bloice.

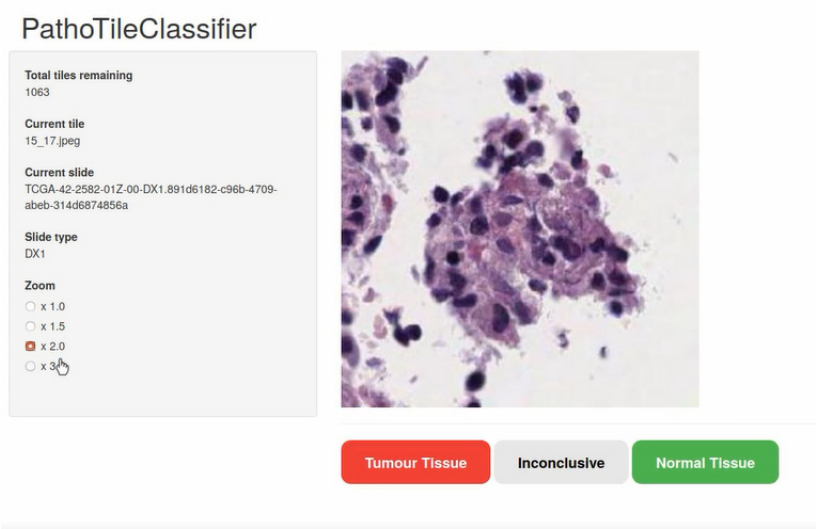

Figure 19: Tile Classifier

Once the 10,000 tiles were classified, these tagged tiles were used to train a neural network to automatically classify the remaining 11.5 million tiles concerning their malignancy.

The model classed around 4 million tiles as tumour tiles with a confidence threshold of 0.7, meaning it is 70% confident the classification is right and there is a 30% probability the image belongs to class “normal tissue”. With this threshold the network performed very well and the data set was not too small as it would have been due to many false negatives at a confidence of 80% or more. The disadvantage, though, is that there were a few false positives in it.

This step was useful because it reduced the size of the data set: for further training of the model only tumour tiles were required. The neural network that was in use is labelled a ResNet-101. It is a pre-trained model that has already been trained with an ImageNet dataset. Many everyday objects like cars, trees, animals are included in this dataset. Although this seems akward, it has been shown that such pre-trained image neural networks are also suitable to differentiate histopathological images. Our image tiles have a size of 224x224 pixels and the model is a very large and deep one with 101 layers. 8 000 of the 10 000 tiles were used for training. The remaining 2000 tiles were subsequently used for testing.

2.2.4 Gene predictor model

HRD status is defined by the presence of various mutations, first and foremost a BRCA 1 or BRCA 2 mutation. If *BRCA* is mutated, it is very likely that HRD is also present (49).

In this step, the neuronal network should learn to recognise from the tiles classified as malignant from the previous step whether such a mutation is present or not. If so, this case would be considered to be HRD positive.
For this purpose, the model was given 368 780 BRCA1/BRCA2 tiles and 291 756 HRD = negative tiles for training. The subsequent test was performed with 88,362 BRCA1/BRCA2 tiles and 53 092 HRD = negative tiles, respectively. The network was trained in 4 epochs and should classify BRCA mutations from cases with a negative HRD status. The model is the ResNet18 model that also belongs to the Convolutional Neuronal Networks. The related number 18 means, that the data go through 18 layers. It is a pre-trained programme that has already been trained with several million images from everyday life and can therefore classify images into a category on the basis of various features (67).
The input data are always images with a pixel size of 224x224. During training, each pixel is given weights that can then be changed in the so-called unfrozen layers of the CNN depending on how well they match the output data. In the frozen layers, the weighting can no longer be changed. The more layers are frozen, the less can be changed in the weights. However, if all layers were left unfrozen, the training would take far too long. To classify BRCA vs. HRD negative, only the late layers were initially left trainable. It is common practice to freeze most weights in a network if using a pre-trained network, so that the first few epochs train only the later layers of the network, and once this is stable the network is unfrozen and the network is trained again, typically for 1 or 2 further epochs. A complete run of all input data is called an epoch (68).
The accuracy for predicting BRCA mutations was calculated and a confusion matrix for BRCA1/BRCA2 vs. HRD negative was then created.

2.3 Mutation data

The mutation data of the individual cases could be downloaded the same way from the TCGA as the whole slides. As described in section *DNA repair mechanisms* only four genes were defined as so called “HRD genes”.
Based on this a case table was created, restricted to these relevant HRD genes. Figure 20 shows exemplarily the first fifteen rows of this table.

Case_Numbe	Tumor_Sample_Barcode	BRCA1	BRCA2	PTEN	CHEK1	Sum_of_Gene_Occurence	HRD
1	TCGA-04-1331-01A-01W-0486-08	0	1	0	0	1	1
2	TCGA-04-1332-01A-01W-0488-09	0	0	0	0	0	0
3	TCGA-04-1336-01A-01W-0486-08	0	0	0	0	0	0
4	TCGA-04-1341-01A-01W-0486-08	0	0	0	0	0	0
5	TCGA-04-1342-01A-01W-0486-08	0	0	0	0	0	0
6	TCGA-04-1343-01A-01W-0486-08	0	0	0	0	0	0
7	TCGA-04-1346-01A-01W-0486-08	0	0	0	0	0	0
8	TCGA-04-1347-01A-01W-0488-09	0	0	0	0	0	0
9	TCGA-04-1348-01A-01W-0492-08	0	0	0	0	0	0
10	TCGA-04-1349-01A-01W-0492-08	0	0	0	0	0	0
11	TCGA-04-1353-01A-01D-1526-09	0	0	0	0	0	0
12	TCGA-04-1356-01A-01W-0492-08	0	1	0	0	1	1
13	TCGA-04-1357-01A-01W-0492-08	1	0	0	0	1	1
14	TCGA-04-1360-01A-01W-0492-08	0	0	0	0	0	0
15	TCGA-04-1361-01A-01W-0492-08	0	0	0	0	0	0

Figure 20: Case table, HRD status. HRD = 0 stands for a negative HRD status and HRD = 1 for an existing HRD.

The individual cases are identified via the tumour sample barcode. If one of the listed mutations was present, this is represented by a "1" in the corresponding column. As soon as one of the HRD genes is present in a case, the HRD status was set as positive, i.e. marked with a "1". Using this table, the HRD status was assigned to the images. Whenever a mutation was present in any of the named genes, the case and the related whole slides were considered HRD positive.

3 Results

3.1 Frequence of HRD positive cases

Across all 436 cases, 54 cases showed a mutation in one of the defined HRD genes. This corresponds to a percentage of about 12%. BRCA1 is the most common mutation; it affects 24 of the 54 HRD positive cases.

HRD		
	Frequency	Percent
HRD -	382	87,6
HRD +	54	12,4
Total	436	100,0

Table 1: Occurence of HRD among all 436 HGSOC Cases

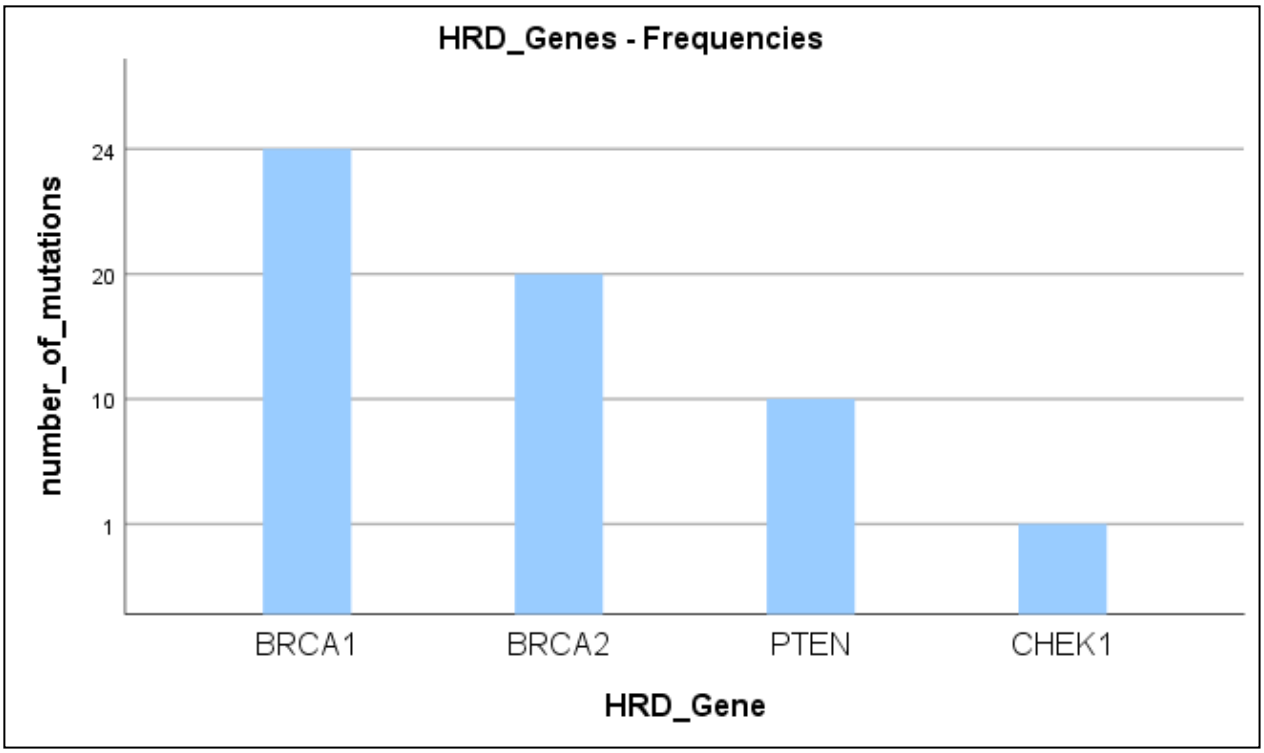

Figure 21: Frequency distribution among the HRD genes

3.2 Overall survival correlation to HRD

The Kaplan-Meier curve in figure 22 shows the overall survival in the two groups (HRD positive = 1, HRD negative = 0) over a time of around 5000 days to last follow up. A clear difference in the survival of the two groups is evident. The HRD positive cases, shown as red curve, have a higher probability of survival than the HRD negative cases at any time. Furthermore, the graph displays the one-year, five-year and ten-year survival in the two groups. Survival rates at one-, five and ten years within the HRD positive group were 92%, 55% and 25%, respectively, while the HRD negative group had rates of 84%, 34% and 18%.

The corresponding logrank test in table 2 yielded a p-value of p = 0,024 (significance level α = 0.05). Hence the difference shown in the diagram is significant.

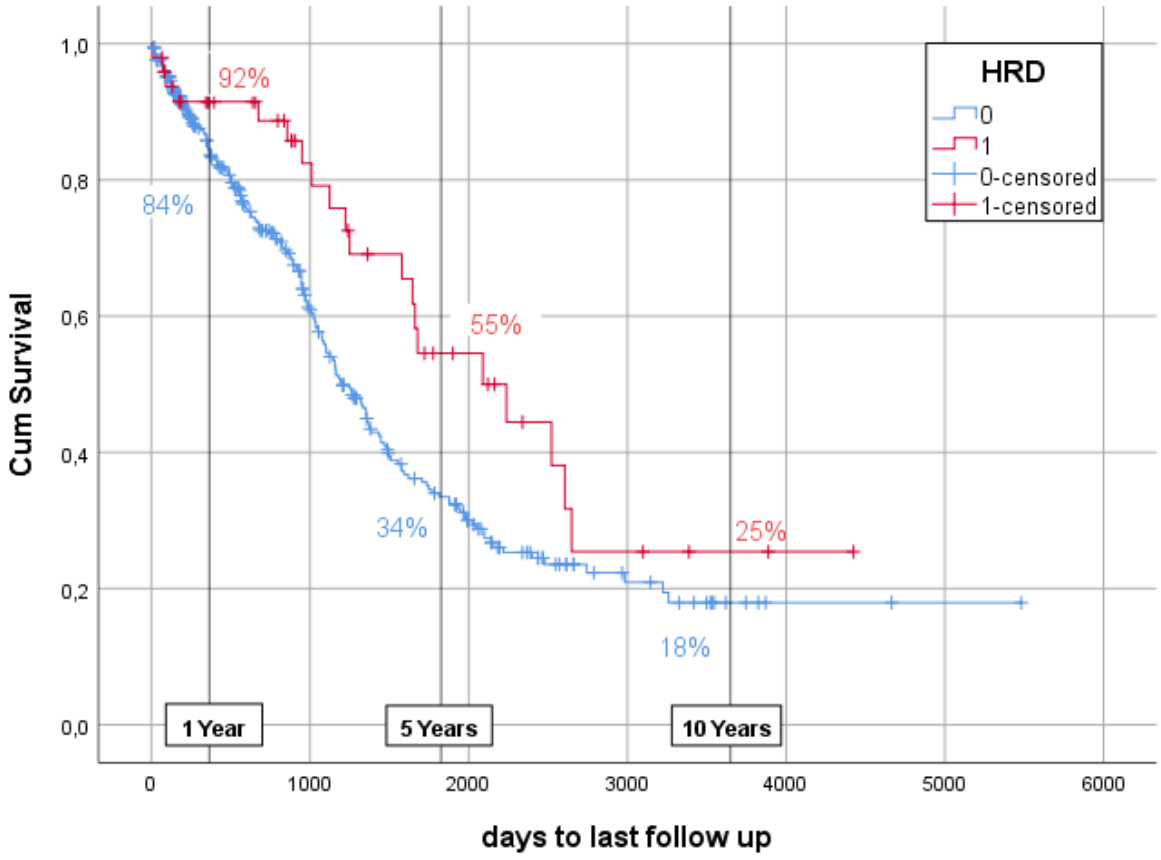

Figure 22: Kaplan-Meier curve

Overall Comparisons			
	Chi-Square	df	Sig.
Log Rank (Mantel-Cox)	5,115	1	,024
Test of equality of survival distributions for the different levels of HRD.			

Table 2: LogRank Test

		HRD		
		negative	positive	Total
vital_status	no data	1	0	1
	alive	156	30	186
	dead	224	24	248
	No follow up	1	0	1
Total		382	54	436

Table 3: Crosstabulation regarding the vital status in HRD groups.

With the help of the cross table (Table 3) a hazard ratio of 1.32 has been calculated. Having no mutation in one of the HRD genes is thus 1.3 times worse regarding prognosis than with mutation.

3.3 Prediction of tumour in image tiles

The 10 000 pre-classified tiles were at the disposal of the model. 8000 tiles were used for training, the remaining 2000 for the final test. The model achieved an accuracy of 83.9%. This means that with a probability of almost 84%, it correctly predicted whether the tissue was tumour or not. The confidence threshold was 0.7, precision 0.92 and recall 0.80. Table number 4 lists the different values. For this confidence threshold of 0.7, only predictions by the model were accepted as correct if the confidence of the network was over 0.7. Precision indicates the rate at which tumour predictions are correct. In this case 92% of the tumour prediction was correct. Recall indicates the rate at which non-tumor tiles were classified as non-tumour. So 80% of non tumour tiles were correctly classified as non-tumour. By raising the threshold precision becomes very high, but recall becomes very low.

Confidence threshold	0.7
Classified as tumour	755
Is actually tumour	682
Precision	0.92
Recall	0.80

Table 4: Results of malignancy prediction

3.4 Prediction of BRCA mutations

A classification of tiles with BRCA mutations versus HRD negative tiles was performed. In the experiment, the prediction of BRCA mutations by the neural network was revealed to have an accuracy of 45.3%. While approximately 32 000 out of 88 000 tiles were correctly classified as BRCA, approximately 21 000 out of 53 000 tiles were incorrectly classified as BRCA. Figure 23 shows some details of the model run and figure 24 shows the corresponding confusion matrix.

Model: ResNet18

Epochs: 4 (One cycle learning)

Data Augmentation: No

BRCA1/BRCA2 training set cases: 29 (75 Whole Slides), test cases: 10 (24 Whole Slides)
HRD0 training set cases: 30 (74 Whole Slides), test cases: 10 (24 Whole Slides0)
Total: 197 Whole Slides

Training Set: 660,536 tiles
Test Set: 141,454 tiles
Total: 801,990 tiles

368,780 BRCA1/BRCA2 training tiles.
88,362 BRCA1/BRCA2 test tiles.

291,756 HRD0 training tiles.
53,092 HRD0 test tiles.

Figure 23: Details of the model run.

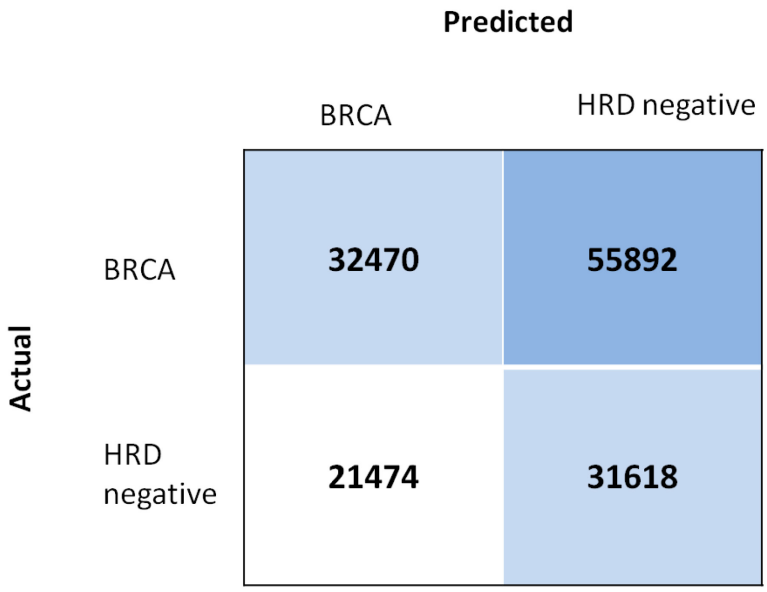

Figure 24: Confusion matrix for BRCA1/BRCA2 positive vs. HRD negative. Upper left and lower right are those tiles that were correctly classified.

The accuracy has been calculated as follows (69):

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN}$$

$$\text{Accuracy BRCA} = \frac{32470+31618}{32470+31618+21474+55892} = \mathbf{0.45}$$

	precision	recall	f1-score	support
BRCA	0.60	0.37	0.46	88362
HRD0	0.36	0.60	0.45	53092
accuracy			0.45	141454
macro avg	0.48	0.48	0.45	141454
weighted avg	0.51	0.45	0.45	141454

Figure 25: Classification report

This translates in Sensitivity and Specifity as follows:

$$\text{Sensitivity} = \frac{TP}{TP+FN}$$

$$\text{Sensitivity of BRCA prediction} = \frac{32470}{32470+55892} = \mathbf{0.37}$$

$$\text{Specifity} = \frac{TN}{TN+FP}$$

$$\text{Specifity of BRCA prediction} = \frac{31618}{31618+21474} = \mathbf{0.60}$$

$$\text{f1-Score} = \frac{2x\,TP}{2x\,TP+FP+FN} = \mathbf{0.46}$$

The f1-Score also named harmonic mean is an overall score that depicts how well the machine learning model is performing by combining both precision and recall. It gives a realistic value considering both values. It is necessary to have a high precision **and** a high recall to get a high f1. The advantage of f1-score over accuracy is that the data set can be unbalanced. This means that we do not have biased results even if the number of samples in one class is larger than in the other (70).

4 Discussion

A reliable prediction of the HRD status by means of histological slides could not be proven in this work. On the other hand, the prediction of the malignancy of the tissue was far better. In any case, the work has shown that there can be great value in artificial intelligence for use in medical research.

In 54 of 436 cases (12%), one or more of the HRD genes were mutated which indicates a homologous recombination deficiency. The most frequently mutated gene was the *BRCA1* gene. A significant difference in survival rates between HRD positive and HRD negative groups exists ($p = 0{,}024$, significance level $\alpha = 0.05$). A comparison of the survival rates at one-, five and ten years suggests that it is more favourable to carry a HRD gene mutation. One possible reason for this is the better response of patients with HRD to chemotherapy: Similarily, it has been found that breast cancer patients with standard anthracycline- and

taxane-based neoadjuvant therapy show a better response to the therapy in case the patient is HRD positive compared to *BRCA1/2* wild-type patients. With chemotherapies damaging the DNA and the simultaneously inactive homologous recombination repair mechanism, tumour cells perish more easily. (71)

Also Myriad, a molecular diagnostic company, describes that patients with HRD respond better to drugs that can affect DNA stability, like PARP inhibitors or platinum based chemotherapy (42).

For predicting the malignancy our deep neuronal network achieved an accuracy of about 84% in the test set with a confidence threshold of 0.7. The performance of the prediction of the HRD status was inferior. Only about 45 % accuracy could be achieved. A major reason for that result might be the use of only too few cases. CNN usually need a big amount of training data to work properly. An effective training of the network was therefore not possible. However, one small success could be noted. The specifity of the BRCA prediction resulted in 60% that means HRD negative cases were correctly not classified as BRCA at a probability of 60%. With such a small amount of data, this suggests that stronger results could be achieved if appropriate measures are taken.

So far, only limited literature on this topic was found to be available. However, there are works that have described something similar. As example Coudray et al., describe in their paper “Classification and mutation prediction from non–small cell lung cancer histopathology images using deep learning” how they proceed to train the Inception Net on mutations in non-small cell lung cancer (65). This research group trained the deep learning model to classify a tissue into two different carcinoma types of the lung or into normal tissue by using whole-slide images from TCGA. This functioned successfully with an average area under the curve of 0.97. In addition, they trained their network to predict the ten most frequent mutations in adenocarcinoma of the lung, by having exclusively slides as input. In this task, they reached an average area under the curve of 0.733 to 0.856.

Furthermore, there are studies that show our thoughts in regard to the effectiveness of PARP inhibitors, but this in endometrial carcinoma. A women with stage IVB (T1bNxM1) high-grade serous endometrial carcinoma got both first-line and second-line chemotherapy. The disease stayed stable under those therapies. No improvement could be shown. A gene analysis revealed a homologous recombination disorder. Thus, an olaparib therapy (PARP inhibitor) was given. Under this therapy, a decrease in tumour size and swelling of the regional lymph nodes was observed. (72)

Several limitations of our study should be considered. First, during literature research, it was noticed that there are different opinions about which genes can be classified as HRD genes and which cannot. While Heeke et al. determined *ARID1A, ATM, ATRX, BAP1, BARD1, BLM, BRCA1/2, BRIP1, CHEK1/2, FANCA/C/D2/E/F/G/L, MRE11A, NBN, PALB2, RAD50, RAD51, RAD51B,* and *WRN* as HRD genes (40), in the ARIEL3 study even more genes were declared as HRD(73). It was therefore necessary to study the literature thoroughly in order to decide which of these genes should be considered as HRD genes for further investigation. A recent paper Takaya et al. shows that actually only some of the declared HRD genes really have an effect on the cell. Actually, only *BRCA1, BRCA2, PTEN* and *CHEK1* would effectively influence the tumorigenesis, progression and sensitivity to PARP inhibitors (43). Therefore, it made sense to limit the selection of the HRD genes to which the neural network should be trained to these four. If the number and type of these defined genes changes, this logically leads to different results, both in terms of statistical analysis and the outcome of the neural network.

Moreover, in terms of data collection, a difficulty arose regarding the quality and quantity of histological slides. Although the TCGA database provides a lot of histological slide images corresponding to the patient cases, only a small part of them are formalin-fixed paraffin-embedded tissues and consequently prove high quality. The rest are frozen tissues, which have undergone changes in appearance during preparation. Since the neural network should be fed exclusively with high quality slides, we first wanted to limit our choice to the so called *DX slides* (paraffin-embedded). Only in 75 of 436 cases those high quality slides were available. Unfortunately, that was not enough, so we decided to include even the inferior frozen section slides. Overall it must be said that the small amount of cases was the main problem. With only 39 BRCA mutated cases, the sample size was very small, making it extremely difficult to train the model sufficiently. As a result, the accuracy score remained very low. In addition, it must be considered that mutation data may have been lost due to the creation of masked MAF files during preparation of the DNA. This of course distorts the data set and can lead to poor ML performance. Although the number of BRCA mutations found is plausible and reflects the data from the literature, it cannot be assumed with certainty that all of the other mutations were registered.

Furthermore, running the NN in the absence of a highly performing hardware was challenging. The hardware was not equipped to tune the network by an exhaustive “grid search” of hyper-parameters. That’s why an optimal setup, which could increase the accuracy, was not feasible. The experiment, which is considered the first approach in this

still unexplored field, has shown that promising results can be achieved by reducing the limitations.

In summary, this work presents an early approach to medical research using artificial intelligence. However, the prediction of HRD status by a neuronal network, which remains a very exciting and important project, warrants confirmation in future studies.

5 References

1. IARC. World fact sheets 2018 [Available from: http://gco.iarc.fr/today/data/factsheets/populations/900-world-fact-sheets.pdf.
2. Cancer IAfRo. Global Cancer Observatory [Available from: http://gco.iarc.fr/.
3. Austria A. MANUAL DER GYNÄKOLOGISCHEN ONKOLOGIE: Christian Schauer, Alain G. Zeimet; 2021 [Available from: https://ago-austria.at/epitheliales-ovarialkarzinom-atiologie-risikofaktoren/.
4. S3-Leitlinie Diagnostik, Therapie und Nachsorge maligner Ovarialtumoren 2019 [Available from: https://www.awmf.org/uploads/tx_szleitlinien/032-035OL1_S3_Maligne-Ovarialtumoren_Diagnostik-Therapie-Nachsorge_2019-03.pdf.
5. Moch GHHKH. Pathologie. 6 ed2019.
6. Fathalla MF. Incessant ovulation--a factor in ovarian neoplasia? Lancet (London, England). 1971;2(7716):163.
7. Fathalla MF. Incessant ovulation and ovarian cancer - a hypothesis re-visited. Facts, views & vision in ObGyn. 2013;5(4):292-7.
8. Kroeger PT, Jr., Drapkin R. Pathogenesis and heterogeneity of ovarian cancer. Current opinion in obstetrics & gynecology. 2017;29(1):26-34.
9. Cannioto RA, Trabert B, Poole EM, Schildkraut JM. Ovarian cancer epidemiology in the era of collaborative team science. Cancer Causes Control. 2017;28(5):487-95.
10. Kurman RJ. 2014.
11. Kim J, Park EY, Kim O, Schilder JM, Coffey DM, Cho CH, et al. Cell Origins of High-Grade Serous Ovarian Cancer. Cancers (Basel). 2018;10(11).
12. Masoodi T, Siraj S, Siraj AK, Azam S, Qadri Z, Parvathareddy SK, et al. Genetic heterogeneity and evolutionary history of high-grade ovarian carcinoma and matched distant metastases. British Journal of Cancer. 2020;122(8):1219-30.
13. Natarajan S, Foreman KM, Soriano MI, Rossen NS, Shehade H, Fregoso DR, et al. Collagen Remodeling in the Hypoxic Tumor-Mesothelial Niche Promotes Ovarian Cancer Metastasis. Cancer Res. 2019;79(9):2271-84.
14. Razak Amanullah N, Poothiode U, Vilasiniamma L. Expression of p53 in epithelial ovarian tumors. 2020;63(2):235-40.

15. Ramalingam P. Morphologic, Immunophenotypic, and Molecular Features of Epithelial Ovarian Cancer. Oncology (Williston Park, NY). 2016;30(2):166-76.
16. Netinatsunthorn W, Hanprasertpong J, Dechsukhum C, Leetanaporn R, Geater A. WT1 gene expression as a prognostic marker in advanced serous epithelial ovarian carcinoma: an immunohistochemical study. BMC Cancer. 2006;6:90-.
17. Kotake Y, Naemura M, Murasaki C, Inoue Y, Okamoto H. Transcriptional Regulation of the p16 Tumor Suppressor Gene. Anticancer Research. 2015;35(8):4397.
18. Raj N, Attardi LD. The Transactivation Domains of the p53 Protein. Cold Spring Harb Perspect Med. 2017;7(1):a026047.
19. Lisio MA, Fu L, Goyeneche A, Gao ZH, Telleria C. High-Grade Serous Ovarian Cancer: Basic Sciences, Clinical and Therapeutic Standpoints. International journal of molecular sciences. 2019;20(4).
20. Dubeau L. The cell of origin of ovarian epithelial tumors and the ovarian surface epithelium dogma: does the emperor have no clothes? Gynecologic oncology. 1999;72(3):437-42.
21. Finch A, Shaw P, Rosen B, Murphy J, Narod SA, Colgan TJ. Clinical and pathologic findings of prophylactic salpingo-oophorectomies in 159 BRCA1 and BRCA2 carriers. Gynecologic oncology. 2006;100(1):58-64.
22. Lee Y, Miron A, Drapkin R, Nucci MR, Medeiros F, Saleemuddin A, et al. A candidate precursor to serous carcinoma that originates in the distal fallopian tube. The Journal of pathology. 2007;211(1):26-35.
23. Nik NN, Vang R, Shih I-M, Kurman RJ. Origin and Pathogenesis of Pelvic (Ovarian, Tubal, and Primary Peritoneal) Serous Carcinoma. 2014;9(1):27-45.
24. Przybycin CG, Kurman RJ, Ronnett BM, Shih Ie M, Vang R. Are all pelvic (nonuterine) serous carcinomas of tubal origin? The American journal of surgical pathology. 2010;34(10):1407-16.
25. Li HX, Lu ZH, Shen K, Cheng WJ, Malpica A, Zhang J, et al. Advances in serous tubal intraepithelial carcinoma: correlation with high grade serous carcinoma and ovarian carcinogenesis. International journal of clinical and experimental pathology. 2014;7(3):848-57.
26. Corzo C, Iniesta MD, Patrono MG, Lu KH, Ramirez PT. Role of Fallopian Tubes in the Development of Ovarian Cancer. Journal of minimally invasive gynecology. 2017;24(2):230-4.

27. Kurman RJ, Shih Ie M. The Dualistic Model of Ovarian Carcinogenesis: Revisited, Revised, and Expanded. Am J Pathol. 2016;186(4):733-47.
28. Lengyel E. Ovarian Cancer Development and Metastasis. The American Journal of Pathology. 2010;177(3):1053-64.
29. Kurman RJ, Shih I-M. The origin and pathogenesis of epithelial ovarian cancer: a proposed unifying theory. The American journal of surgical pathology. 2010;34(3):433-43.
30. 2014.
31. Liang M, Zhao J. Protein expressions of AIB1, p53 and Bcl-2 in epithelial ovarian cancer and their correlations with the clinical pathological features and prognosis. European review for medical and pharmacological sciences. 2018;22(16):5134-9.
32. McAlpine JN, Porter H, Köbel M, Nelson BH, Prentice LM, Kalloger SE, et al. BRCA1 and BRCA2 mutations correlate with TP53 abnormalities and presence of immune cell infiltrates in ovarian high-grade serous carcinoma. Modern Pathology. 2012;25(5):740-50.
33. Moschetta M, George A, Kaye SB, Banerjee S. BRCA somatic mutations and epigenetic BRCA modifications in serous ovarian cancer. Ann Oncol. 2016;27(8):1449-55.
34. Ledermann JA, Drew Y, Kristeleit RS. Homologous recombination deficiency and ovarian cancer. Eur J Cancer. 2016;60:49-58.
35. Krokan HE, Bjørås M. Base excision repair. Cold Spring Harb Perspect Biol. 2013;5(4):a012583-a.
36. Easton DF, Lesueur F, Decker B, Michailidou K, Li J, Allen J, et al. No evidence that protein truncating variants in BRIP1 are associated with breast cancer risk: implications for gene panel testing. J Med Genet. 2016;53(5):298-309.
37. Marth C. Key Concepts of PARP Inhibition.
38. John M. Philosophical Logic and Artificial Intelligence. 1 ed: Kluwer Academic Publisher; 1989. 222 p.
39. Alfred Nordheim RK. Molekulare Genetik. 10 ed2015.
40. Heeke AL, Pishvaian MJ, Lynce F, Xiu J, Brody JR, Chen WJ, et al. Prevalence of Homologous Recombination-Related Gene Mutations Across Multiple Cancer Types. JCO Precis Oncol. 2018;2018.
41. Detect somatic and germline mutations for HRR Deficiency [Available from: https://www.agilent.com/en/product/next-generation-sequencing/amplicon-based-next-generation-sequencing-ngs/onco-genetics-amplicon-amplification/suremastr-hrr-

764821?elqTrackId=0c27deb619bf405a948abe0a875970b3&elq=598ac51f984c4e5aa6110dc400062e34&elqaid=5110&elqat=1&elqCampaignId=4192#productdetails.
42. Genetics M. myChoice HRD® Update 2016 [Available from: https://investor.myriad.com/static-files/8ef72264-ca8c-4b01-b0e9-35fa83c63889.
43. Takaya H, Nakai H, Takamatsu S, Mandai M, Matsumura N. Homologous recombination deficiency status-based classification of high-grade serous ovarian carcinoma. Sci Rep. 2020;10(1):2757.
44. Telli ML, Timms KM, Reid J, Hennessy B, Mills GB, Jensen KC, et al. Homologous Recombination Deficiency (HRD) Score Predicts Response to Platinum-Containing Neoadjuvant Chemotherapy in Patients with Triple-Negative Breast Cancer. Clinical Cancer Research. 2016;22(15):3764.
45. K. Aktories UF, F. Hofmann, K. Starke. Allgemeine und spezielle Pharmakologie und Toxikologie. 12 ed2017.
46. Lee JM, Ledermann JA, Kohn EC. PARP Inhibitors for BRCA1/2 mutation-associated and BRCA-like malignancies. Ann Oncol. 2014;25(1):32-40.
47. Taylor KN, Eskander RN. PARP Inhibitors in Epithelial Ovarian Cancer. Recent Pat Anticancer Drug Discov. 2018;13(2):145-58.
48. Gelmon KA, Tischkowitz M, Mackay H, Swenerton K, Robidoux A, Tonkin K, et al. Olaparib in patients with recurrent high-grade serous or poorly differentiated ovarian carcinoma or triple-negative breast cancer: a phase 2, multicentre, open-label, non-randomised study. The Lancet Oncology. 2011;12(9):852-61.
49. da Cunha Colombo Bonadio RR, Fogace RN, Miranda VC, Diz M. Homologous recombination deficiency in ovarian cancer: a review of its epidemiology and management. Clinics (Sao Paulo). 2018;73(suppl 1):e450s.
50. Michels J, Vitale I, Galluzzi L, Adam J, Olaussen KA, Kepp O, et al. Cisplatin Resistance Associated with PARP Hyperactivation. Cancer Research. 2013;73(7):2271.
51. Mirza MR, Monk BJ, Herrstedt J, Oza AM, Mahner S, Redondo A, et al. Niraparib Maintenance Therapy in Platinum-Sensitive, Recurrent Ovarian Cancer. 2016;375(22):2154-64.
52. Timmerman D, Van Calster B, Testa AC, Guerriero S, Fischerova D, Lissoni AA, et al. Ovarian cancer prediction in adnexal masses using ultrasound-based logistic regression models: a temporal and external validation study by the IOTA group. Ultrasound in obstetrics & gynecology : the official journal of the International Society of Ultrasound in Obstetrics and Gynecology. 2010;36(2):226-34.

53. Moon A, Bourdeth A, Jerez R, Alger J, Chuang L. Evaluation of Ovarian Neoplasms in Honduras: Characteristics and Diagnostic Concordance Between Ultrasound, Tumor Markers and Histopathology. Gynecologic Oncology Reports. 2019;30:100501.
54. Penson RT. Harrisons Manual of Oncology 2nd Ed.
CHAPTER 54
Ovarian Cancer [Available from: https://doctorlib.info/oncology/harrisons-manual-oncology/54.html.
55. Wang Z, Jensen MA, Zenklusen JC. A Practical Guide to The Cancer Genome Atlas (TCGA). Methods in molecular biology (Clifton, NJ). 2016;1418:111-41.
56. Institute NC. The Cancer Genome Atlas Program - National Cancer Institute [Available from: https://www.cancer.gov/about-nci/organization/ccg/research/structural-genomics/tcga.
57. Liang Z, Powell A, Ersoy I, Poostchi M, Silamut K, Palaniappan K, et al., editors. CNN-based image analysis for malaria diagnosis. 2016 IEEE International Conference on Bioinformatics and Biomedicine (BIBM); 2016 15-18 Dec. 2016.
58. Rajaraman S, Antani SK, Poostchi M, Silamut K, Hossain MA, Maude RJ, et al. Pre-trained convolutional neural networks as feature extractors toward improved malaria parasite detection in thin blood smear images. PeerJ. 2018;6:e4568.
59. Scerri M, Grech V. Artificial intelligence in medicine. Early human development. 2020;145:105017.
60. Lison PJLTG, 1. An introduction to machine learning. 2015;35.
61. Was ist Machine Learning? [Available from: https://www.bigdata-insider.de/was-ist-machine-learning-a-592092/.
62. Understanding Neural Networks: What, How and Why? [Available from: https://towardsdatascience.com/understanding-neural-networks-what-how-and-why-18ec703ebd31.
63. Zhao Z-Q, Zheng P, Xu S-t, Wu X. Object Detection with Deep Learning: A Review. CoRR. 2018;abs/1807.05511.
64. PARLOFF R. From 2016: Why Deep Learning Is Suddenly Changing Your Life. Fortune. 2016.
65. Coudray N, Ocampo PS, Sakellaropoulos T, Narula N, Snuderl M, Fenyo D, et al. Classification and mutation prediction from non-small cell lung cancer histopathology images using deep learning. Nature medicine. 2018;24(10):1559-67.

66. Ritterhouse LL, Nowak JA, Strickland KC, Garcia EP, Jia Y, Lindeman NI, et al. Morphologic correlates of molecular alterations in extrauterine Müllerian carcinomas. Modern Pathology. 2016;29(8):893-903.
67. resnet101 [Available from: https://de.mathworks.com/help/deeplearning/ref/resnet101.html;jsessionid=81a2bcf4010e5d7ccd8b5731bb4b.
68. Sagar R. What Does Freezing A Layer Mean And How Does It Help In Fine Tuning Neural Networks [Available from: https://analyticsindiamag.com/what-does-freezing-a-layer-mean-and-how-does-it-help-in-fine-tuning-neural-networks/.
69. Baratloo A, Hosseini M, Negida A, El Ashal G. Part 1: Simple Definition and Calculation of Accuracy, Sensitivity and Specificity. Emerg (Tehran). 2015;3(2):48-9.
70. Chicco D, Jurman G. The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. BMC Genomics. 2020;21(1):6-.
71. Telli ML, Hellyer J, Audeh W, Jensen KC, Bose S, Timms KM, et al. Homologous recombination deficiency (HRD) status predicts response to standard neoadjuvant chemotherapy in patients with triple-negative or BRCA1/2 mutation-associated breast cancer. Breast Cancer Research and Treatment. 2018;168(3):625-30.
72. Nakamura K, Aimono E, Tanishima S, Imai M, Nagatsuma AK, Hayashi H, et al. Olaparib Monotherapy for BRIP1-Mutated High-Grade Serous Endometrial Cancer. 2020(4):283-90.
73. Coleman RL, Oza AM, Lorusso D, Aghajanian C, Oaknin A, Dean A, et al. Rucaparib maintenance treatment for recurrent ovarian carcinoma after response to platinum therapy (ARIEL3): a randomised, double-blind, placebo-controlled, phase 3 trial. Lancet (London, England). 2017;390(10106):1949-61.