The Ploidy Studies Of Ovarian Cancer: Searching For New Methods Of Diagnosis

Robert Lapcevic
Carroll College, Helena, MT
THE PLOIDY STUDIES OF

OVARIAN CANCER:

SEARCHING FOR NEW METHODS OF DIAGNOSIS

Submitted in Partial Fulfillment of the Requirements of Graduation with Honors
to the Department of Mathematics at Carroll College, Helena, Montana

Robert Daniel Lapcevic

April 9th, 1998
Advisor: 
Marie Vanisko
Mathematics and Engineering

Reader: 
Murphy Fox
Sociology and Literature

Reader: 
Marilyn Schendel
Biology
Acknowledgements

There are so much thanks needing to be said for the work that went into this thesis. First of all, I must thank Dr. Sam Ballon for giving me the research position this past summer, without it none of this will be possible. Also, all of his colleagues and staff at the Women’s Cancer Institute of Northern California, for all the help with records, filing, and stress relief. Next, there is a whole group at Carroll College to thank, my advisor, Marie Vanisko, and my readers, Murphy Fox and Marilyn Schendel. Their patience, caring, and understanding were what made all of this possible, especially as the due date neared. Also to all of my friends and family, who, along with the aforementioned people, pushed me to work hard and finish what I started. Thanks again Carroll College for the past four years of memories.
# Table of Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page #</th>
</tr>
</thead>
<tbody>
<tr>
<td>Purpose</td>
<td>5</td>
</tr>
<tr>
<td>Introduction to Genetics</td>
<td>6</td>
</tr>
<tr>
<td>Literary Review</td>
<td>11</td>
</tr>
<tr>
<td>Introduction to Ovarian Cancer</td>
<td>14</td>
</tr>
<tr>
<td>Introduction to My Study</td>
<td>18</td>
</tr>
<tr>
<td>Data and Results</td>
<td>20</td>
</tr>
<tr>
<td>Conclusion</td>
<td>29</td>
</tr>
<tr>
<td>Works Cited</td>
<td>31</td>
</tr>
</tbody>
</table>
Purpose

This paper stems from research information accumulated over the summer of 1997 on the topic of ovarian cancer. My role in this was statistical accumulation, computation, and initial analysis. The study appealed to me due to the intertwined usage of math and biology it entailed. Specifically, a background in biology aided in comprehending the affliction of ovarian cancer and the underlying genetics that control it. The original purpose of this study involved compiling statistics on the symptoms, pathology reports, and flow cytometry data accumulated over the past few years from ovarian cancer patients at the Women’s Cancer Center of Northern California. From this, a new method of classifying the extent of the disease in new patients would be developed. The results of analyzing this data could improve upon the current method of diagnosis, which involves stages and grades. This would allow the treating physician a better chance at informing the patient as to the severity of the disease from which she suffers. It would also allow the patient to be more informed as to the chances of survival, the complications, and the realism of reoccurrence. The specific genes that maintain cell cycles and their division, possible mutations, and how all of it is regulated through the expression of certain genetic elements are explained in this paper. The possible benefits of this study include more thorough examinations, better treatments, and informed prognoses. These ideas are experimental, yet their implications could bring about major changes in the treatments of ovarian cancer.
Introduction to Genetics

The human body is a complex machine with a complex origin. The combination of an egg from a female with the sperm of a male results in fertilization. The resulting zygote becomes a morula, a large number of cells arranged in a ball, in the first six days, and then a blastocyst, a hollow ball of cells, over the next two days (8). It implants itself into the wall of the endometrium, and by the time four weeks has passed, the embryo is the size of a pea and most of the organs have began to take shape (5). Within another four weeks, it is called a fetus, with the features of a human. After twenty-eight weeks of continued growth and development a baby is born. It continues growing beyond puberty into an adult, able to reproduce and start this cycle all over again. But specialized tools are needed to create the state of complexity seen within an adult human being.

The underlying controlling factor is deoxyribonucleic acid, DNA for short, sequences found in the nucleus of both the egg and sperm. These DNA sequences make up the code of life with only four different nucleotides. Each nucleotide has a phosphate group, a deoxyribose sugar molecule, and a nitrogenous base. This is the variable region and can be either a purine, double-ringed type, or a pyrimidine, a single-ringed type (5). Purines are adenine and guanine and the pyrimidines are cytosine and thymine. The DNA molecule itself consists of two strands of nucleotides held together by phosphate bonds. The two strands bind together by hydrogen bonding between corresponding nucleotides, adenine with thymine and guanine with cytosine, in a three-dimensional structure called a double helix. The sequence of bases regulates the formation of proteins and therefore will manage all of the physiological processes that occur throughout growth and development as well as the functions of daily life.
There is a significant amount of DNA contained in both the sperm and egg. Fitting all of this DNA into the nucleus of any cell can be difficult, due to the small size of the cell and the miles of DNA used in coding. This problem is solved by packaging the DNA into chromosomes. First, the DNA winds around proteins called histones. The histones 2A, 2B, 3, and 4 combine to form an octamer which the DNA wraps itself around almost twice. It combines with the H1 histone to form a nucleosome, and the nucleosomes then combine into more complex structures by forming highly ordered supercoils. Each double helix of DNA can be condensed into a visible structure called a chromosome. Regions of the chromosome, which are more tightly coiled, are called heterochromatin and those less tightly wound regions are called euchromatin. It is the euchromatin regions which are transcribed and translated to form the proteins, and these control the processes of the cell and ultimately the human body (5).

In a normal human body cell, there are 46 chromosomes: 23 from the mother and 23 from the father. Twenty-two of these chromosomes from each parent are similar in their genetic contents. The 23rd chromosome in each of these donors is the sex chromosome, an X or a Y. If both parents donate an X then a female is produced, whereas if the father donates a Y then it creates a male. The entire 46 form the genome, and the DNA codes for the proteins in the human body.

Once the egg and sperm combine, the cell begins to rapidly divide, replicating its DNA each time before division takes place. Each division creates two new cells out of one old one, and each of the new cells has a complete set of chromosomes [46]. This leads to development from the zygote to an embryoblast and eventually to a fetus-with a unique chromosomal pattern.
Due to the inheritance of these chromosomes from parents, each cell receives two similar chromosomes. These pair with one-another to make twenty-three pairs of homologous chromosomes. A particular segment of a chromosome is called a gene, and its location at a certain spot on one of the chromosomes is called a locus. On the homologous chromosomes, the pair of corresponding loci contain two different forms of the gene. These different forms are called alleles. The alleles could be the same, homozygous, or different, heterozygous. In the heterozygous mix, one allele maybe dominant and the other recessive. Homo- or heterozygous combinations of alleles make up the genotype of an organism. The observed expression of characteristics is termed the phenotype and is controlled by the genotype.

One way of examining the physical structure of the chromosomes on a macroscopic level is the karyotype, a mapping of chromosomes. On a karyotype, the two homologous non-sex chromosomes would be paired together and then given a common numerical label of one through twenty-two. The X [ˈs] and/or Y are called the twenty-third chromosomes or just by their letters. This mapping of specific chromosomal regions shows similarities between different cells, different human beings, and even within different organisms. The relationship between the physical map and actual genes can be difficult to determine, but a search is on to map the whole human genome, all of the chromosomes, finding the loci of every gene.

The cell cycle, an ordered sequence of events in the life of a dividing cell, is another important part of the growth and development of human beings and cells. This is a cycle composed of four parts:

1) “G1” – a growth phase occurring after mitosis but before replication.
2) "S-phase" – a period when the DNA replicates within a cell.

3) "G2" – a growth phase occurring after replication but before mitosis.

4) "M-phase" – a period when mitosis occurs.

The G1, S, and G2 periods cell cycle are termed interphase, a time during the cell cycle when the cell is not in mitosis. Two important things occurring during the G1 phase are:

1) The cell decides to commit to replication by going past the restriction point. Once past this it cannot go back.

2) The cell synthesizes all of the factors and enzymes needed to replicate.

During the S phase, the cell performs replication. Replication is the basic process of methodically copying the DNA of each of the 46 chromosomes. Therefore, the cell has two copies of the genome for a short period until separation during mitosis. In G2, the cell readies itself for mitosis by making the factors needed to perform mitosis. These processes make up the period of interphase (1).

Mitosis is a process during which the already replicated chromosomes are divided equally in an organized manner, with 46 moving to one end of the cell, while the other 46 move to the opposite end. Once this has occurred, cell division occurs, separating each old cell into two new cells, both with a complete genome. Then the whole process starts over again. Sometimes the cells do not divide and this leads to complications because there are now two nuclei in the cell, each containing a copy of the entire genome (1).
Ploidy is another term which is associated with the number of cell chromosomes (2). The ploidy of a cell is its integer value, “N”, and represents the number of unique chromosomes for that organism. Somatic or body cells are normally diploid [2N], having 2 copies of each unique chromosome, one from the mother and one from the father. Gamete or sex cell is haploid [N]. In the human, the number “N” is 23. A cell is termed polyploid if it has more than two complete sets of chromosomes [3N, 4N]. A cell with either more than or less than the normal pair of homologous chromosomes is called aneuploidy. In this case, the number of chromosomes is not an integer multiple of “N” [2N-1, 2N+1]. Aneuploidy can be caused by several different circumstances, but usually involves incomplete separation of homologous chromosomes or replicated DNA strands during cell division. This happens in cancerous cells quite frequently. Due to the rapid divisions occurring in cancerous cell nuclei, the chromosome number of these cells is not an integer multiple of “N”, thus aneuploidy (2).
Literary Review

Specifically, this study looks at two important genes dealing with cell proliferation. One is the p53 gene dealing specifically with the function of tumor suppression. The other gene examined in this paper is the MDR1 gene. Its function is the interaction of the cell with certain chemotherapeutic agents, and its resistance to them. My literary review explored the p53 tumor suppressor gene. This review covers the current knowledge of the p53 gene and how its expression relates to the topic of ovarian cancer including: its effects within the diseased cells as well as the organism as a whole, and new therapies designed around the p53 gene and its protein products.

The p53 gene codes for a 53-kilodalton protein used in the suppression of cancerous tumor cell growth. The wild-type p53 gene creates a protein product which has the ability to halt cell growth. It has been proven that when significant DNA damage had been done, the wild-type p53 gene’s transcription levels rose as did the amount of p53 protein within the cell. The high levels of this protein are the key in forcing the cell into programmed cell death. Cells with DNA damage and thus high p53 protein levels are pushed into apoptosis. This is a programmed cell death in which the cell literally commits suicide, definitely a benefit to the organism if it is a cancerous cell.

The exact mechanisms of how exactly the p53 gene and its protein affect the cell is still being heavily researched. The protein stimulates the production of another protein, which then inhibits the enzymes involved in the cell cycle. Specifically, it causes other genes to be stimulated and they cause the division of a cell to slow and stop. One of these genes is the Cip1 gene which acts like a “brake on cell division” (7). The protein product created by this gene inhibits the actions of several Cdk proteins. This causes cell division...
to stop completely. The p53 protein also works as a transcription factor acting to stimulate the production of several other proteins. One in particular is the WAF1 gene which possibly controls the tumor suppressive effects of p53. It is present in all normal cells, yet is not found in many of cancer cell lines. All play a role in the inhibition of the cell cycle. This is how the normal cell division is halted.

When a mutated p53 gene is allowed to produce protein, it deprives the cell of the benefits of the apoptotic behavior which the wild-type provides. Also, it spurs on resistance to chemotherapeutic drugs, radiation, and promotes abnormal cell growth. The abnormal growth stimulated by this "oncogene" leads to severe cancerous growth if not detected and treated early. Mutation of p53 is a prominent feature in about fifty percent of all cancers. About ninety percent of the mutations observed are due to missense mutations in the genetic code. A missense mutation involves a base-pair substitution within a gene. It alters a codon, changing a single amino acid in the primary sequence of its gene product. Studies have found missense mutations in about forty percent of the cases of breast cancers, fifty percent of lung cancers, and in many other types of cancer, including that of the ovaries.

According to Curtis C. Harris, "A growing stream of papers has correlated p53 status with patient prognosis" (7). Because of this, a profound amount of research has been done on the p53 gene and its protein product. Studies on different types of carcinogens, which cause missense mutations in the gene, have become the current area of research. By finding ways to prevent the mutations or to reverse their effects, the elimination of several types of cancer is possible, including some ovarian tumors. As Jean Marx, one scientist working in the field, said, "...any therapy that restores lost p53
function would have great power to affect a diverse set of cancers” (9). Yet, prevention is just one means of stopping the growth of cancer. Other methods have hypothesized the idea of simply inserting a good copy of the p53 gene into the cells. This method is to be tried later this year [1998]. A new focus has been on determining how to repair such a mutation. The way to correct the mutational change would be to make a configurational change in the three-dimensional shape of the protein, possibly reactivating the p53 protein. This structural manipulation is, of course, easier said than done. There is also an immunotherapy method, which vaccinates the cell against cancerous growths by alerting the immune system of the host organism to a mutant form of the protein (6). All of these methods offer signs of hope and possibly a cure.
Introduction to Ovarian Cancer

Now let us specifically examine ovarian cancer, its classification by the current system, and perhaps some innovations in treatment that are being used currently to treat this disease.

My research at the Women’s Cancer Institute of Northern California over the Summer of 1997 was geared toward finding a better method of identifying symptoms and predicting prognoses of patients. The current system cannot explain why certain patients who seemed at greater risk upon inspection do better than others whose cancer diagnosis was more favorable. It cannot explain the relapse of cancer in certain patients while others recover without further complications. Why does a stage III grade II person live cancer free after surgery, while a stage II grade I case have recurrent problems for many years to come, or worse pass away despite the minimal affect placed on the organism by the carcinoma? The explanations are mysteries to all studying the problem. These mysteries demands greater research and resolution, thus allowing a better understanding by the doctor and patient, their prognosis, and chances of relapse.

The carcinoma of the ovary is a common type of cancer in the pelvic region. It tends to undergo metastasis, the spreading of cancer cells, and then its curability becomes problematic. This type of cancer is curable if detected in the early stages; however it can grow deadly if allowed to metastasize. Jeffery L. Stern, M.D. is quoted as saying, “Ovarian cancer is the fourth largest killer of women in the United States” (10). There are some frightening statistics - one in 70 contracting ovarian cancer, and one in every hundred women die of this disease.
There are actually five types of histological classification of this carcinoma. Three of these five were used and examined in the research performed this past summer. They are:

1) serous – tissues containing quantities of serum or other watery fluid (8)
2) mucinous – tissues containing quantities of mucus or glycoproteins (8)
3) endometrioid – tissues containing quantities of hormonally affected tissues (8)

There is also clear and undifferentiated types of histology, yet they were not examined in the study. There is also a classification system that currently defines the cancers by a system of stages and grades.

Stages of cancer deal specifically with how far the carcinoma has spread throughout the rest of the organism. Stage I deals with cancer that has afflicted the ovaries. A stage II carcinoma has metastisized to the pelvic region. Stage III has the cancer moving into the abdominal cavity [peritonium] of the body. The shedding of cells by the cancerous ovary tissues leads to a spread of the disease into the peritonium. From here, the disease can get out of control, affecting other organs of the organism and thus leading eventually to a stage IV cancerous growth. This is when the cancer is seen throughout many of the organs. It has become involved within the lymph nodes of the organism and affects the entire body (8).

The grades of cancer also affect the prognosis of the patient. Currently, there are four types of grades defined by those involved in classifying and treating this disease. The cancerous cells are classified by how they look under a microscope, the more differentiated the better. Grade 0 is the most well-differentiated, almost looking like a
normal cell culture. Grades one through three move from well-differentiated [grade 1] to poorly-differentiated [grade 3]. By taking into account the grade, stage, and histology of the carcinoma, a prognosis can be made and a treatment can be formulated.

Currently, there are many different ways ovarian cancer can be treated. First of all, if the ovaries are affected, they are removed. Next, a pathological study of other suspected tissues is performed and surgical removal of the cancerous areas is prescribed. Depending on how far advanced the cancer is, a surgical method of tumor debulking, as it is called, can be done. This entails removal of all cancerous cell tissue as best as possible with a scalpel. Chemotherapy is then prescribed after this to follow-up and ensure that all of the cancerous cells are killed or removed. A “second look” is taken after about the sixth chemotherapy admission as a precaution.

Due to the complications of undergoing surgery, minimally invasive procedures like laparoscopy are now being tested. The administration of chemotherapy with such drugs as:

1) Taxol – a drug which binds microtubules of rapidly dividing cells, inhibiting cell division (8)

2) Cisplatin – a drug which binds guanine bases in rapidly dividing cells, inhibiting replication of DNA (8)

3) Carboplatin – another drug also affecting DNA replication of the rapidly dividing cells (8)

They are a regular part of the treatment for cancer surgery recovery, but these harsh chemicals carry adverse effects to the body in trying to kill all of the cancerous tissue. Radiation is also a method used for those cases in which chemotherapy is not
strong or effective enough. This also has significant side-effects and the benefits must be weighed against the possible harm that these latter methods cause. Other methods are currently being researched, and there are always new drugs earning FDA approval that minimize the side effects and maximize the harm done to cancerous tissues. Perhaps, one day the ultimate cure will be discovered, a miracle drug or procedure that does only good, with no side effects. Until then, however, the options must be weighed and decisions made as to the best possible treatment, one case at a time, one patient at a time.
Introduction to My Study

This study on ovarian cancer examined DNA ploidy as a possible factor in predicting the severity of cancer. Knowledge gained in ploidy studies could help in the search for a better classification system than that of the stage and grade currently used. Genetics and ploidy of this disease are possibly correlated to the deadliness of the disease and also may be a predictor for recovery rates, chances for relapse, etc.

All data gathered for this report, analyzed, and used to indicate some correlations and other non-associations, was obtained mainly from the analysis performed mechanically by a company out of Irvine, CA called Oncotech. The data was obtained by a means termed flow cytometry. This process first takes the solid cancerous tumor mass and disaggregates it with the use of enzymes, physical teasing, or centrifugation. The cells are then placed in a liquid suspension. The liquid suspension is filtered until a sample stream is created. Within the stream, cells are allowed to flow one at a time in a single-file line through the center of the stream. Centering of the cell flow is accomplished by applying an equal pressure on all sides of the stream. Particles in the stream flow through a laser beam at a certain point, and the scattering of light and fluorescence by the particles is measured by receptors. Light scattering occurs due to the striking of light waves on various structures of the cell. These waves hit the membrane of the cell, the nucleus, and other materials within the cytoplasm. Another measurable characteristic is fluorescence. The use of certain dyes and stains, which cause the eminence of certain wavelengths of light, bind to different biochemical structures within the cell. Wavelength and intensity emitted are detected by receptors. All of the information collected by the receptors is then processed by the use of beam splitters, light
filters, and photomultipliers to name a few. Then light signals are converted to electronic information pulses, and these pulses are analyzed by and converted into data bits by a computer database. It stores each point into lists which will later be used to create a graphic representation of the information in the form of a histogram (4).

This method of examining the cancerous tissues lends itself specifically to several important data studies. Specifically, the ploidy count, a DNA index, and the S-phase percentage all are measured by this method of analysis. The ploidy of a cell, mentioned earlier, is computed by measuring the amount of DNA bound by a dye, namely propidium iodide. The use of computer-analysis helps create a histogram with information as to the DNA content in certain fractions. This staining procedure targets for certain genes also, specifically p53 and MDR1. The fractions were of G1%, S%, and G2-M%. The G1% peak is divided by a reference peak and that is then used to calculate the DNA index. From this index number, a reading is made to determine the ploidy. The S-phase percentage is also measured by the second peak of the histogram [S%].
Data and Results

The exploration of data enacted during this research project was originally intended to find a new means of diagnosing, classifying, and updating the prognosis for ovarian cancer patients. Several different studies are performed, looking at how both of the terms in the current classification system, stage and grade, relate to different cancer cell characteristics. Stage and grade will be correlated with certain measured characteristics, such as DNA Index, S-Phase percentage, and ploidy. If a correlation is found, this could help the medical profession gain a stronger grasp on one of the largest yearly killers, especially in America. There is much to be gained from the knowledge and statistical studies done here, even if no correlations are found. The study would be valuable in directing future researchers. In the words of my mentor and supervisor of this project, Dr. Sam Ballon, “Well, if there is no correlation what-so-ever, then it is not a total loss. It confirms my feelings all along that we are paying a lot of money for reports that are really telling us nothing worthwhile” (2). This could quite possibly give the company doing the analysis a slight loss of customers, but it could save money for consumers.

The first relationship examined is how both the stage and the grade compare to the final result, that is, how each of the patients is currently doing. Five catagories are involved in the results:

1) NED is an acronym for “No evidence of disease”.
2) AWD stands for alive with disease.
3) CDP/EXP means continued disease progression or expired, the patient has died.
Attesting to the quality of surgeons at the institute where this data was taken, the NED group was quite large.

Comparing both the stage \([r_s = 0.026]\) and the grade \([r_s = 0.125]\) with the results category, no significant correlation is apparent. This could be due to a lack of data points, for example: the stage I, stage IV, and grade I groupings.

Testing stage and grade with the DNA index became the next step. These graphs showed an increasing trend, though not significant, of the DNA level moving more toward a higher value [stage: \(r_s = 0.005\), grade: \(r_s = 0.027\)]. The cells of higher stages [III and IV] tend to have more DNA in their nuclei. This result is probably produced by the lack of division of the nuclei or the failure of chromosomes to separate during mitosis. Both of these explanations occur in cancerous cells.
The boxplot for the grade did not show as strong results as the stage did when compared with the DNA index. The grade I had a median of 1.0 and are thus all diploid. This makes sense because grade I is the most well-differentiated of all cell types. The other two stages had many more data points, and both medians were about 1.5. This meant that their DNA index was higher than that of a diploid cell [1.0], therefore the cells of grades II and III were more poorly differentiated. The data showed an increase according to predictions for the general trends. The higher the stage or grade, the larger the DNA index value, or at least the less likely the cells are to have a score of 1.0 signifying a diploid cell.

Another interesting correlation is how the S-phase values corresponded to the results of the patients, while coding the data points with either stage \([r_s = 0.345]\) or grade \([r_s = 0.330]\). Both the stage and the grade show a correlation with the S-phase values at a 0.05 level of significance.
The percentage of S-phase within the cells was mainly between ten to twenty percent, and the data never exceeded thirty percent. There is a significant correlation of 0.05 between the S-phase and the DNA level \([R = 0.257, r_s = 0.263]\). Note from Figure 3 that as the DNA level rises, so does the S-phase percentage. This relationship was explored as it related to stage, grade, results, and histology categorizations.

![Figure 3. Scatter Plot of S-Phase vs. DNA Index](image)

The stage I cells have too few values and are not spread out enough on the scatter plot in Figure 3, thus no line of best fit is drawn. This could have been predicted because it is the least invasive of all four stages of ovarian cancer, therefore harder to detect and less data available. However, examining the stage II and stage III data was more interesting. Both sets of data seemed to express a wide range of values according to both the S-phase and DNA index axises. The line of best fit for stage II data has the same slope as the line for the total population, see Figure 3. This shows that there is a positive correlation between the stage II and the data for S-phase and DNA index; as one increases so does the other. The stage III data also the same slope as the line of best fit for
the total population. The stage III line is shifted up from the stage II line as seen in Figure 3. This could be indicating that there is a directly proportional relation between S-phase and DNA index when categorized by stage. It also indicates that as the stage value get larger, a shift in the S-phase intercept value occurs. This shift of values explains why stage III is metastasized throughout a greater area. The larger the number of cells in S-phase, the more that are undergoing replication. If the cells are cancerous and are rapidly dividing, it is probable that their DNA index will be in the aneuploid range, a value on the DNA axis greater or less than 1. Stage IV had few data points also, but still had enough values to get a line of best fit. This line has a higher slope than the other two, stage II and III. This different slope still followed what seems to be the general trend, as S-phase increases so does the DNA index when examining the category of stage.

![Figure 4. Scatter Plot of S-Phase vs. DNA Index](image)

Examining the grade as the category label for the scatter plot of S-phase percentage versus DNA index, the same general pattern occurs as with the stage values. The values of the data on the S-phase percentage axis were mostly in the 10 to 20 percent
range with a few values located in the thirty percent range. For grade I, the scatter plot in Figure 4 shows the upward sloping line showing a direct proportion between S-phase and DNA index, as seen with stage in Figure 3. This trend is the same for the values of grades II and III also. None of the lines of best fit have the same slope, but the two with a larger number of data points are closer to the line for total population. Neither is as close as stage III of Figure 3, but both appear to be getting closer to the slope of the total population.

![Figure 5. Scatter Plot of S-phase vs. DNA Index](image)

Upon seeing trends in the earlier studies, shown in Figures 3 and 4, a scatter plot of the S-phase versus DNA index with the category of results was examined. The extreme results categories, NED and CDP/EXP, have slopes similar to the line of best fit for the total population. This is interesting because the trend is saying that those with ovarian cancer can result in either a complete recovery or can perhaps eventually die as S-phase and DNA index both increase. This does not make a lot of sense, because how
can a cancerous cell that has the same S-phase and DNA index value be equally likely to
metastasize or to stop dividing and die. There is also the AWD category in results. The
straight line shown on Figure 5 means there is no correlation between the S-phase and
DNA index plot when categorized with the AWD results. The study of this scatter plot
needs further research done on it, possibly with more data gathered on this topic.

A last plot looked at using the scattering of data in the S-phase versus DNA index
is that categorized with histology \( r_s = -0.123 \). The histology studies were obtained by
examining the pathology reports of the treating hospitals. These reports focused on three
of the distinct types of cell character: serous, mucinous, endometriod, or a combination of
these three. Data gathered in this area was then examined with the patient results and
various other categories.

The scatter plot shown in Figure 6 shows a few interesting trends. It also has the
direct proportionality seen between the values for S-phase and DNA index when
categorizing it with histology. A couple of categories did not show a line of best fit, while
those that did all had varying slopes. Only the serous type of cells showed a slope that almost matched the line of best fit for the total population. The other two lines for endometrioid and a grouping of serous and mucinous cell types show an upward trend with a positive slope, but the slope is not near the fit line for the total population. The reasoning for this could be due to small sample sizes in a couple of the categories, or a few outliers skewing the data a bit.

These studies were the primary focus of my summer research project. Hardly any new information comes from these data studies. Yet it does confirm previous studies indicating that as stage increases, then the patient’s prognosis lessens, and that as S-phase and DNA index values increase, the stage and grade values seem to also increase.

Another correlation needing further examination is that of the histology of the cancerous tissues and how they relate to patient results. The data points are most abundant in the “no evidence of disease” column. This could reflect on the doctors’ ability more than on the histology of the cells. The other types of cells, both the mucinous and endometrioid, also are diversely spread on a scatter plot. Due to this scattering of the data, there does not appear to be any apparent correlation between the histology and S-phase percentage \( r_s = -0.101 \) or the patient results \( r_s = 0.232 \). Despite the lack of correlation, this does not suggest discontinuing the histology studies. Perhaps, working with a larger or more diverse data set would demonstrate the correlation that was anticipated.

Another study performed on the data set is that of the MDR1 gene and the effects of the chemotherapeutic drug Cisplatin, and the patients’ results. When looking at the histoscore assigned the MDR1 percent data, there is no correlation shown by the scatter
plot [Figure 7] of the data \( r_0 = -0.094 \). Some sort of relation possibly starts to form, but
there are not enough data points to make an accurate determination of how the gene’s
overexpression can cause a resistance to chemotherapeutic agents. This would be worth
further investigation with the use of more data points if at all possible.

Figure 7. Scatter Plot of MDR1\% vs. Cis\%
Conclusion

The study of ovarian cancer, with all of its interesting qualities, provides an excellent research subject. The potential good done in furthering knowledge of this affliction benefits society as a whole. The search for a correlation between different measurable aspects of ovarian cancerous tissue and the resulting patient prognoses proved to be one aspect that showed some promise. In examining the category of stage as it relates to a graph of the S-phase versus the DNA index readings, a possibly significant correlation can be seen. The stage II and III lines of best fit seem to have slope similar to the line of best fit of the total population. As the S-phase increases, so does the DNA index values. The S-phase is a measure of the amount of the cells which are undergoing replication and the DNA index measures the amount of DNA in the cell in comparison to a normal diploid cell [1.0]. This relationship makes sense in cancer cells because as the replication rate increases, so does the chance for errors and thus cells with more or less DNA are aneuploid. Future studies should include a larger data set, random selection of data, and possibly elimination of outliers when computing statistical correlations.

Other findings of this study include the grade and stage versus the DNA index and the histology versus final diagnosis. The stage and grade had an observed correlation between their values and the level of DNA index at low values. This could be due to small values of the stage I and grade I categories. The histology versus results showed a large number of serous cells being correlated to a positive prognosis. There were only a few of the other categories and this made studying the data a lot more difficult.

To conclude, a recommendation should be made to continue research in these areas. The benefits outweigh the costs, for life is the most precious resource. A few
studies should be repeated to find a more significant correlation of the data. More data would help in the studying of this disease. Also, randomly sampling the data and having more efficient ways of identifying the information in patients’ charts would be an improvement. Other information to be studied includes:

1) A correlation between the lymph node infectious percentage versus prognosis
2) Age versus prognosis or stage and/or grade
3) Racial background versus stage and/or grade
4) Economic status versus recovery rate or prognosis

These are just a few examples where future research can go to investigate the killer known as ovarian cancer. With the rise in knowledge of this disease and the human genome, technology may hold the key to unlocking today’s mysteries.
Works Cited


10) http://wccenter.com/ovary.html