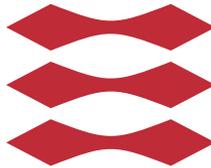


# Quantitative tumor heterogeneity assessment on a nuclear population basis

Anne-Sofie Wessel Lindberg  
s082831

DTU



In cooperation with

visiopharm  
TURNING IMAGES INTO KNOWLEDGE

2014

Technical University of Denmark  
Department of Applied Mathematics and Computer Science  
Matematiktorvet, building 303B,  
2800 Kongens Lyngby, Denmark  
Phone +45 4525 3351  
[compute@compute.dtu.dk](mailto:compute@compute.dtu.dk)  
[www.compute.dtu.dk](http://www.compute.dtu.dk)  
IMM-M.Sc-s082831

# Abstract

---

20 % of all treatments of breast cancer aren't successful due to the heterogeneity in the cell population. The heterogeneity can be measured as the response from various biomarkers which also is used in the determination of the type of treatment.

226 TMA cores stained with either ER or Ki67 were aligned with their neighbor slice stained with PCK and the tumor cells with positive and negative response to the biomarker were segmented using the software VIS. An automatic method for a better visualization of the heterogeneity in the cell population response from the two biomarkers Ki67 and ER was developed in form of a heatmap that illustrated the percentage of positive cells in small areas.

The visualization was further used as a guidance to find the largest area with highest response (hottest hotspots) and an automatic calculation of the heterogeneity score in this area was performed. 110 of the 226 TMA cores were scored by a pathologist. The automatic calculated scores for the Ki67 TMA cores were compared with the pathologist scores. The scores calculated in the hottest hotspot were not significantly different from the pathologist scores but were in general 7 % higher than the pathologist scores. The automatic calculated scores for the ER TMA cores were also compared with the pathologist scores. The scores were not found significantly different from the the pathologist scores but were in general 1.3 % lower.

The impact on calculating the scores in hotspots was investigated and it was found that scores from the first, second, third and fourth hottest hotspot weren't significantly different from the pathologist scores but scores calculated randomly outside the hottest hotspot were different from the pathologist scores.

**Keywords:** *breast cancer, digital image analysis, Ki67, ER, cancer heterogeneity, heterogeneity score, heatmap, hotspot detection*



# Resume

---

20 % af alle behandlinger af brystkræft er ikke succesfulde grundet heterogenitet i tumorcellernes population. Heterogeniteten kan måles ved at måle responset fra en eller flere biomarkører. Biomarkørerne anvendes også til at bestemme hvilken behandling der skal gives.

226 TMA cores farvet enten med biomarkøren ER eller Ki67 blev alignet med deres naboslice farvet med PCK. Herefter blev tumorceller, der reagerede positivt og negativt på biomarkøreren, segmenteret ved brug af softwaren VIS. En automatisk metode blev udviklet til at visualisere heterogeniteten i cellepopulationen, et såkaldt heatmap, for både de Ki67 farvede og ER farvede billeder. Heatmapet viste den procentvise andel af positive celler i forskellige områder.

Heatmapet blev yderligere anvendt til at finde det største område med højeste procentvise andel af positive celler, et såkaldt hotspot. En automatisk beregning af heterogenitetsscoren i det fundne hotspot blev udviklet. 110 af de 226 anvendte TMA cores var blevet scoret af en patolog. De udregnede scores for de Ki67 farvede billeder blev sammenlignet med scorene fra patologen. Der blev ikke påvist en signifikant forskel mellem de udregnede scores og patologscorene, men de udregnede scores var generelt set 7 % højere end patologscorene. Det samme blev gjort for de udregnede scores for de ER farvede billeder og her blev der heller ikke påvist en signifikant forskel mellem de udregnede scores og patologscorene. De udregnede ER scores var generelt set 1.3 % lavere end patologscorene.

Betydningen af at regne scorene i et hotspot blev også undersøgt og det blev fundet at scorene regnet i det største hotspot, det næststørste, det tredje største og det fjerde største ikke kunne påvises at være signifikant forskellige fra patolog scorene mens scores regnet tilfældigt uden for det største hotspot var signifikant forskellig fra patolog scorene.

**Nøgleord:** *brystkræft, digital billedanalyse, Ki67, ER, cancer heterogenitet, heterogenitetsscore, heatmap, detektion af hotspot*

# Preface

---

This project was a 32.5 ECTS master thesis carried out at DTU Compute at the Technical University of Denmark in collaboration with the company Visiopharm. The project is the completion of the education Medicine and Technology at the Technical University of Denmark. The project was carried out in the period October 2013 to March 2014.

The aim of the project was to visualize heterogeneity in breast cancer and calculate a quantitative measure based on the cell population in histology images stained with two different biomarkers. The used data set consists of a set of around 240 TMA cores stained with the biomarkers Ki67 or ER and some of them scored by a pathologist together with around 240 neighbor slice stained with PCK.

The output of the project is a set of heatmaps visualizing the heterogeneity and used for guidance to find so-called hotspot where a quantitative measure of biomarker response was calculated based on the nuclear population basis. The scores were compared with the scores provided by the pathologist.

Anne-Sofie W. Lindberg

Anne-Sofie Wessel Lindberg  
Lyngby, 31 March-2014



# Acknowledgements

---

I would like to thank the following people for their assistance and support throughout this project:

A special thanks to my supervisors Professor Rasmus Larsen and Professor Knut Conradsen for their support, supervision on the weekly meetings and for always having their doors open.

I would also like to give a special thanks to my supervisor Michael Lippert from Visiopharm for providing the data set, the project idea and for feedback throughout the project.

And a thanks to associate Professor Allan Aasbjerg Nielsen for help with the semivariogram analysis.



# Contents

---

<b>Abstract</b>	<b>i</b>
<b>Resume</b>	<b>iii</b>
<b>Preface</b>	<b>v</b>
<b>Acknowledgements</b>	<b>vii</b>
<b>1 Introduction</b>	<b>1</b>
<b>2 Theory</b>	<b>3</b>
2.1 Breast cancer . . . . .	3
2.1.1 Types of breast cancer . . . . .	3
2.2 Immunohistochemistry . . . . .	5
2.2.1 Preparation of tissue . . . . .	6
2.2.2 PCK stain . . . . .	6
2.2.3 Ki67 stain . . . . .	6
2.2.4 ER stain . . . . .	7
2.3 Cancer heterogeneity . . . . .	9
2.4 Heterogeneity scores . . . . .	10
2.4.1 H-scores . . . . .	10
2.4.2 Ki67 scores . . . . .	10
2.5 Semivariogram analysis . . . . .	11
<b>3 Data</b>	<b>15</b>
3.1 TMA cores and their structure . . . . .	15
3.2 Removal of bad TMA cores . . . . .	16
3.3 Scores on data . . . . .	19
<b>4 State of the art</b>	<b>21</b>

<b>5</b>	<b>Methods</b>	<b>25</b>
5.1	Image processing in VIS . . . . .	25
5.1.1	Alignment using virtual double staining . . . . .	25
5.1.2	Segmentation of cells . . . . .	27
5.1.3	Examples of the segmentation of cells . . . . .	28
5.2	Outline of TMA core . . . . .	29
5.3	Tumor masks . . . . .	31
5.3.1	Amount of tumor tissue inside TMA core . . . . .	33
5.4	Blob detection . . . . .	33
5.5	Heatmap calculation . . . . .	36
5.5.1	Semivariogram analysis . . . . .	36
5.5.2	Heatmap . . . . .	38
5.5.3	Sliding window function . . . . .	38
5.5.4	The Gauss filter used . . . . .	40
5.5.5	Heatmap parameters . . . . .	43
5.6	Hotspot detection . . . . .	45
5.6.1	Detecting largest hotspot . . . . .	45
5.6.2	Field Of View (FOV) . . . . .	48
5.6.3	Calculation of heterogeneity score at hottest hotspot . . . . .	50
<b>6</b>	<b>Results</b>	<b>51</b>
6.1	Segmented cells and their semivariogram . . . . .	51
6.2	Amount of tumor tissue inside TMA core . . . . .	54
6.3	Heatmaps . . . . .	55
6.3.1	Ki67 Heatmaps . . . . .	55
6.3.2	ER Heatmaps . . . . .	60
6.4	Hotspot detection . . . . .	64
6.4.1	Ki67 Heatmaps with fixed size FOV and adjusting FOV . . . . .	64
6.5	Mean, variance and histogram for heatmaps . . . . .	66
6.6	Ki67 scores from hottest hotspot . . . . .	68
6.6.1	Bland Altman plot . . . . .	69
6.6.2	Linear regression models . . . . .	70
6.7	Ki67 scores outside the hottest hotspot . . . . .	76
6.7.1	Linear regression models . . . . .	76
6.8	Cut off at 14 % for Ki67 scores . . . . .	83
6.9	Ki67 scores in the hottest, second, third and fourth hottest hotspot . . . . .	84
6.10	ER scores from hottest hotspot . . . . .	86
6.10.1	Model $y = \beta x$ . . . . .	86
<b>7</b>	<b>Discussion</b>	<b>91</b>
7.1	Amount of tumor tissue . . . . .	91
7.2	Heatmaps and the spatial correlation . . . . .	91
7.3	Hotspot detection and FOV . . . . .	92
7.4	Quantative measured scores . . . . .	92

---

7.4.1	Increasing variance for increasing Ki67 scores . . . . .	92
7.4.2	Classification of patients based on Ki67 scores above and below cut off value . . . . .	93
7.4.3	Evaluation of the score results and the fitted regression models . . . . .	94
<b>8</b>	<b>Conclusion</b>	<b>95</b>
<b>A</b>	<b>Semivariogram function by Allan Aabjerg Nielsen</b>	<b>107</b>
<b>B</b>	<b>Amount of tumor tissue in ER and Ki67 stained images sepa- rately</b>	<b>109</b>
<b>C</b>	<b>Regression model for Ki67 scores from the hottest hotspot</b>	<b>111</b>
<b>D</b>	<b>Log model for Ki67 scores from the hottest hotspot</b>	<b>115</b>
<b>E</b>	<b>Regression model for Ki67 scores outside the hottest hotspot</b>	<b>119</b>
<b>F</b>	<b>Log model for Ki67 scores outside the hottest hotspot</b>	<b>123</b>
<b>G</b>	<b>Regression model for the ER scores</b>	<b>127</b>
<b>H</b>	<b>Regression models for Ki67 scores in the second, third and fourth hottest hotspot</b>	<b>129</b>
	<b>Bibliography</b>	<b>137</b>



# Introduction

---

In 2011 over 508.000 died from breast cancer worldwide. Breast cancer is the most prevalent type of cancer in women world wide[1]. One of the main problems in diagnosing and treating breast cancer is that the population of breast cancer cells is extremely heterogeneous.

There are two main ways to measure cancer heterogeneity; The first one deals with classification of different cell types and measurement of parameters within the same cell type using Haemotoxylin and Eosin (H&E) stained images. The second is a description of the cells' response to various biomarkers and is the one investigated in this project.

Biomarkers are used to evaluate specific features for the individual cancer e.g. aggressiveness and presence of estrogen receptor as those used in this project. The responses to the different biomarkers are used to plan the individual treatment and if it should be just one treatment or a combination of treatments. Due to the heterogeneity, cells will respond different to the biomarker and the distribution of responses will not be homogeneously distributed within a tissue sample.

To make sure patients get the right treatment the response to the biomarker needs to be calculated at a location presenting the highest response to the biomarker and due to the extreme heterogeneity this might be hard to identify.

The aim of this project was to develop a visualization of the heterogeneity in tumor tissue stained with two different biomarkers and furthermore to develop

a method for measuring the heterogeneity on a more objective scale and compare the scores with subjective scores measured by a pathologist. The impact of measuring in areas with high heterogeneity is discussed and an analysis is performed to clarify the use of these.

## 2.1 Breast cancer

Breast cancer is the most prevalent type of cancer for women worldwide. In 2011 over 508.000 died from breast cancer. Even though the number seems extremely high breast cancer is also one of the cancer types with highest survival rates reaching from 40 % in low-income countries to over 80 % in North America, Sweden and Japan [2].

In Denmark the survival rate is around 50 % and there is around 3600 new cases per year. Breast cancer is characterized by being an extremely heterogeneous type of cancer both regarding proliferation, metastasis and cell distribution [3].

This section contains a more medical explanation of two of the most common breast cancer types and an illustration of the cells involved.

### 2.1.1 Types of breast cancer

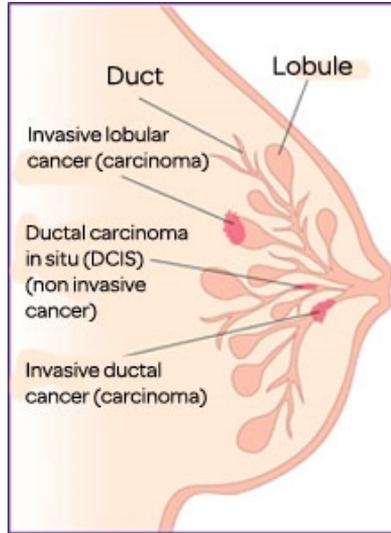
There are two main types of breast cancer, a ductal and a lobular type. Around 80 % of all breast cancers are ductal and around 10 % are lobular. The remaining 10 % are other special types such as tubular, papillary and metaplastic and will not be further described in this report [3].

The two main types are described further in the next sections.

#### **2.1.1.1 Ductal carcinoma**

The ductal carcinoma originates in the milk ducts and can be both non-invasive (DCIS - Ductal Carcinoma In Situ) or invasive (IDC - Invasive Ductal Carcinoma). Often the non-invasive is just a pre-step and develops into the invasive type when the cancer cells break through the walls of the ducts and grows into the surrounding tissue, primarily fat tissue. Both DCIS and IDC are illustrated on figure 2.1. Depending on the type and metastasis status the treatment of ductal carcinoma is a combination of the following [4]:

- Lumpectomy - breast preservation operation
- Mastectomy - surgical removal of partial or whole breast
- Axillary lymph node dissection - removal of lymph nodes under the arm due to metastasis
- Radiation
- Chemotherapy
- Hormonal therapy
- Biologic targeted therapy



**Figure 2.1:** Illustration of originate sites for Ductal Carcinoma In Situ (DCIS), Invasive Ductal Carcinoma (IDC) and Invasive Lobular Carcinoma (ILC). Image originates from [5].

### 2.1.1.2 Lobular carcinoma

The lobular carcinoma originates in the milk producing lobules and can both be non-invasive and invasive as well. Around 30 % of the non-invasive develop into invasive cancer. The invasive type (ILC - invasive lobular carcinoma) is illustrated on figure 2.1 where the anatomy of the breast also can be seen. This type of breast cancer affects primarily women above 55 years old.

## 2.2 Immunohistochemistry

Immunohistochemistry (IHC) is as the name refers a combination of immunology, histology and chemistry. The immunology used is the knowledge about antibodies binding to antigens which can be used in histological images. Histological images are tissue samples that have been fixed, sliced, stained and placed on a glass slide. To find specific cells containing specific antigens, the tissue sample is stained with a mixture containing the antibodies for these antigens.

The binding between the antigen and antibody often gives rise to a colored

histochemical reaction which makes it possible to see the cells that contain the antigen[6].

### 2.2.1 Preparation of tissue

The method used for making stained slices is described shortly in this subsection. A further description of the Tissue Micro Array (TMA) cores can be found in section 3.1.

First the tissue is cut out and afterwards fixed using paraformaldehyde. A hollow needle is used to take out a cylinder shaped sample of the tissue, which is embedded in a block of paraffin. Afterwards the paraffin block is cut into 4  $\mu\text{m}$  thick slices. Each slice is then stained with different kinds of stains. The used stains in this study are the biomarkers PCK, Ki67 and ER which are further described in the following three sections.

### 2.2.2 PCK stain

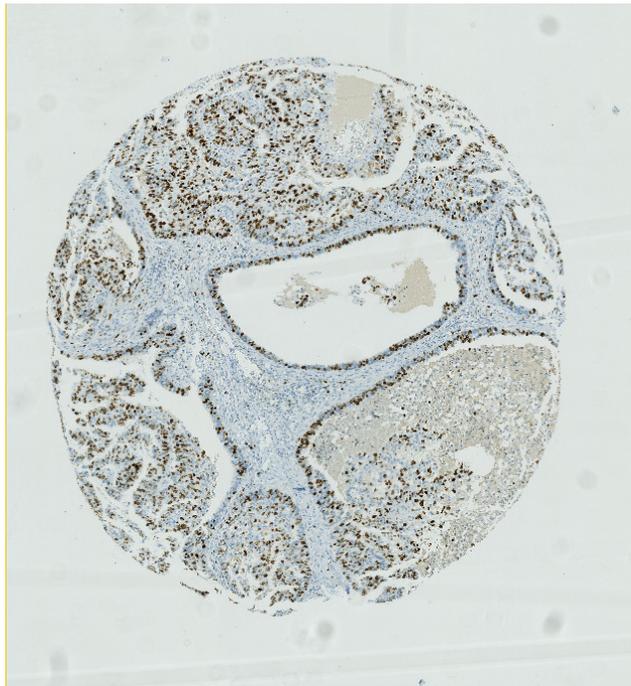
PCK stands for Pan Cytokeratin and is a mixture of the two antibodies AE1 and AE3. Cytokeratins are keratin-containing filaments that are found in the cytoskeleton of epithelial cells. Since almost all cases of breast cancer are adenocarcinomas and therefore originates from epithelial cells it is an important tool to visualize epithelial cells in histology images. It is known that the intermediate filament protein expression is the same in tumor cells as the cells they originate from[7] and therefore the stain can be used to mark regions with tumor cells.

There can be more than one type of cytokeratins in cells and the used stain is therefore a mixture of two clones of anti-cytokeratin monoclonal antibodies called AE1 and AE3. AE1 detects cytokeratin 10, 14, 15, 16 and 19 and AE3 detects cytokeratin 1, 2, 3, 4, 5, 6, 7 and 8. The reason why the stain is called Pan Cytokeratin is that it covers "all" cytokeratins. There are a few cytokeratins that it doesn't cover e.g. cytokeratin 17 and 18 that are present in hepatoma and therefore it can't be used for investigation of liver cancer [8], [9], [3].

### 2.2.3 Ki67 stain

Ki67 is a protein that only is present during the active phases of cell division. The biomarker can therefore be used to mark cell proliferation and thereby

indicate how aggressive the cancer is. Aggressive cancer is cancer cells dividing rapidly and creating large or many tumors, that may spread during a short time period. The cells where Ki67 is active are colored brown while the cells that are negative for Ki67 are colored blue. An example of a Ki67 stained image can be seen on figure 2.2.



**Figure 2.2:** An example of a histology image stained with the biomarker Ki67. The brown cells are cells that reacts positive to the stain and therefore is Ki67 positive. The blue cells are Ki67 negative cells. The histology image is a TMA core.

An aggressive cancer has a poor prognosis but is easier to treat with chemotherapy since chemotherapy destroys rapid dividing cells [10], [11], [12].

### 2.2.4 ER stain

ER stands for Estrogen Receptor and is a staining type that as the name refers stains cells that contain estrogen receptors. An estrogen receptor is localized

inside the cell and it binds estrogen with high affinity<sup>1</sup>.

Positive stained cells are colored brown and negative cells are colored blue. The intensity of the brown color indicates how many estrogen receptors the cell contain and this has an impact on the score named the H-score which is described in section 2.4.1.

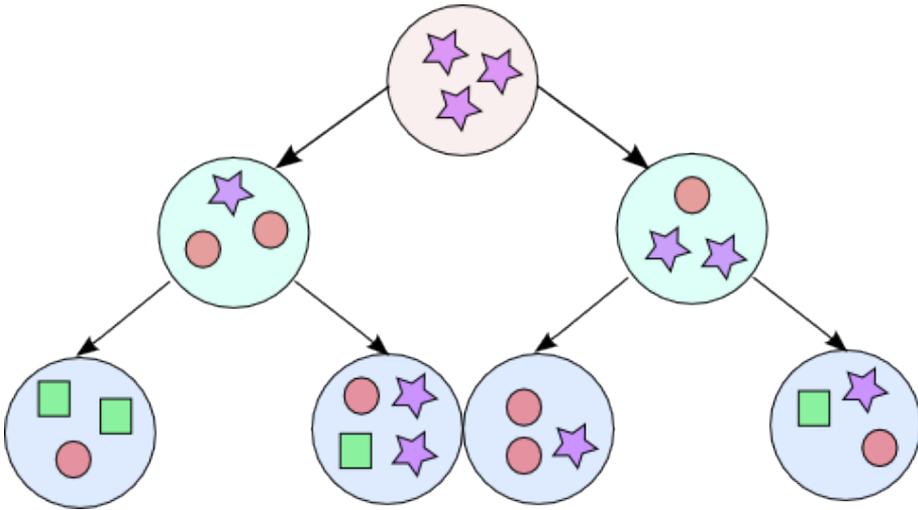
Around 75 % of all breast cancer carcinomas have estrogen receptors [13]. Cancers with estrogen receptors respond well to endocrine treatment and therefore have a better prognosis than estrogen negative carcinomas.

---

<sup>1</sup>Affinity in chemistry means the tendency of two compounds to combine to each other.

## 2.3 Cancer heterogeneity

A tumor consist of cancer cells which all originates from the same mother cell. A normal cell division consists of a mother cell that divides into two similar daughter cells. In cancer the daughter cells can deviate a bit from the mother cell due to genetic changes. An illustration of this can be seen on figure 2.3, where circles represent cells and the symbols represent genes inside the cells.



**Figure 2.3:** An illustration of cell division for cancer cells. The circles represent cells and the symbols represent different kind of genes. For some divisions the cancer cells' genes change slightly which in the end causes different types of cancer cells resulting in cancer heterogeneity.

As it can be seen from the figure the gene set (illustrated by symbols) changes a bit from each cell division making the tumor cells slightly different from each other. This happens in some cell divisions but not in all cell divisions. Cancer with high heterogeneity is more difficult to treat since it contains many different cells, that will not respond in the same way to treatment. Therefore cancers with high heterogeneity need to be treated with a mixture of treatments[14].

The heterogeneity can be quantified in different ways and in this project it is described by looking at the response to two different biomarkers.

## 2.4 Heterogeneity scores

To quantify the response to a biomarker a score is calculated. This section defines the H-score and the ER score used for ER stained images and the Ki67 score used for Ki67 stained images.

### 2.4.1 H-scores

ER stained images are scored with an H-score given as

$$\begin{aligned}
 H_{score} = & (\% \text{ of cells stained at intensity category } 1 \times 1) \\
 & + (\% \text{ of cells stained at intensity category } 2 \times 2) \\
 & + (\% \text{ of cells stained at intensity category } 3 \times 3)
 \end{aligned} \tag{2.1}$$

where category 1 is weak brown stain intensity, category 2 is intermediate brown stain intensity and category 3 is strong brown stain intensity.

The H-score lies in the interval 0 to 300, where 300 equals to 100 % of tumor cells stained strongly[15].

The ER stained images can also be evaluated using the percentage of positive cells as a score. This is named the ER score and is simply done by summing the % in each category.

### 2.4.2 Ki67 scores

The Ki67 score is the number of positive cells stained by the Ki67 biomarker calculated in a chosen area. The Ki67 score is given as

$$\text{Ki67}_{score} = \frac{\text{Number of positive cells}}{(\text{Number of negative cells} + \text{Number of positive cells})} \times 100 \tag{2.2}$$

The Ki67 score lies in the interval 0 to 100 %, where 100 % equals to all cells stained positive by the Ki67 biomarker.

## 2.5 Semivariogram analysis

Semivariogram analysis can be used to find the spatial correlation. The semivariogram measures the variability of a variable in a 2D or 3D space. In this case it is in the 2D space and the spatial variability is given as the variogram which is a function of distance and direction:

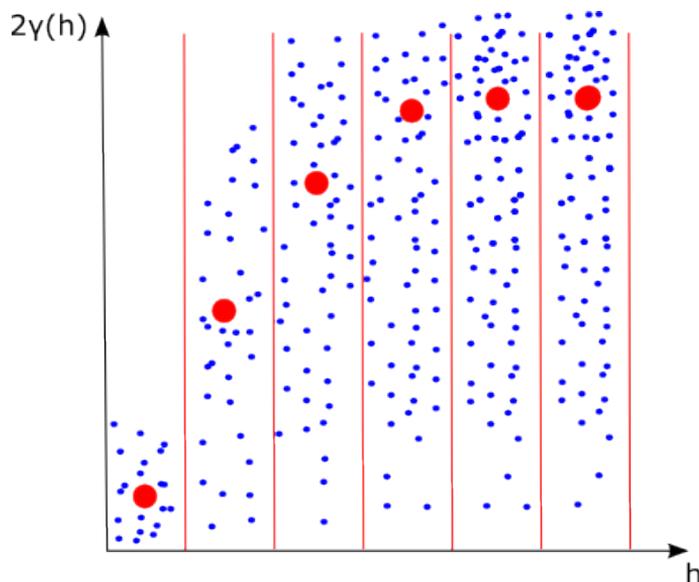
$$2\gamma(r, h) = E\{[Z(r) - Z(r + h)]^2\} \quad (2.3)$$

where  $r$  is the location in space and  $h$  is the displacement vector between two observations.

The semivariogram  $\gamma$  is the half of the variogram.

To calculate a graph of the semivariogram, the semivariogram value for all point pairs is divided into bins and the mean value of each bin is used as a data point for the semivariogram graph. The value from the first bin is divided by 2 to minimize the nugget effect which is described in the following.

An illustration of this can be seen on figure 2.4 where the calculated variogram value between point pairs is illustrated by blue points, the bins with red lines and the mean in each bin as a red dot.

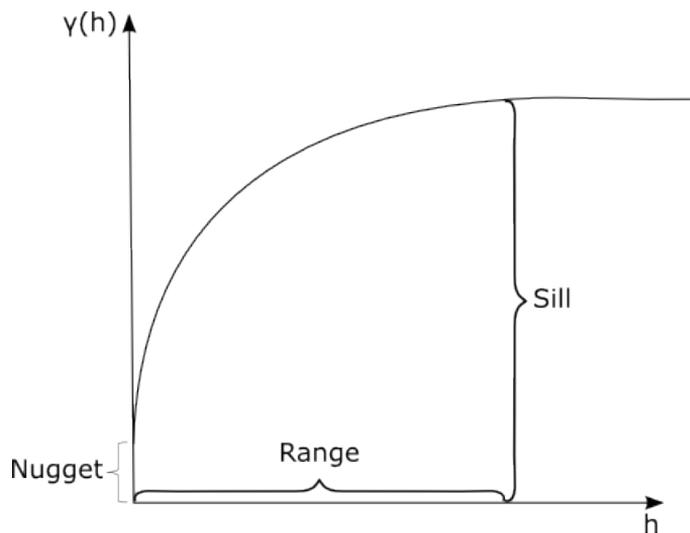


**Figure 2.4:** An illustration of the construction of a semivariogram. The variogram value as a function of the size of the displacement vector  $h$  between points is shown. The bins are marked by red vertical lines and the red dot in each bin is the mean value. The mean value in the first bin is divided by 2 to minimize the nugget effect. To transform the values into a semivariogram the mean values are divided by 2.

The generated graph follows a spherical model which flattens at some value. Therefore a spherical model can be fitted to the semivariogram and the fitted model has three characteristics:

- **Range of influence** - The distance between points where the correlation between them ends. Characterized on the model as the value on the x-axis where the graph flattens.
- **Sill** - The sample variance characterized as the semivariogram value (on the y-axis) when the correlation stops.
- **Nugget** - Theoretically, the semivariogram value should be 0 at lag 0, but due to measurement errors or spatial variation at distances smaller than the sampling interval the value often is larger than 0 and this is called the nugget effect. The value is characterized on the graph as the intercept of the model.

The three characteristics is illustrated on a spherical model on figure 2.5.



**Figure 2.5:** Range, sill and nugget on a spherical model.

The spherical model is given as

$$\gamma(h) = \begin{cases} 0 & |h| = 0 \\ C_0 + C_1 \left[ \frac{3}{2} \frac{|h|}{R} - \frac{1}{2} \frac{|h|^3}{R^3} \right] & 0 < |h| \leq R \\ C_0 + C_1 & R < |h| \end{cases} \quad (2.4)$$

where  $R$  is the range,  $C_0$  is the nugget effect and  $C_0 + C_1$  is the sill.



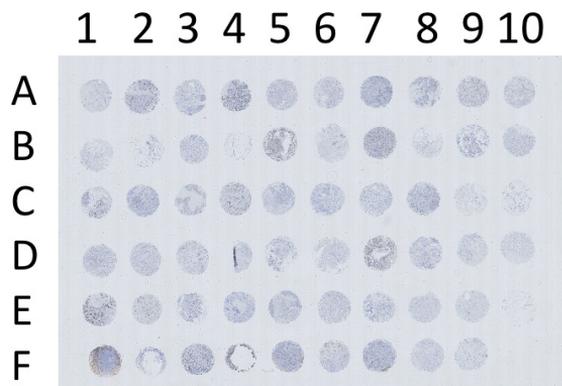
This chapter contains a description of the used data. Furthermore the data quality is evaluated and a short description of the scores provided by a pathologist is given.

## 3.1 TMA cores and their structure

Each histology image used is an image taken of a TMA core. TMA stands for Tissue Micro Array and is a 4  $\mu\text{m}$  thin stained slice of tissue from a cylinder block made by plugging a needle down in a paraffin fixed block of tissue. Each TMA core is placed on a glass slide that contains up to 60 TMA cores in total. This slide is scanned with an optical scanner that takes high resolution digital images of all TMA cores.

One slide contains up to 60 TMA cores either stained with ER or Ki67. The neighbor slice (in the cylinder block) to each TMA core is stained with PCK and placed at another slide at the same position in the TMA core grid. The TMA core grid consists of 6 rows and 10 columns, where the rows are named A-E and the columns 1-10, so each TMA core is named with these coordinates e.g. A4, B7 etc.

Figure 3.1 shows a virtual slide where the TMA cores are stained with Ki67. The coordinate system is also shown on the figure.



**Figure 3.1:** An example of a virtual slide and the coordinate system. It contains 59 TMA cores stained with Ki67. Core E10 is missing.

In this project, 2 slides with a total number of 111 TMA cores stained with Ki67 and the corresponding 2 slides stained with PCK are used for the analysis of Ki67. 64 of the TMA cores have been scored by a pathologist.

For the analysis of ER, 2 slides with a total number of 115 TMA cores stained with ER and the corresponding 2 slides stained with PCK are used. 55 of the TMA cores have been scored by a pathologist.

Some TMA cores are removed from the data set which is further described in the next section.

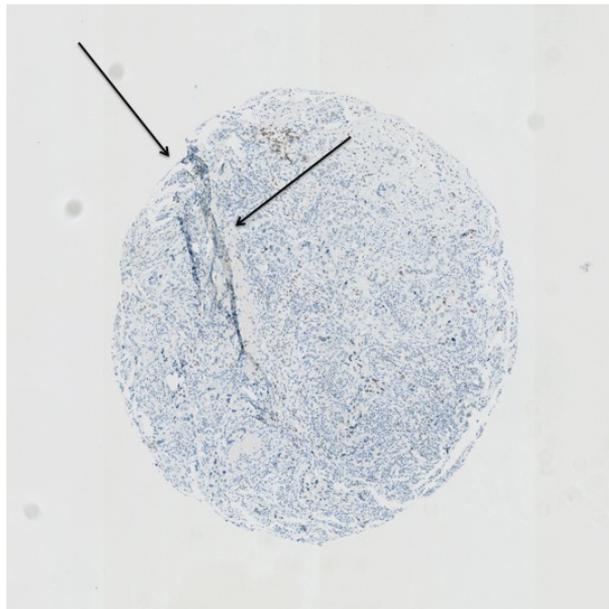
## 3.2 Removal of bad TMA cores

The quality of each TMA core is checked since the alignment (further described in section 5.1.1) between a Ki67 or ER stained core and the corresponding PCK stained core doesn't make sense if the two cores aren't similar due to low quality of the core.

The quality of the core pair is significantly reduced if one of the cores is folded, torn apart or stretched under the production process.

Figure 3.2 shows an ER stained core that is folded. Cores that are folded, torn or stretched are excluded from the data set.

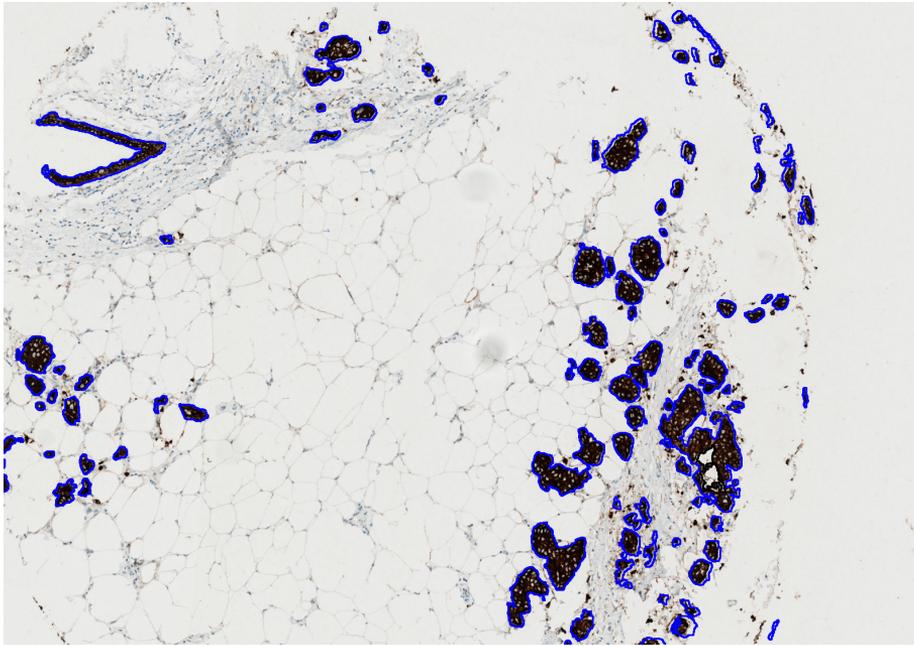
ER or Ki67 cores and their neighbor slice stained with PCK that aren't similar enough are also removed. Bad alignments between those two are excluded as well.



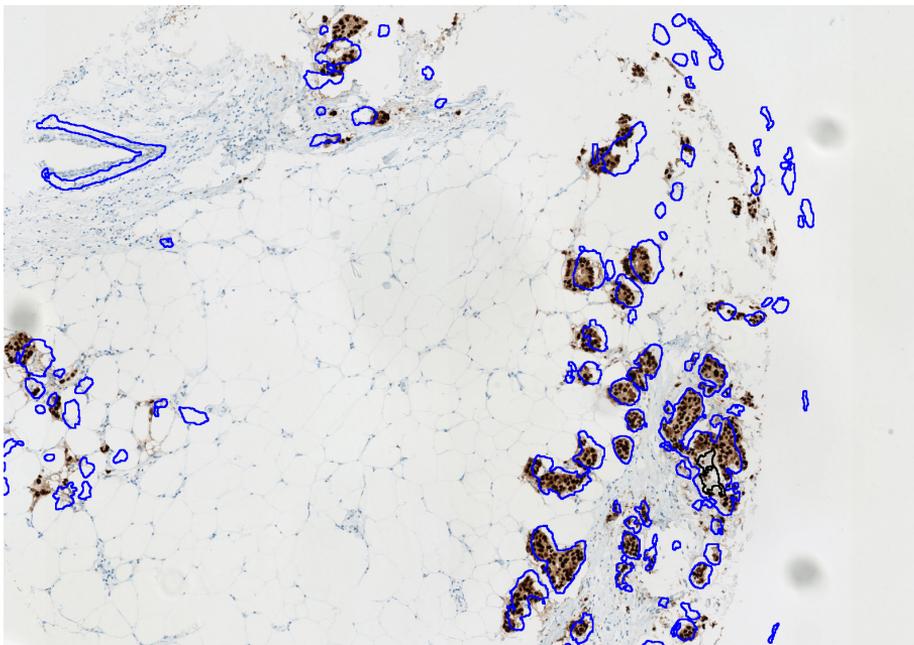
**Figure 3.2:** An ER stained core that is folded. The fold is marked with arrows. Cores that are folded, torn apart or stretched too much are removed since the alignment with PCK doesn't make sense then.

In total 18 out of 226 cores have been removed from the data set, 12 ER stained cores and 6 Ki67 stained cores.

Figure 3.3 shows an example of a core that aren't torn apart, stretch or folded but is excluded as well simply because the ER and the PCK slice are too different.



(a) PCK stained core with tumor regions marked with blue.



(b) ER stained core with tumor regions found on the PCK stained slice marked with blue.

**Figure 3.3:** Two neighbor slices that aren't sufficient similar since the tumor regions found in the PCK stained core (a) doesn't match the cells areas in the ER stained core (b).

Out of 64 scored Ki67 cores 6 cores were removed leaving 58 for analysis of the Ki67 score.

Out of 55 scored ER cores 3 cores were removed leaving 52 cores for analysis of the ER score.

### 3.3 Scores on data

As mentioned, some of the TMA cores stained with Ki67 or ER have been scored by a pathologist.

The scores for Ki67 cores are calculated as % of positives in a chosen area of the TMA core. The standard according to [16] is that the score should be based on around 400 cells which are manually counted by the pathologist in a ROI manually chosen by the pathologist. The chosen ROI should contain the highest number of positives in the core better known as a hotspot. The ROI is normally a fixed size FOV.

The scores for ER cores are H-scores defined in (2.1) or % of positives. The scores provided by the pathologist for the ER cores both contain the H-score and the % of positives in each group (stain intensity). The total % of positives for the ER cores is used instead of the H-score for the analysis since the intensity of positives aren't measured. The score is therefore referred to as the ER score.



# State of the art

---

One of the largest problems in treatment of cancer is the heterogeneity. The heterogeneity in cancer occurs due to small genetic changes in cell division of cancer cells, which results in a tumor consisting of many different cells reacting differently to treatment. This chapter contains a small review of previous research done to understand and score the cancer heterogeneity in breast cancer.

## Cancer heterogeneity

In [14] the division of cancer cells is found to follow the cancer stem cell model. Sometimes a clonal sweep within the tumor takes place which results in a completely overtake of the entire cell population by the new clone. This will not cause tumor heterogeneity since all cells will have the same genes. A heterogenous population occurs when the clone is not able to take over the entire population and thereby only some cells are replaced by the new clone. This happens a number of times and the result is a heterogenous cell population.

## Representativity of TMA studies

In [17] the representativity of using TMA cores was investigated. The concern about using TMA cores is that the small sample of the heterogeneous tumor

tissue isn't representative for the tumor and its heterogeneity. In an experiment using 553 cases of breast cancer over 80 % of ER positives were found using only one TMA core and over 98 % were identified using 4 TMA cores per patient. The core size was also shown to have an effect on the result and the 0.6 mm core was shown preferable. It was also shown that two cores per tumor is preferable to gain a more precise result.

Several studies have been performed to find out how heterogenous a tumor is and decide which treatment would be effective for each individual. The following describes the impact of biomarkers and how to score the heterogeneity based on those.

## **Ki67**

In [12] Ki67 was investigated as a proliferation marker in breast cancer. The correlation between Ki67 and clinical outcome was investigated based on several other studies with over 11.000 patients in total. In 5 out of 6 of the used studies the Ki67 score was used to predict response to chemotherapy in breast cancer. The higher Ki67 score the better response. The correlation was strongest for patients with node negative breast cancer where the correlation for patients with node positive breast cancer was positive but weaker. No correlation between the Ki67 score and endocrine treatment was found.

In [11] the impact of Ki67 as a prognostic and predictive factor on Estrogen Receptor (ER) Positive breast cancer patients was reviewed using several studies. Most studies found that Ki67 is a good prognostic factor for ER+ patients and several indicated that it also could be used as a predictor for chemotherapy. Since Ki67 scoring isn't accepted as a standard yet, the article also implies that a standardization of the staining methods and an automatic image analysis is wanted to get a standardized score for Ki67, since the impact on both treatment and prognosis is clear.

## **Impact of cut off value on survival**

In [18] three different cut off values (10 %, 14 % and 20 %) were used for dividing 369 Hormone Receptor (HR) positive, Human Epidermal growth factor Receptor 2 (HER2) negative, node negative invasive breast cancer patients into high and low risk groups and compare it with their survival. A cut off value of 14 or 20 %

---

was found most reasonable for the determination of high risk patients and their survival rates where higher with these cut off values. The analysis was based on Japanese patients and in [16] a cut off value of 14 % is recommended for use in Europe.

## Hotspot detection

In [19] the reproducibility of Ki67 scores was investigated. Four different methods for finding the hotspot were used and the scores were compared.

One of the methods was performed by an independent pathologist, where the hotspot detection was guided by a grid with squares containing cell counts. A variability of 30-40 % in counts between the pathologists was observed, when they counted in the same FOVs and it was 35-50 % higher when they counted in different, random selected FOVs. Using this method and a cut-off value of 15%, 13% cases out of 237 total were misclassified.

The determination of the hottest hotspot was also investigated to see how often the hottest hotspot was found by a pathologist as the hottest hotspot. In 50 % of the cases the hottest hotspot was chosen as the hottest one. In 23% of the second hottest hotspot was chosen as the hottest one. In 15 % of the cases the third hottest hotspot was chosen as the hottest one and the fourth was also chosen in 15 % of the cases. In 2 % of the cases the fifth hottest hotspot was chosen as the hottest one.

Another method that was investigated for calculation of the Ki67 score was an automatic digital image analysis performed using the software tool VIS[20]. The Ki67 scores were estimated based on area of cells. The results from this method were strongly reproducible and the deviation of undertreated was 4 % and 43 % overtreated. The deviation percentages for the pathologist method were 8-32 % for the untreated and 25-51 % for the overtreated.

## Visualization of heterogeneity

In [21] a HetMap for visualizing the patient's individual heterogeneity using the cell population based on HER2 and ER stained slices was created. The HetMap is a graph of the entire patient population of cells, where a cell-based heterogeneity is on one axis and a slide-level heterogeneity is on the other axis. The

tumor regions was marked by a pathologist and the two scores were calculated. The disadvantage of this method is that HER2 stained slices represent the best case scenario of heterogeneity.

# Methods

---

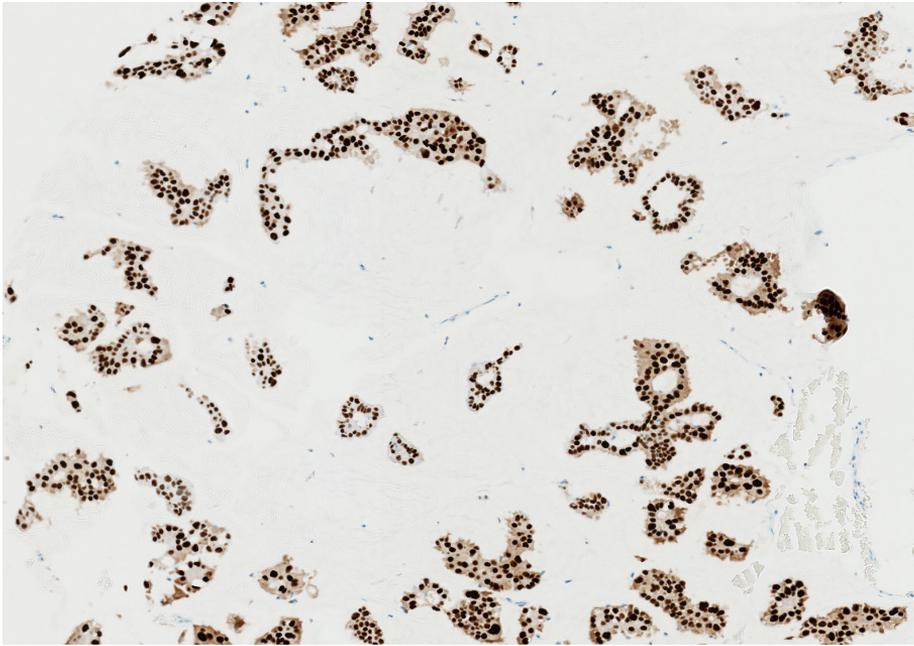
## 5.1 Image processing in VIS

The histology images are preprocessed in Visiopharm Integrator System (VIS)[20] to create segmented images that only contain positive and negative cells for the biomarker. In this section, the steps performed in VIS are described and illustrated.

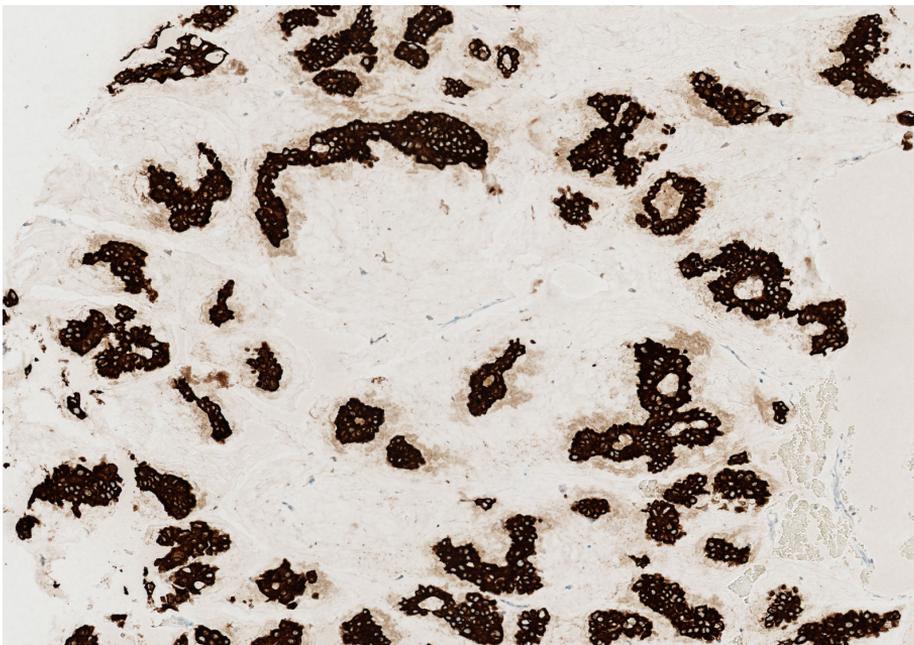
### 5.1.1 Alignment using virtual double staining

To only segment tumor cells an alignment using virtual double staining is performed. The alignment is between the ER or Ki67 stained TMA core and the PCK stained TMA core. The PCK stained TMA core is a neighbor slice to the ER or Ki67 core and therefore the content in the core is assumed to be similar. At the PCK stain slice the epithelial tissue is stained dark brown and since breast cancer originates from epithelial cells the epithelial tissue represents tumor regions.

Figure 5.1 shows an example of two aligned cores, a core stained with ER aligned with the neighbor slice stained with PCK.



(a) ER stained TMA core with positive cells colored brown.



(b) PCK stained slice with epithelial tissue colored dark brown.

**Figure 5.1:** (a) An ER stained TMA core with positive colored brown and (b) the neighbor slice stained with PCK where the epithelial tissue is stained dark brown. The two cores have been aligned.

If the alignment is too poor or the content in the cores is too dissimilar, the TMA cores are removed from the data set. All alignments are visual evaluated and the dissimilarity is evaluated as described in details and illustrated in section 3.2.

### 5.1.2 Segmentation of cells

The positive and negative cells in each ER or Ki67 stained TMA core are segmented and an image only containing two labels is made. As already mentioned the positive stained cells are colored brown and the negative stained cells are colored blue. The intensity of the color can vary between manufactures and the used parameters in the protocol are adjusted to the stain used. The protocol is described in details in the next three sections. The segmentation is only performed within the tumor regions found after the alignment.

#### 5.1.2.1 Preprocessing

The performed preprocessing of each image is:

1. **Red band image** - The red band image of the RGB image is used for the analysis. In the red band image the difference between the brown and blue color is smallest and since the negative cells are blue and the positive cells are brown, this image is used for the segmentation.
2. **Mean filter** - The image is afterwards filtered using a mean filter to minimize noise.
3. **Polynomial filter** - The image is then filtered using a local linear polynomial filter to find the edges of elongated shapes.
4. **Blob detection** - The image is also filtered twice with a blob filter of different sizes. It is done twice to ensure to find small and large round cells.

#### 5.1.2.2 Classification

A classification of the preprocessed image is performed using a simple threshold classifier. The threshold value depends on the stain used since the intensities varies across manufacturers. Labels are assigned based on the classifier.

### 5.1.2.3 Postprocessing

After the classification some post processing steps is performed.

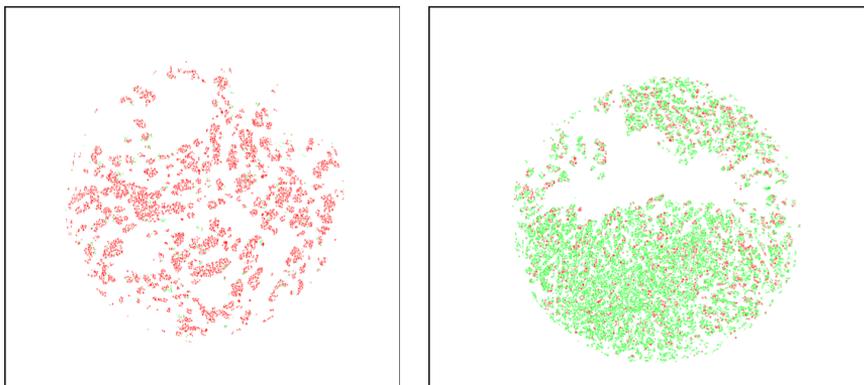
- **Label clean up** - Clean up in labels from the classification. Some criteria are set up for labels to find out whether or not they belong to a cell or the background e.g. based on shape and size. Separation of cells is also performed at this step using watershed algorithm.
- **Erosion** - Erosion of the segmented cells is performed to make sure that the cells stay separated after the export from VIS since the image is down-scaled at export. The erosion is done by assigning a different label to the edge of the cells with a defined width.

The labeled images with positive and negative cells are a segmentation of the cells and the images are exported from the software for further analysis. In the next section some examples of exported images are shown.

### 5.1.3 Examples of the segmentation of cells

The images with the segmented cells are exported from the software with a magnification of 4 as bmp files and this section shows two examples of the exported images used for the analysis.

Figure 5.2 shows some examples of the exported images from VIS with negative (green) and positive (red) cells.



(a) An ER stained image with positive (red) cells and a few negative (green) cells exported from VIS.

(b) An Ki67 stained image with a mixture of positive (red) and negative (green) cells exported from VIS.

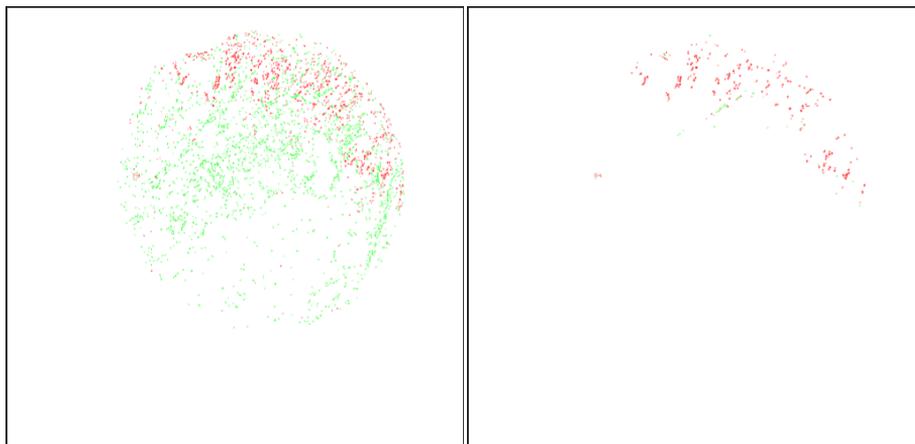
**Figure 5.2:** Examples of exported images from VIS with segmented cells, (a) an ER stained image with positive (red) cells and a few negative (green) cells and (b) a Ki67 stained image with a mixture of positive (red) and negative (green) cells.

## 5.2 Outline of TMA core

The outline of the TMA core is found on each exported image with the segmented positive and negative cells. The outline is used to make a binary TMA core mask for each TMA core.

The mask is multiplied on the image to exclude cells from neighbor TMA cores that might be visible in the image and to exclude cells in fragments of the TMA core that also have been segmented.

The image with segmented tumor cells can't be used to find the outline since the tumor regions don't necessarily cover the whole outline. A segmentation of all cells in the TMA core is therefore performed and the image is used to find the outline. Figure 5.3 shows the segmented cells in the whole core and the segmented cells in the tumor regions in the same TMA core.

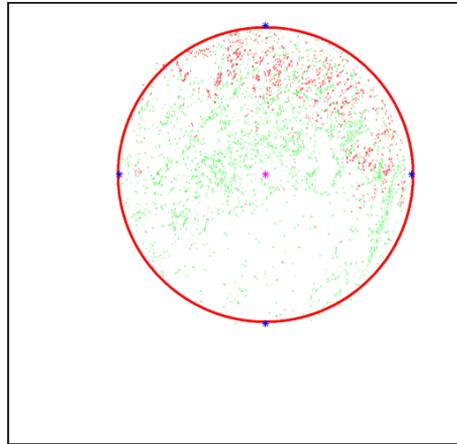


(a) ER positive and negative cells segmented in the whole core. (b) ER positive and negative cells segmented in the tumor regions.

**Figure 5.3:** Example of (a) a TMA core where all cells in the core are segmented and (b) the same TMA core where cells inside tumor regions have been segmented.

Four points on the image with segmented cells in the whole core are found using minimum and maximum x and y coordinates and the best geometric fit for a circle is found by minimizing the orthogonal distances using standard Levenberg-Marquardt optimization [22].

Figure 5.4 shows the same TMA core as on figure 5.3 with the best fitted circle as the outline of the TMA core. The blue points are the 4 points used for the circle fit and the pink point is the center of the fitted circle.



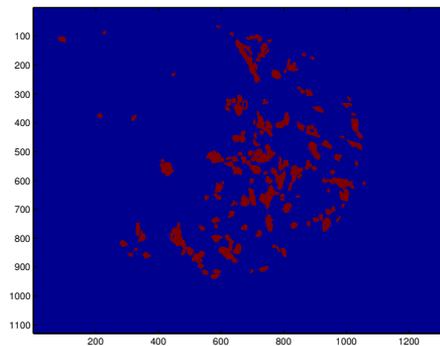
**Figure 5.4:** The same TMA core as on figure 5.3 with the fitted circle (red), the four used points (blue) and the center of the circle (pink).

### 5.3 Tumor masks

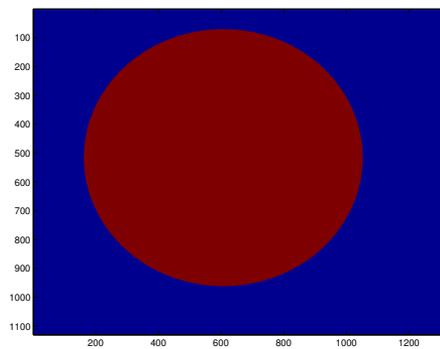
A tumor mask for each TMA core is also made and is e.g. used to calculate the amount of tumor tissue in each TMA core. The tumor regions found in the PCK stained image are exported from VIS as an image and converted into a binary image for all TMA cores.

All tumor masks are multiplied with their binary TMA core mask (described in section 5.2) to remove segmented regions outside the core that occurs due to air, dust or small fractions of epithelial tissue lying around the TMA core.

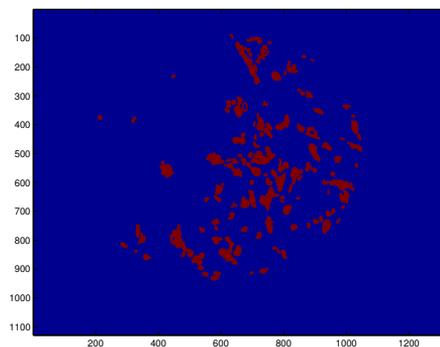
Figure 5.5 shows an example of a tumor mask with tumor regions outside the TMA core and the regions after multiplication with the TMA core mask.



(a) Tumor mask exported from VIS.



(b) TMA core mask found as described in section 5.2.



(c) Tumor mask multiplied with TMA core mask. Regions outside the TMA core have been removed.

**Figure 5.5:** Example of tumor mask with regions outside the TMA core. (a) the raw tumor mask, (b) the TMA core mask and (c) the tumor mask and the TMA core mask multiplied to eliminate small objects labeled as tumor tissue but lying outside the TMA core.

### 5.3.1 Amount of tumor tissue inside TMA core

The amount of tumor tissue in each TMA core is calculated. This is done by calculating the area of each tumor mask and each TMA core mask in pixels.

The percentage of tumor tissue in each TMA core is given as

$$\%_{\text{tumor tissue}} = \frac{\text{Number of pixels in tumor mask}}{\text{Number of pixels in TMA core mask}} \times 100 \quad (5.1)$$

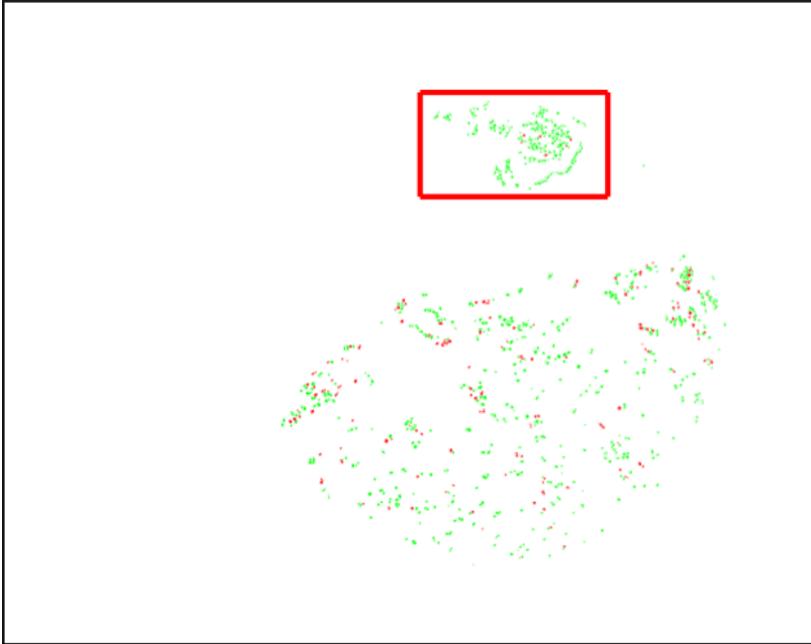
The results can be found in section 6.2.

## 5.4 Blob detection

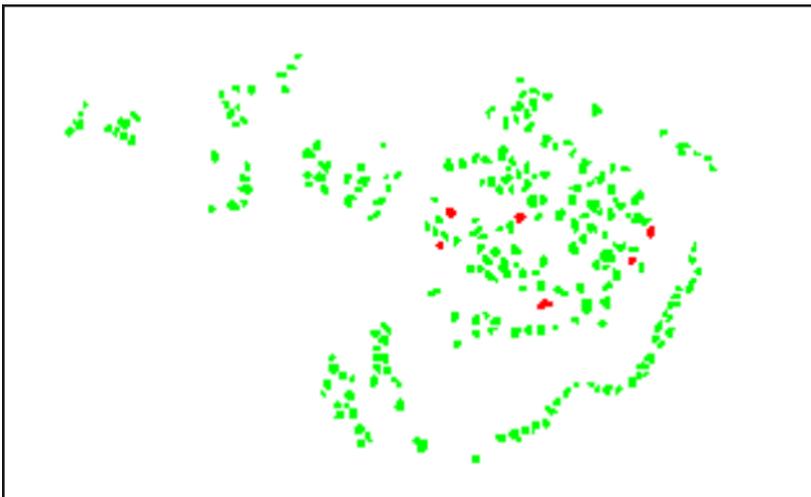
A blob detection in MATLAB is performed on images with segmented cells to find the cells and their center. Since the cells were eroded in VIS to make sure that they stayed separated after the downscaling of the image under export the size of the cells aren't physically representable. Therefore the cells are presented by their center found in a blob detection instead.

A blob detection of the positive cells and the negative cells is performed separately to know if the cell is positive or negative. The positive cells lie in the red band of the RGB image and the negative cells in the green band and a blob detection of the image in each band is therefore performed.

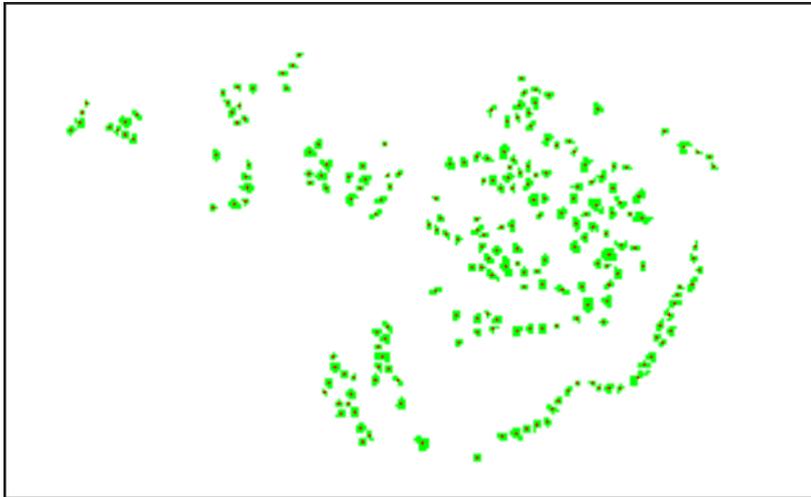
Figure 5.6 shows an example of the two blob detections of the positive (red) and negative (green) cells. For a better visualization a zoom box is used. The zoom box is marked by a red square on image figure5.6a and the rest of the images (figure 5.6b-5.6d) are images inside the zoom box.



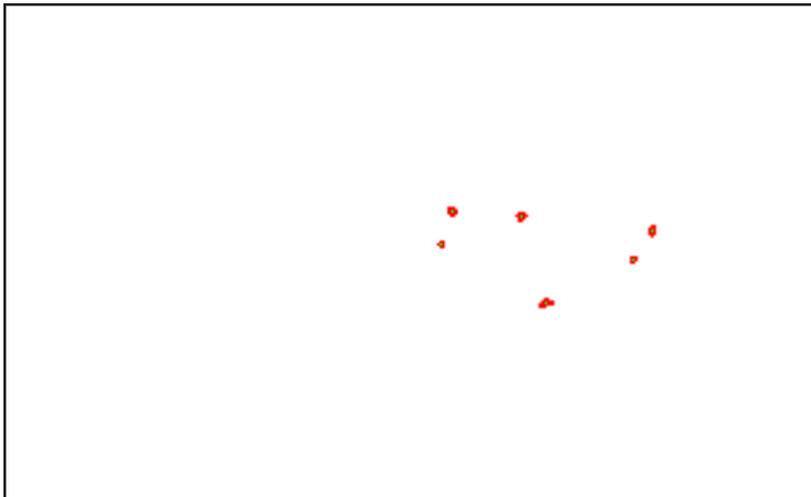
(a) The image from VIS with segmented positive (red) and negative (green) cells. The red square is a zoom box used to visualize the positive and negative cells.



(b) Zoom in on (a) marked by a red square for a better visualization of the positive (red) and negative (green) cells.



(c) The negative cells in the zoom box and their center (marked by red dot) found in the blob detection of the negative cells.



(d) The positive cells in the zoom box and their center (marked by green dot) found in the blob detection of the positive cells.

**Figure 5.6:** Visualization of the blob detection of the positive and the negative cells. (a) the image from VIS with segmented positive (red) and negative (green) cells. The red square is a zoom box used to get a better visualization of the small cells. (b) The image on (a) in the zoom box. (c) the positive cells and their center from the blob detection and (d) the negative cells and their center from the blob detection.

The centers found for the positive and negative cells are put into two images each representing the positive and negative cells. These images are used for the rest of the analysis based on cells.

## 5.5 Heatmap calculation

This section describes the methods used and considerations made about the heatmap generation. First the performed semivariogram analysis is described, then the sliding window function and the gauss filter are described and some heatmap examples are shown.

### 5.5.1 Semivariogram analysis

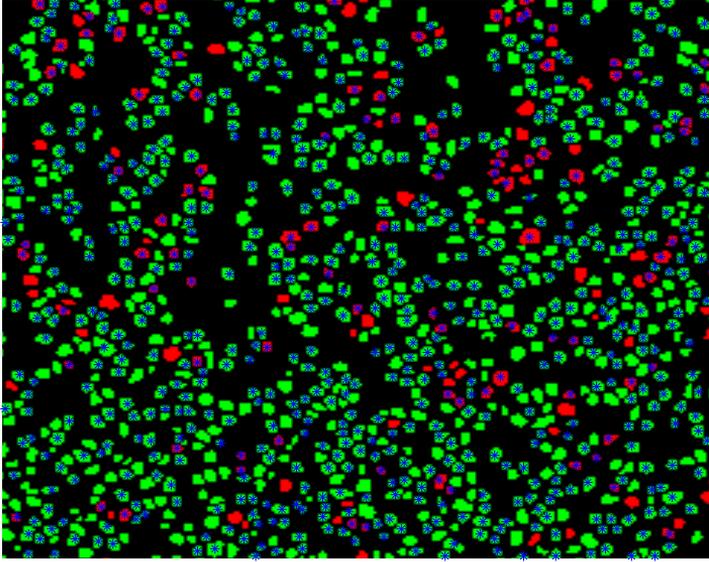
To investigate the spatial correlation, a semivariogram analysis is performed for both the ER stained and the Ki67 stained images. The reason that the spatial correlation is investigated separately in the two image types is that the cells stained by ER might lie differently than the cells stained by Ki67.

The theory for calculating the semivariogram is described in section 2.5. This section contains some examples of the performed semivariogram analysis, the used parameters and other considerations.

The used images for the semivariogram analysis are the images with cells represented by their center from the blob detection put into one image. The image contains one pixel for each cell either positive with value 1 or negative with value 0. The background is valued as Not-a-Number (NaN) to exclude these pixels.

A few of the images contain over 30.000 cells and due to computation speed and working memory capacity only 5000 random sampled cells are used for the calculation of the semivariogram. Most of the images contains between 2000 and 10.000 cells and 5000 is a good representative for images containing more than 5000 cells.

Figure 5.7 shows a zoom on one image with 5000 random sampled cells between 9000 cells total. As it can be seen the cells are well represented and the sampling aren't spread too much so the estimation of the spatial correlation still makes sense.



**Figure 5.7:** Zoom on image where 5000 cells are random sampled between 9000 cells in total. The random sampled cells are marked by a blue star and as it can be seen the sampling well represent the cells. Almost all images contain less than 10.000 cells. The estimated range for this image is 6.54 pixels.

A large matrix  $X$  is set up for the analysis and the coordinates for the negative cells  $N(x, y)$  and the positive cells  $P(x, y)$  are placed in column 1 and 2. The value 0 or 1 depending on whether it is a negative or positive cell is placed in column 3. The set up of the  $X$  matrix is given as

$$X = \begin{bmatrix} & & 0 \\ N(x, y) & & \vdots \\ & & 0 \\ & & 1 \\ P(x, y) & & \vdots \\ & & 1 \end{bmatrix} \quad (5.2)$$

The matrix is used as an input to the function for calculation of the semivariogram provided by Allan Aasbjerg Nielsen. It can be found as appendix A. The

function calculates the semivariogram value for all point pairs provided in the matrix  $X$ .

A spherical model is fitted to the calculated semivariogram to find the range, sill and nugget for the semivariogram as described in section 2.5.

#### 5.5.1.1 Considerations

Images with only positive or negative cells were removed from the semivariogram analysis since it didn't make sense to calculate the correlation between positive and negative cells in these images. The analysis is based on 56 out of 110 ER stained images and 76 out of 114 Ki67 stained images.

The used lag distance is 1 pixel and the number of lags used is 100.

The results from the semivariogram analysis are presented in section 6.1.

### 5.5.2 Heatmap

The heatmap is a visualization of the heterogeneity by visualizing the response to one of the two used biomarkers as the percentage of positive cells in a given area. The heatmap is also used as a guide for finding hotspots to calculate the heterogeneity score in. The generation of the heatmap is described in the following sections.

### 5.5.3 Sliding window function

To generate the heatmap the percentage of positive cells in small areas needs to be calculated.

The images from the blob detection containing positive and negative cells are used.

A sliding window is used for calculation of the percentage of positives in all pixels inside the TMA core defined by the TMA core mask. This is practically done by taking a ROI in both images (image with positive or negative cells) and then multiply them with a circular function. Afterwards the sum in both

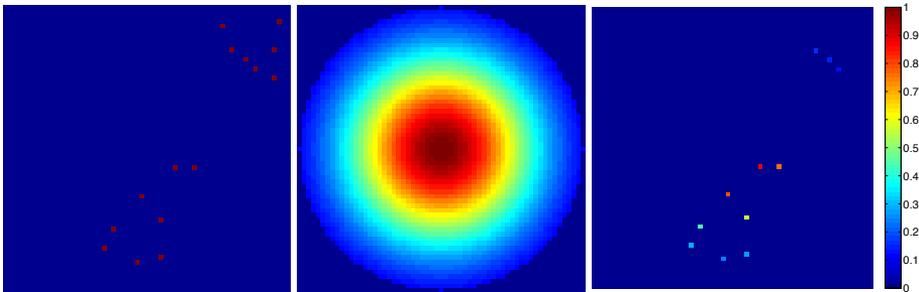
images is calculated and the percentage of positives cells is calculated as

$$\%_{positives} = \frac{\sum Pixels_{positive}}{(\sum Pixels_{positive} + Pixels_{negative})} \quad (5.3)$$

A new ROI is taken out from the image by moving the ROI 1 pixel at the time and thereby sliding through the whole TMA core.

To take the position of the cells inside the ROI into account, the circular window is gauss weighted. The used gauss filter is described in the next section.

Figure 5.8 shows a ROI in the image with positive cells, the circular gauss weighted window function and the result of those two multiplied.



(a) Positive cells in a chosen ROI at same size as the circular gauss window function. (b) The circular gauss window function. (c) Positive cells in the ROI after multiplication with the circular gauss filter. The remaining cells can have values between 0.11 and 1. The sum is 4.68.

**Figure 5.8:** An example of the sliding window function where a ROI of the image with positive cells on (a) is multiplied with the circular gauss filter on (b) and the remaining cells on (c) have values corresponding to their distance to the center of the circle. The values in (c) are summed and corresponds to the value for positive cells in that ROI. The same procedure is done for the negative cell image and the percentage of the positive cells is assigned to the center of the ROI.

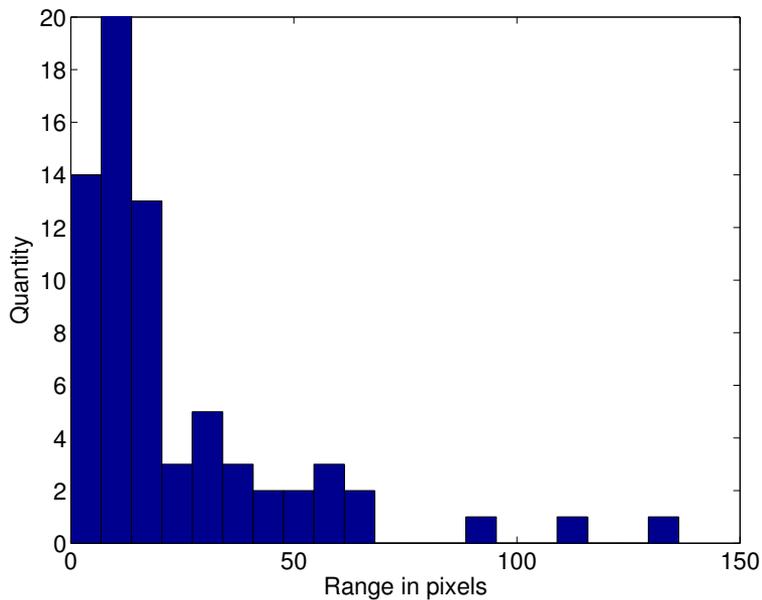
### 5.5.4 The Gauss filter used

The window function is gauss weighted to weight the impact of cells placed close to the center higher than cells placed close to the periphery of the circle. The median of the found ranges in the semivariogram analysis is used as the standard deviation  $\sigma$  in the gaussian function. The used ranges are presented in table 5.1. The reason why the median and not the mean is used is a few outliers that maximizes the mean artificial. The two distributions of range can be seen on figure 5.9.

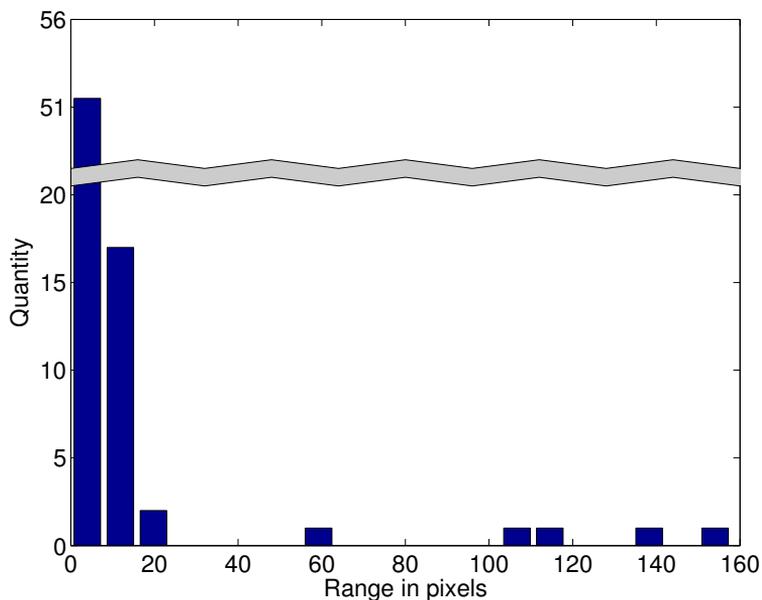
Median of Range	
Ki67	6.93
ER	14.40

**Table 5.1:** The median of the ranges for Ki67 stained images and ER stained images from the semivariogram analyses.

The used window functions for Ki67 stained images and ER stained images can be seen on figure 5.10a and 5.10b.

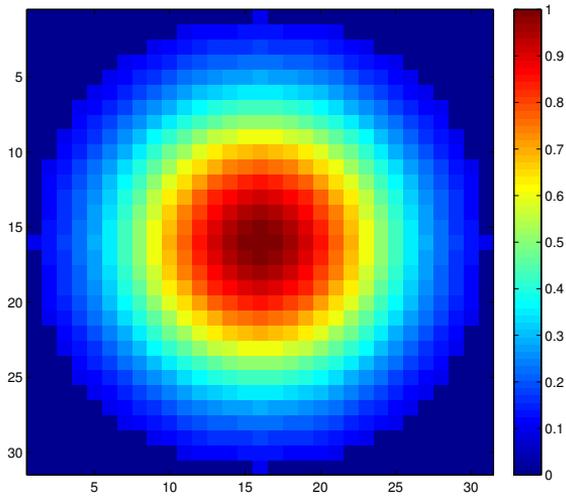


(a) Distribution of ranges from ER stained images.

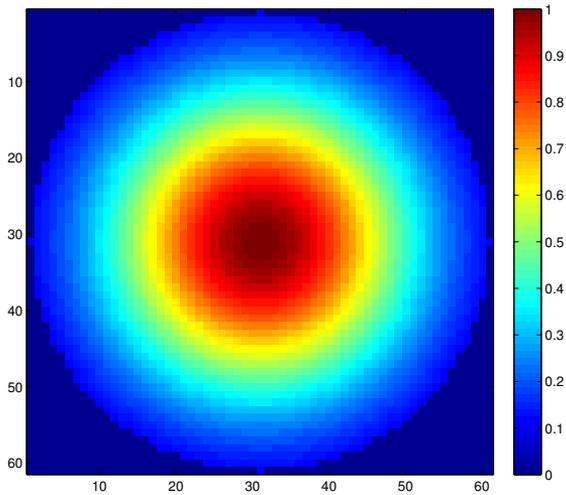


(b) Distribution of ranges from Ki67 stained images. A few outliers are seen.

**Figure 5.9:** The two distribution of range from the two semivariogram analyses. (a) the distribution of range from the semivariogram analysis of ER stained images and (b) the distribution of range from the semivariogram analysis of Ki67 stained images.



(a) The circular gauss weighted window function used for Ki67 stained images. The standard deviation  $\sigma$  of the used gaussian function is 6.93. The size of the window function is 31 x 31.



(b) The circular gauss weighted window function used on ER stained images. The standard deviation  $\sigma$  of the used gaussian function is 14.4. The size of the window function is 61 x 61.

**Figure 5.10:** The two circular gauss weighted window functions used for the heatmaps for the Ki67 and ER stained images respectively.

### 5.5.5 Heatmap parameters

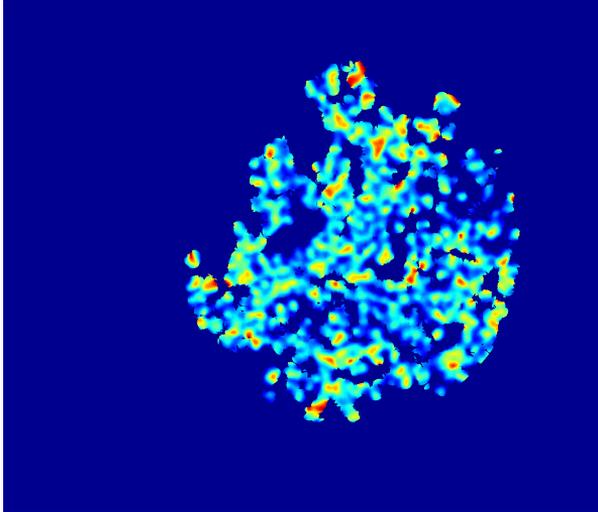
A few parameters for the heatmaps are calculated and used to describe the heatmaps. The parameters are only calculated inside the tumor regions of the heatmap, which means that the heatmap is multiplied with both the TMA core mask and the tumor mask.

A calculated heatmap and the heatmap multiplied with tumor and TMA core mask is shown on figure 5.11.

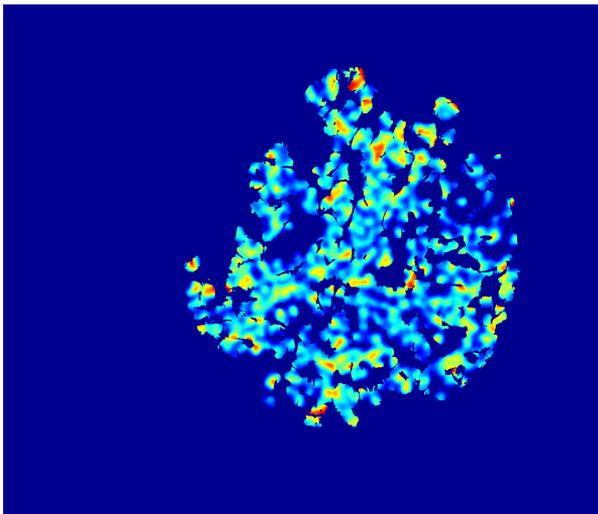
The calculated parameters for each heatmap are:

- Mean value  $\mu$
- Variance  $\sigma^2$
- Histogram

The mean value and variance for all heatmaps and the cumulated histogram for some heatmaps can be found in section 6.5.



(a) Calculated heatmap for a Ki67 stained image.



(b) The same heatmap as on (a) but after multiplication with tumor mask and TMA core mask.

**Figure 5.11:** (a) Calculated heatmap for a Ki67 stained image and (b) the same heatmap but after multiplication with the tumor mask and TMA core mask done before calculation of mean value and variance.

## 5.6 Hotspot detection

A detection of the hottest hotspot in the heatmap is performed and is used to calculate a heterogeneity score in the TMA core, either the Ki67 score or the ER score.

The method for finding the hottest hotspot is described in this section and the calculation of score is described in section 5.6.3.

The results can be found in section 6.4.

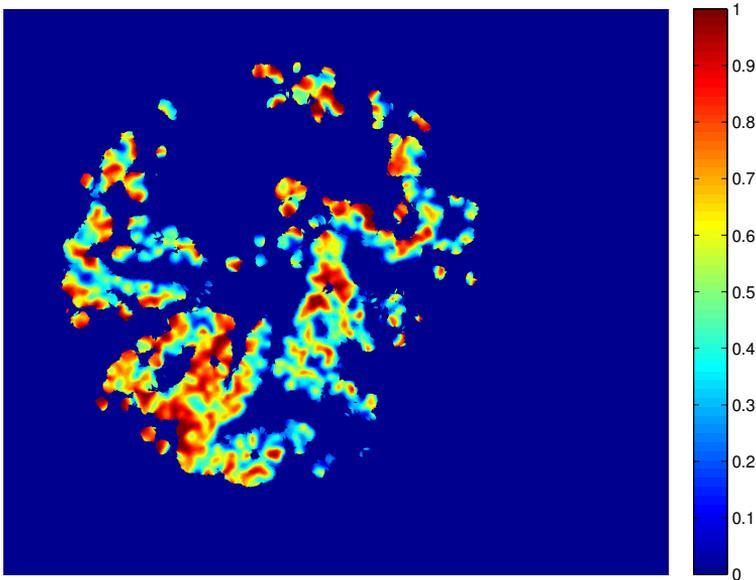
### 5.6.1 Detecting largest hotspot

The largest hotspot in the heatmap is the largest area with the the relative maximum in the heatmap as pixels value. In most heatmaps the value will be 1 or close to 1 but for some heatmaps based on images with few positives the maximum might be 0.5 or lower. Therefore the relative maximum is used instead of using 1.

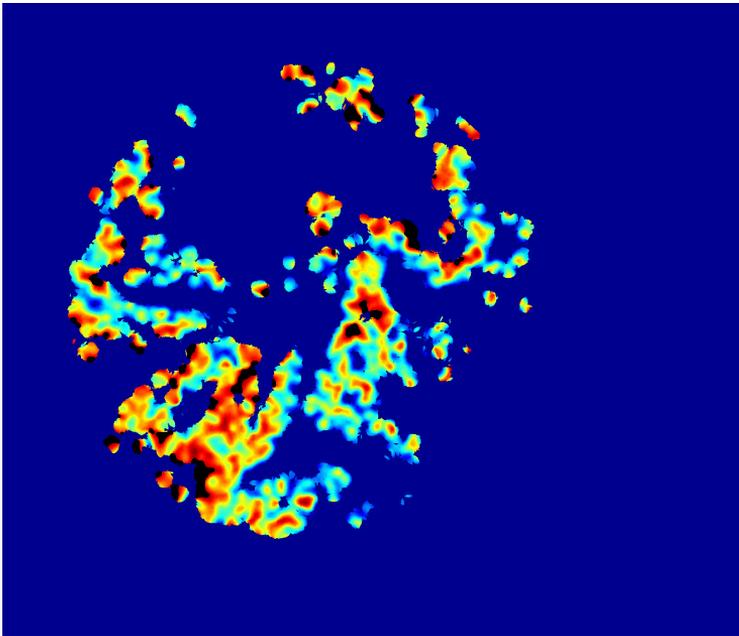
The following steps are done to find the largest hotspot

- **Threshold** - A simple threshold of the heatmap with the relative maximum of the heatmap as the threshold value.
- **Blob detection** - A blob detection of the found areas with the relative maximum value if more than one exist. The blobs are sorted by size and the largest blob is used as the hottest hotspot.
- **Dilation until 60 % of max value using seed growing** - The largest blob found is dilated by including neighbor pixels iteratively if their value is larger than 60 % of the maximum value. The iterative process is started at the edge of the blob found by an edge detection. The maximum number of iterations is 100. The number is set to avoid a too large hotspot resulting from wrong placement of the FOV. The hotspot is dilated to make sure the FOV contains a representative distribution of the neighborhood of the hotspot.
- **Center** - The center of the dilated blob is calculated.

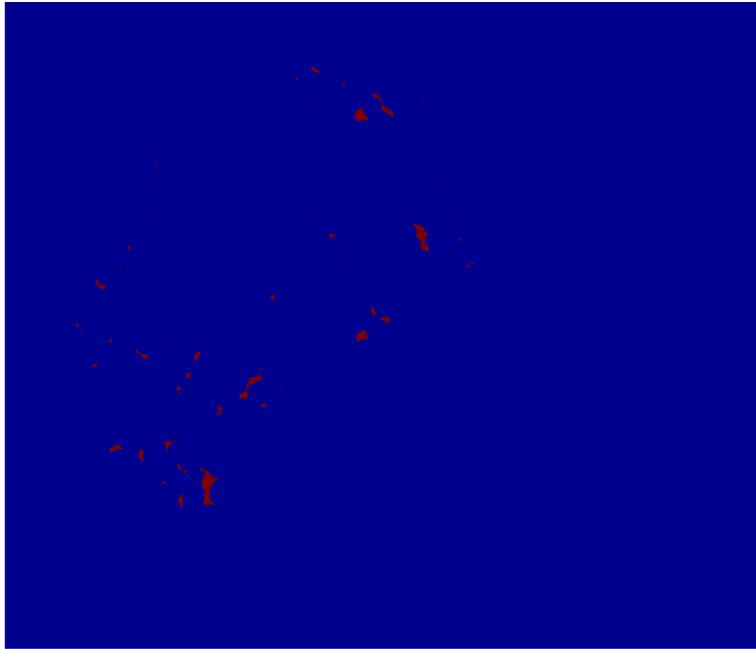
Figure 5.12 illustrates the process of finding the largest hotspot.



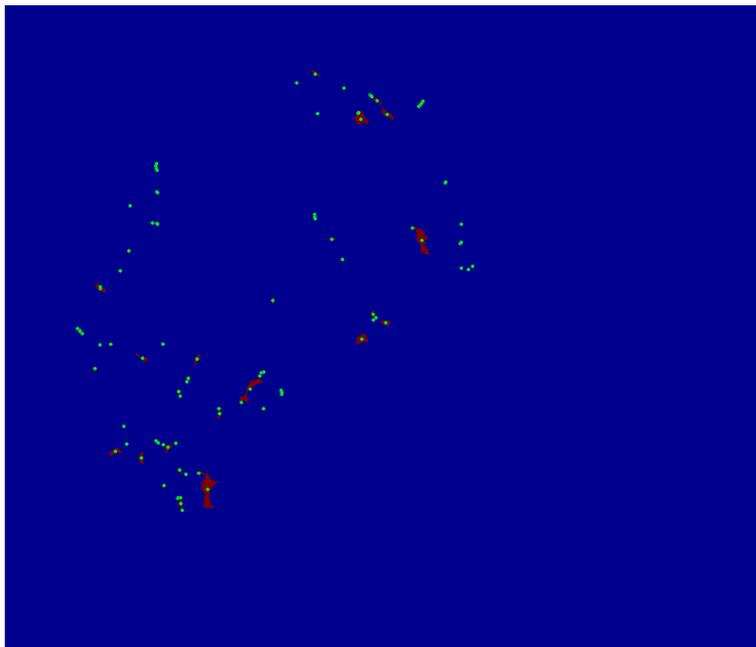
(a) The heatmap used for the illustration of hottest hotspot detection.



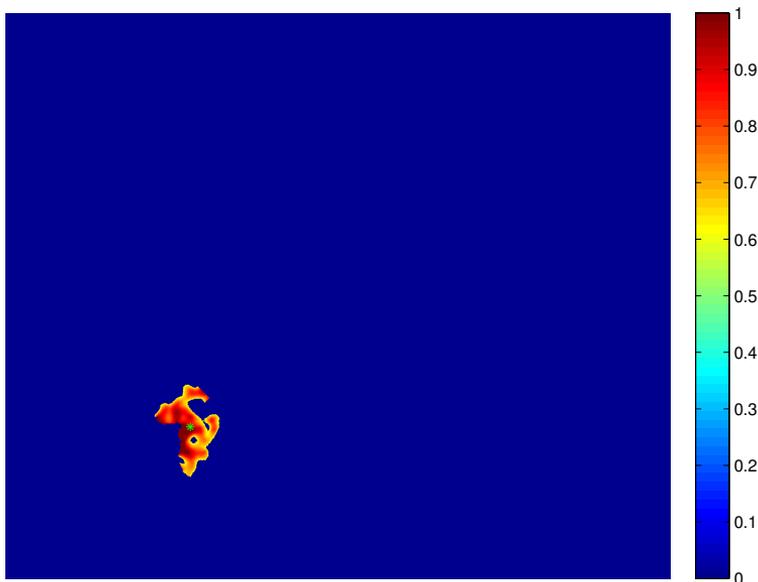
(b) The maximum value is marked with black at the heatmap.



(c) The image after threshold with the relative maximum as threshold value.



(d) The threshold image with centers of blobs marked by a green dot found in the blob detection. The largest blob is chosen based on area.



(e) The found largest blob dilated using seed growing - which is done iteratively by including neighbor pixels with values larger than 60 % of the maximum value. The maximum number of iterations is 100. The center of the hotspot is marked by a green star.

**Figure 5.12:** Illustration of the process with detecting and dilating the hottest hotspot. (a) the heatmap, (b) relative maximum value on the heatmap marked by black, (c) result of threshold with the relative maximum value, (d) blob detection of the threshold image to find largest blob and (e) dilated hottest hotspot using seed growing with max 100 iterations.

### 5.6.2 Field Of View (FOV)

The score is calculated inside a Field Of View (FOV). The FOV used by the pathologist is of defined size but should minimum contain 400 cells [16]. The pathologist places the FOV manually at a site where the heterogeneity is representative and often in the hottest hotspot which is found subjectively.

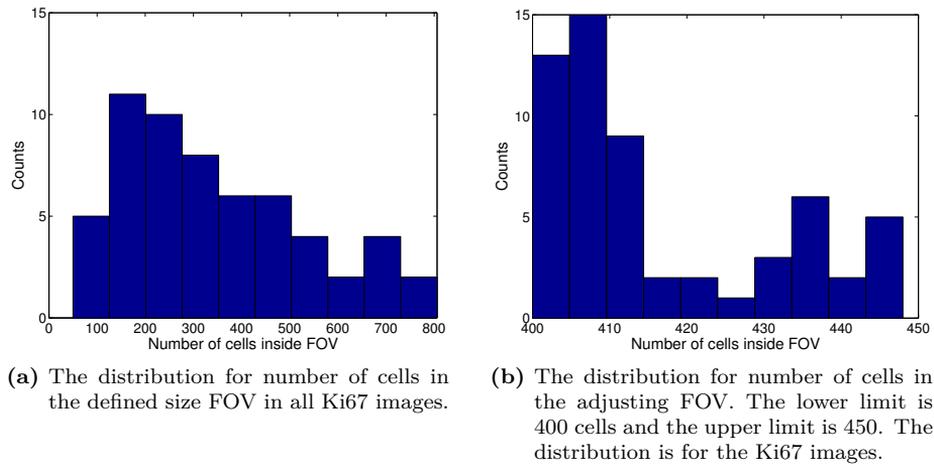
Two types of FOV is used for calculation of the heterogeneity score, a fixed size FOV and an adjusting FOV that vary in size depending on the number of cells inside the FOV.

The size of the FOV is  $88 \mu\text{m} \times 88 \mu\text{m}$  and the TMA core is between 1700-2000

$\mu\text{m}$  in diameter for comparison.

Most ER stained images is either 100 % positive or negative (score 0) which means that the FOV placed at a hotspot doesn't make sense since either the whole core is a 'hotspot' or no hotspots exist. Therefore the FOV for ER stained images is placed randomly if the size of the hotspot is above the size of the defined size FOV.

Figure 5.13 shows the distributions of cells inside the two types of FOVs.



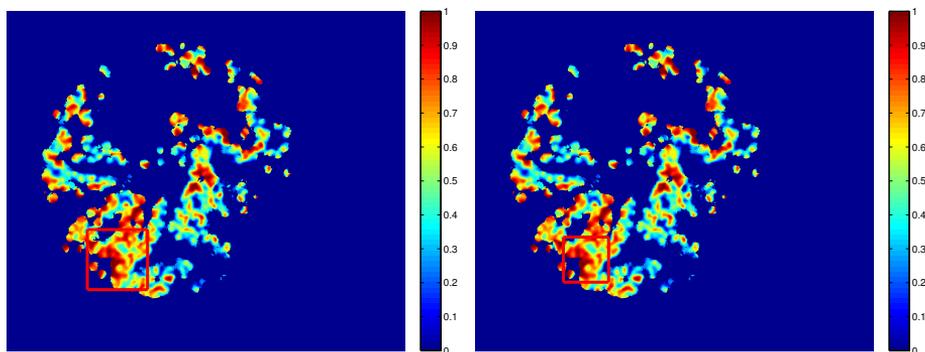
**Figure 5.13:** The distribution of cells in the FOV for the two types of FOV, (a) the defined size FOV and (b) the adjusting FOV with lower limit of 400 cells and upper limit of 450 cells.

### 5.6.2.1 Considerations about the chosen FOV type

The distribution for the defined size FOV on figure 5.13a varies between 100 cells and up to 800 cells in a FOV, where most FOVs' contain less than 400 cells. The FOV size was tried varied, but the distribution of cells within and around the hotspot varies too much to find a size that made sure that the FOV contained at least 400 cells and a realistic maximum of cells that would make sense to count manually. Therefore the adjusting FOV was chosen.

Figure 5.14 shows the two FOV types on the same heatmap as used for the illustration on figure 5.12. From the figure it can be seen that the amount of cells in the hotspot is greater than 450 for the defined size of FOV and the

adjusting FOV is therefore decreased to make the FOV contain between 400 and 450 cells.



(a) The same heatmap as used in the illustration of the hotspot detection on figure 5.12; here with FOV of defined size. (b) The same heatmap as used in the illustration of the hotspot detection on figure 5.12; here with adjusting FOV which is decreased compared to the FOV of defined size on (a).

**Figure 5.14:** The two types of FOV, (a) a FOV of defined size and (b) a FOV with adjusting size until it contains minimum 400 cells. In this case the FOV is decreased a bit compared to the FOV of defined size.

### 5.6.3 Calculation of heterogeneity score at hottest hotspot

The heterogeneity score is calculated inside the FOV as percentage of positives cells out of the total sum of cells. The cells are counted in the positive and negative cell images from the blob detection with one pixel representing each cell. Analysis of the calculated scores and comparison between them and the pathologist scores are described in section 6.6 for Ki67 and section 6.10 for ER scores.

A calculation of the heterogeneity score outside the hottest hotspot is also performed for Ki67 images and is presented in section 6.7.

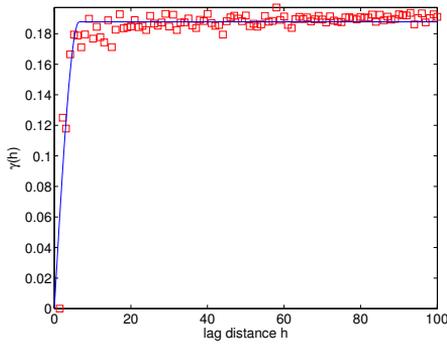
A linear regression model is fitted to the calculated scores and the best fitted model is found. A F-test of each model's slope different from 1 is performed to investigate whether the scores are significantly different from the pathologist scores. All fitted models and statistical tests are described in sections 6.6, 6.7 and 6.10.

## 6.1 Segmented cells and their semivariogram

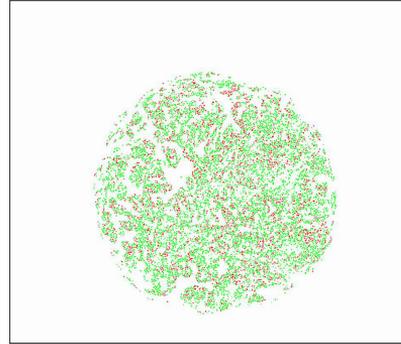
This section contains results for some of the images with segmented cells and their semivariogram with a fitted spherical model.

Figure 6.1 and figure 6.2 show images with segmented cells for two Ki67 stained images and the calculated semivariogram with a fitted spherical model.

Figure 6.3 and figure 6.4 show images with segmented cells for two ER stained images and the calculated semivariogram with a fitted spherical model.

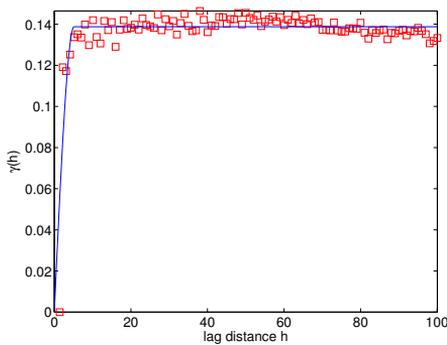


(a) Semivariogram (red squares) with fitted spherical model (blue line) for the Ki67 stained image shown on (b).

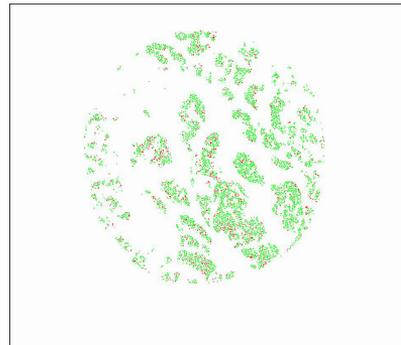


(b) A Ki67 stained image with a great mixture of closely lying positive cells (red) and negative cells (green).

**Figure 6.1:** (a) The semivariogram (red squares) with a fitted spherical model (blue line) for the Ki67 stained image on (b).

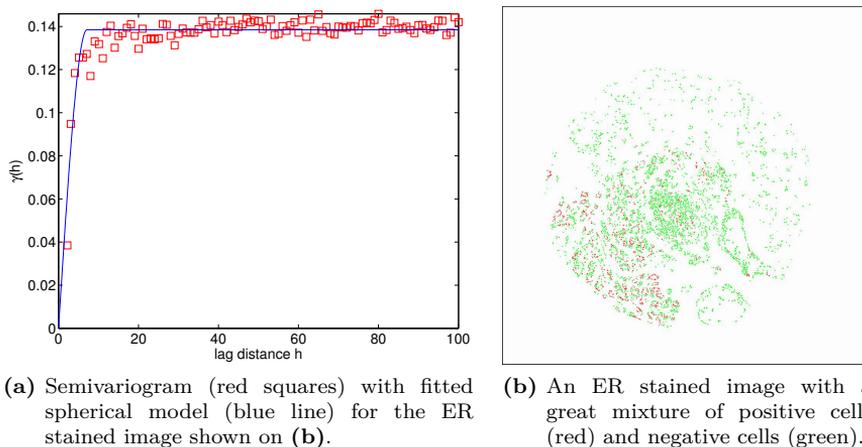


(a) Semivariogram (red squares) with fitted spherical model (blue line) for the Ki67 stained image shown on (b).

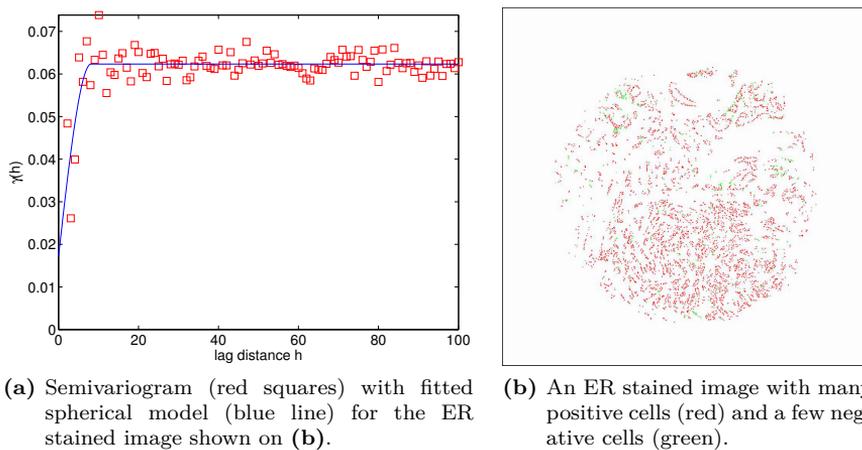


(b) A Ki67 stained image with more spread areas with positive cells (red) and negative cells (green).

**Figure 6.2:** (a) The semivariogram (red squares) with a fitted spherical model (blue line) for the Ki67 stained image on (b).



**Figure 6.3:** (a) The semivariogram (red squares) with a fitted spherical model (blue line) for the ER stained image on (b).

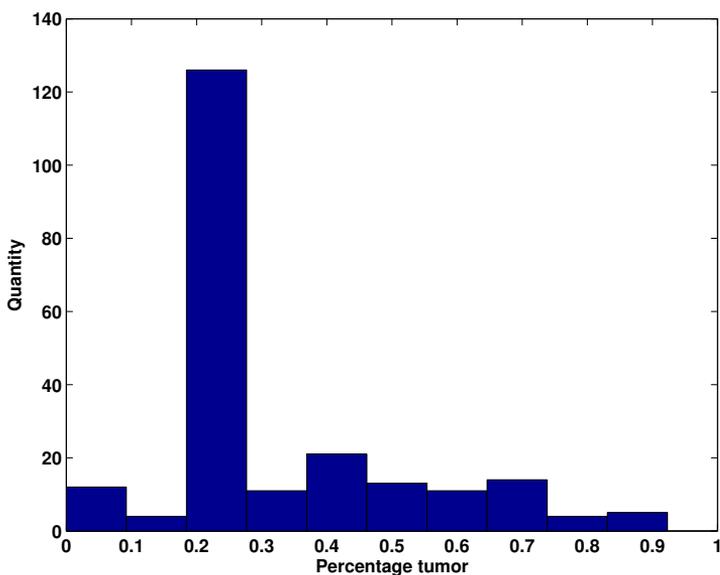


**Figure 6.4:** (a) The semivariogram (red squares) with a fitted spherical model (blue line) for the ER stained image on (b).

## 6.2 Amount of tumor tissue inside TMA core

This section contains the distribution of tumor tissue in percentage in the TMA cores in the dataset.

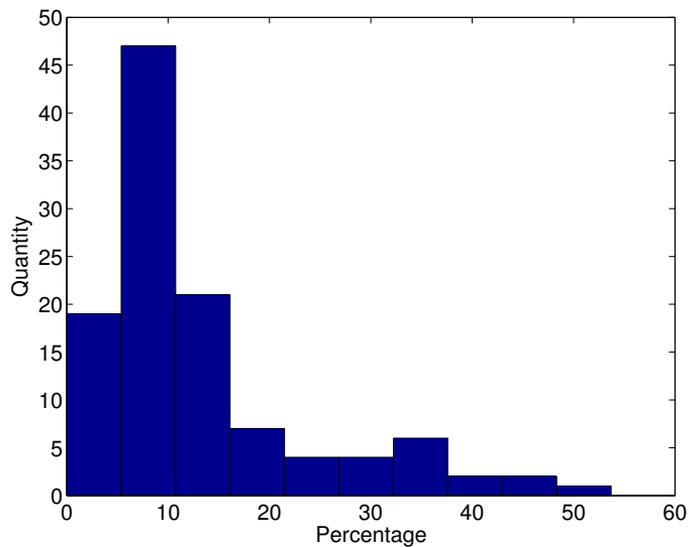
Figure 6.5 shows the total distribution of percentage of tumor tissue in all TMA cores. The distributions for Ki67 and ER stained cores can be found in appendix B.



**Figure 6.5:** Distribution of percentage tumor tissue inside TMA core for all stained cores.

It is clearly seen that most TMA cores contain between 20-30 % tumor tissue. The median of the histogram for all TMA cores is 26.7 %.

Figure 6.6 shows the distribution of tumor tissue in the FOV as percentage of the total amount of tumor tissue in the TMA core.



**Figure 6.6:** Distribution of tumor tissue inside the FOV as percentage of the total amount of tumor tissue in the TMA core.

The results are discussed in section 7.1.

## 6.3 Heatmaps

This section contains some of the calculated heatmaps for the Ki67 stained and ER stained images. The chosen heatmaps show the three types of heatmaps that exist in the two types of stain.

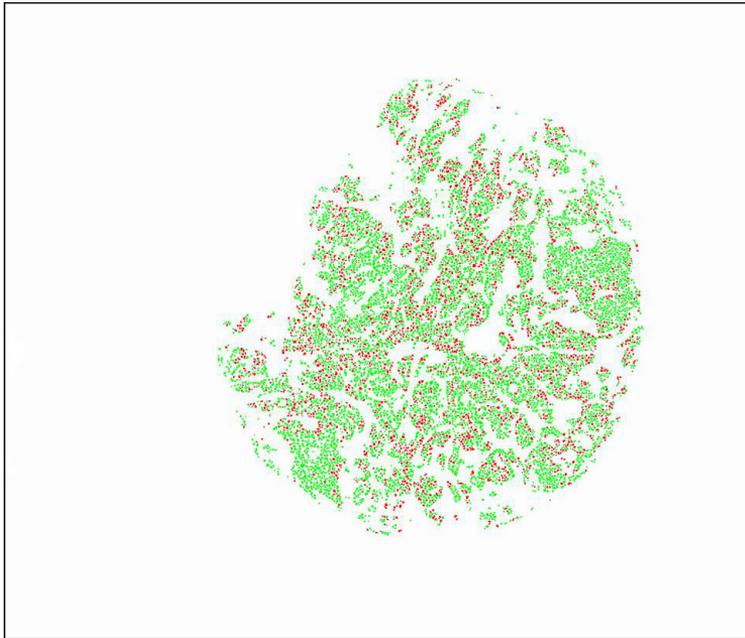
### 6.3.1 Ki67 Heatmaps

Figure 6.7 shows the Ki67 image with a great mixture of positive and negative cells and the calculated heatmap. The hotspots are clearly seen on the heatmap but they are hard to identify on the image with the segmented cells.

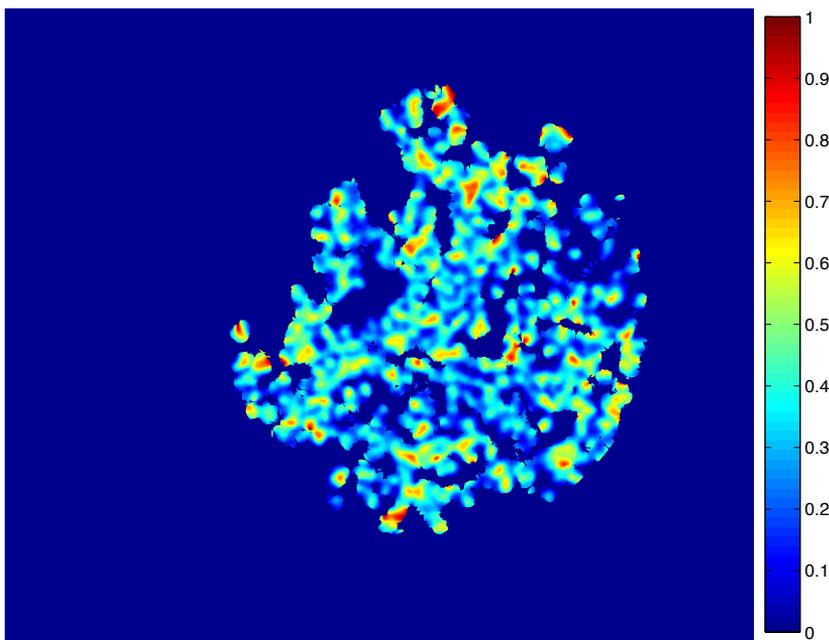
Figure 6.8 shows the Ki67 image with a great mixture of positive and negative cells and the calculated heatmap. Areas with positive cells are clearly seen on

the segmented cell image and the heatmap helps visualize the density of the cells in these areas.

Figure 6.9 shows the Ki67 image with few cells mostly negative cells and a few positive cells and the calculated heatmap. Hotspots are hard to identify on the segmented cell image due to the varying density of negative cells. The heatmap clearly shows the hottest hotspot and a few other hotspots.

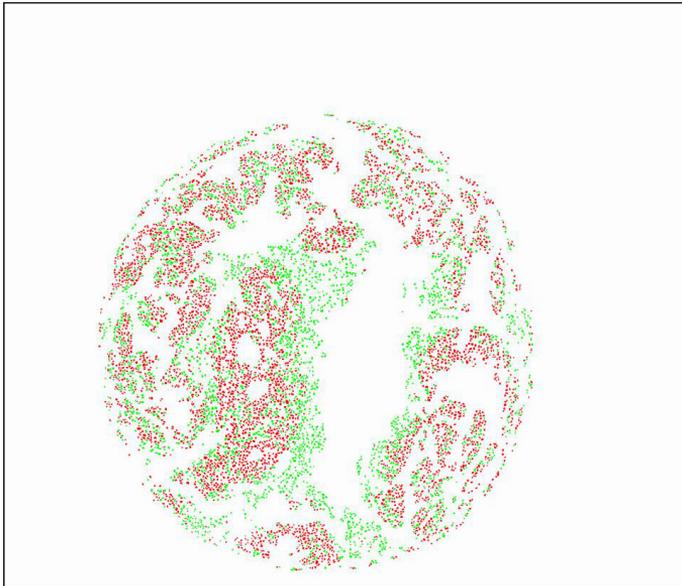


(a) Image with positive (red) and negative (green) Ki67 stained cells.

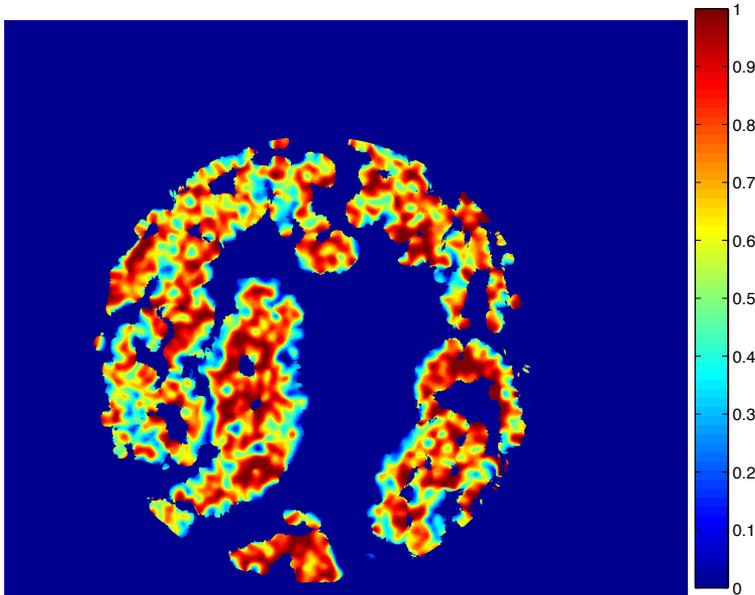


(b) Calculated heatmap for the image on (a).

**Figure 6.7:** (a) A Ki67 stained image with a great mixture of positive (red) and negative (green) cells and (b) the calculated heatmap. The hotspots are clearly seen on the heatmap but hard to identify on (a).

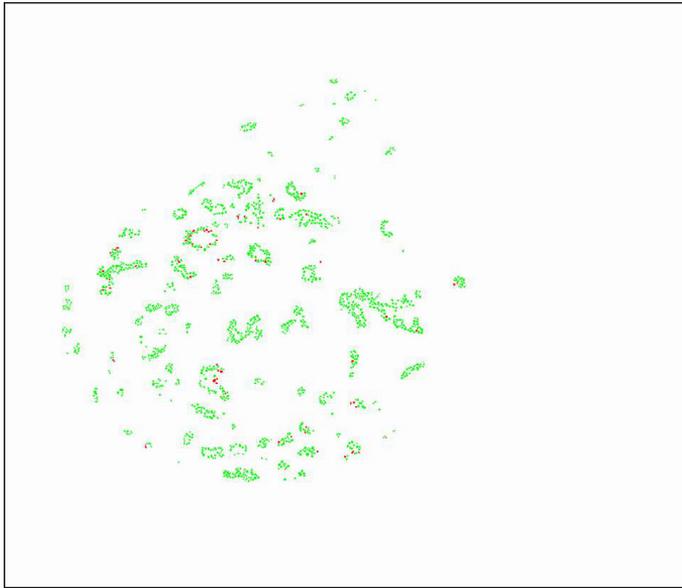


(a) Image with positive (red) and negative (green) Ki67 stained cells.

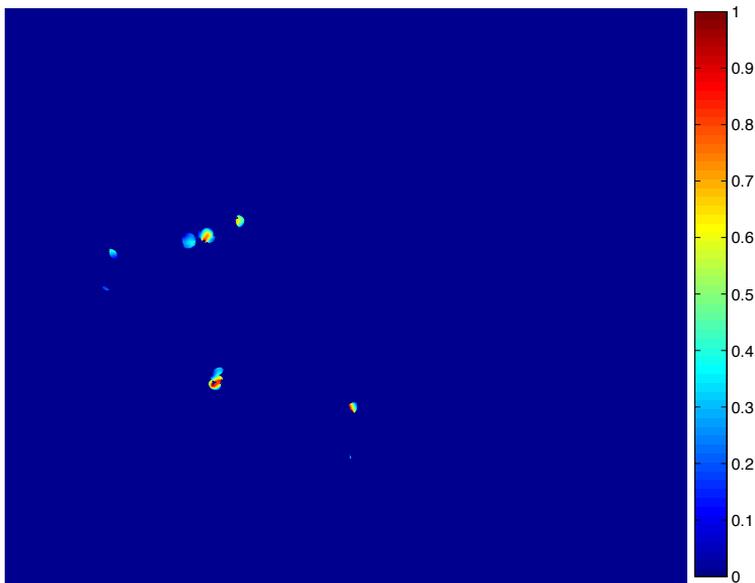


(b) Calculated heatmap for the image on (a).

**Figure 6.8:** (a) A Ki67 stained image with large areas with only positive (red) or negative (green) cells and (b) the calculated heatmap. The hotspots are obvious on the cell image but the cell densities are hard to incorporate in a manual selection of the hotspot. The heatmap clearly shows areas with high density of positive cells and eliminates areas with only negative cells or areas with many negative cells densely packed and few positives.



(a) Image with a few positive (red) and many negative (green) Ki67 stained cells.



(b) Calculated heatmap for the image on (a).

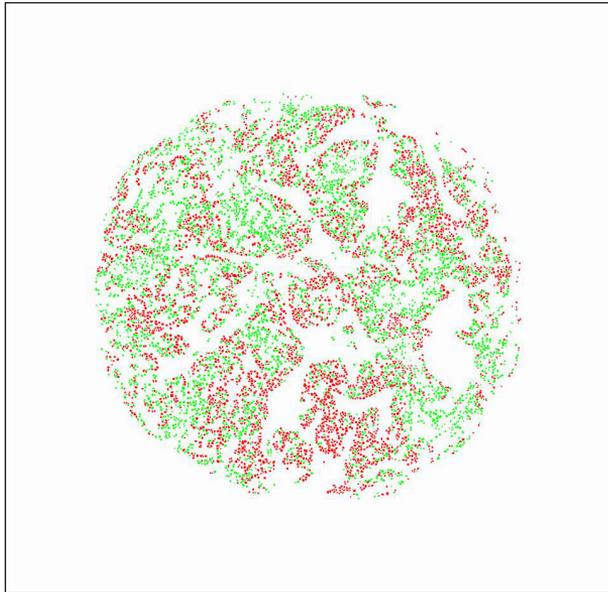
**Figure 6.9:** (a) A calculated heatmap for a Ki67 stained image (b) with few tumor cells and almost all negative (green) to the Ki67 stain. Areas with positive cells are easy to identify, but the density of positive cells versus the density of negative cells can be hard to measure by eye and the heatmap clearly points out areas with more positive cells than negative in the few areas where positive cells occur.

### 6.3.2 ER Heatmaps

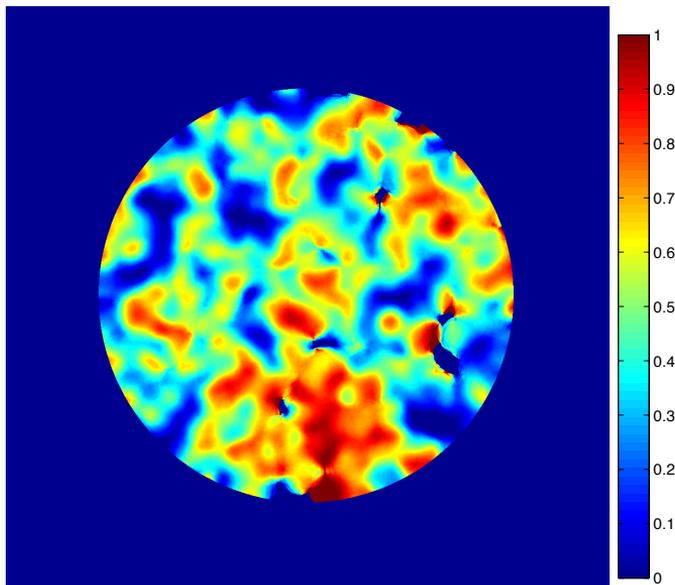
Figure 6.10 shows the ER image with a great mixture of positive and negative cells and the calculated heatmap. The hotspots are easily seen on the heatmap and the cell densities are taken into account compared to just finding the hotspot on the cell image.

Figure 6.11 shows an ER image with almost only positive cells and the calculated heatmap. The hotspot detection in this kind of image doesn't make sense and when the hotspot size is larger than the fixed size FOV, the score is just calculated at a random place.

Figure 6.12 shows an ER image with one positive cell and many negative cells together with the calculated heatmap. Hotspot detection might influence the score a bit due to the low cell density of the negative cells but the conclusion will be a pure negative ER image. Most ER images and their heatmap are like the example on this image and the one on figure 6.11.

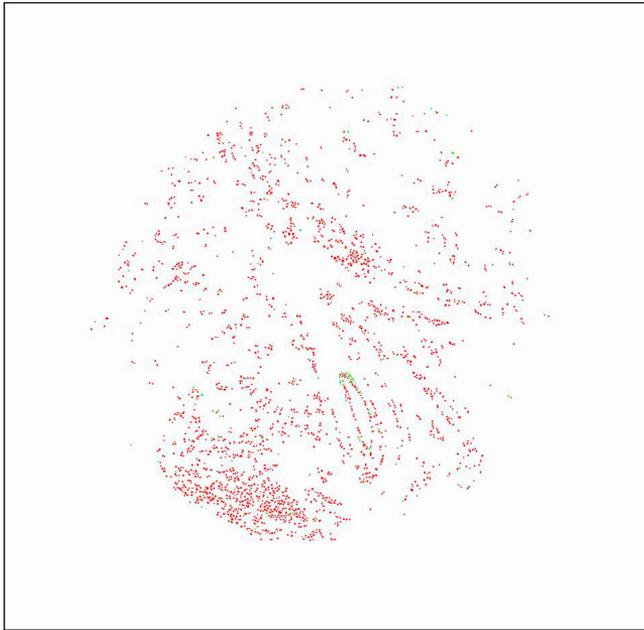


(a) Image with positive (red) and negative (green) ER stained cells.

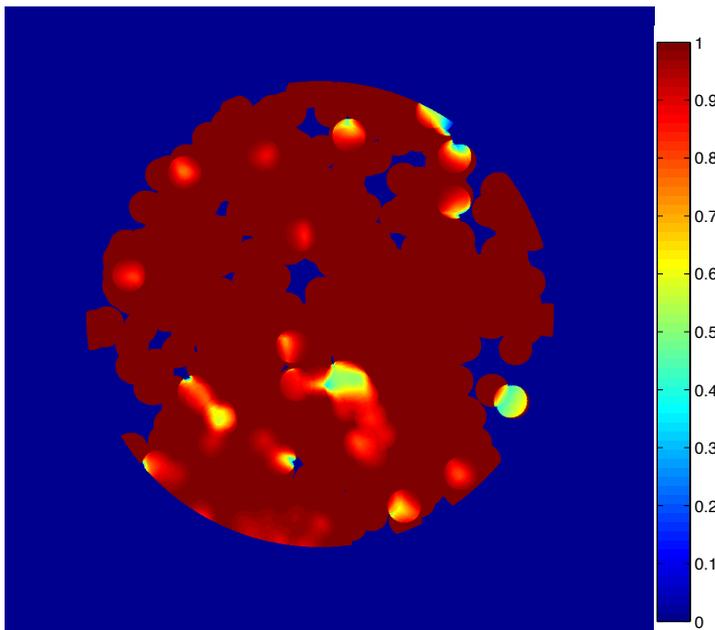


(b) Calculated heatmap for the image on (a).

**Figure 6.10:** (a) An untypical ER stained image with a great mixture of positive (red) and negative (green) cells and (b) its heatmap. Most ER stained images have either positive cells or negative cells. In this case a mixture of positive and negative cells exists and the hotspot is easy to identify. The heatmap can nevertheless be used for the exact finding of the hottest hotspot and not just a hotspot.

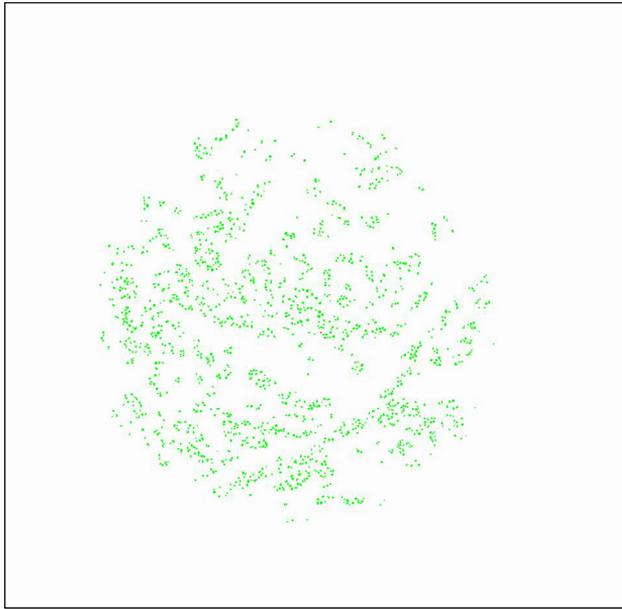


(a) Image with positive (red) and negative (green) ER stained cells.

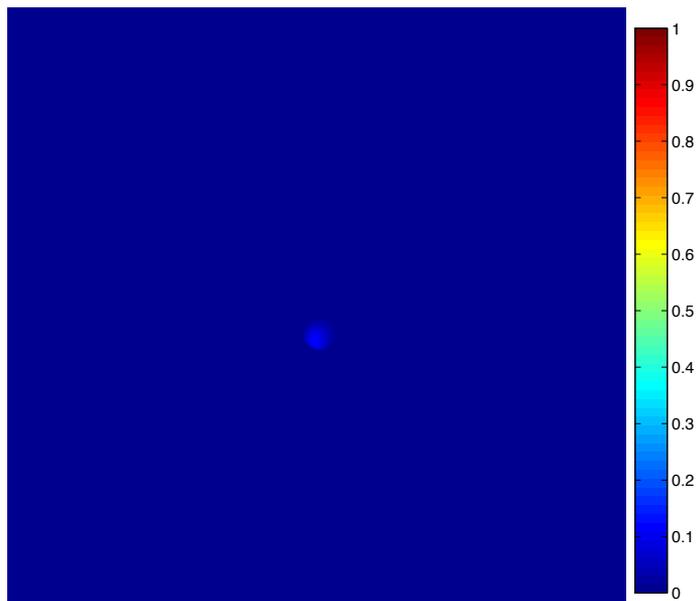


(b) Calculated heatmap for the image on (a).

**Figure 6.11:** (a) A typical positive ER stained image with almost only positive (red) cells and a few negative (green) cells and (b) its heatmap. A manual selection of hotspot is easy and the heatmap can be used to support the choice.



(a) Image with one positive (red) and many negative (green) ER stained cells.



(b) Calculated heatmap for the image on (a).

**Figure 6.12:** (a) A typical negative ER stained image with almost only negative (green) cells and only one positive (red) cell and (b) its heatmap. The positive cell is hard to identify on the image and the heatmap therefore nicely illustrate its presence. The impact on the total score is not substantial and the heatmap can only be used as an illustration.

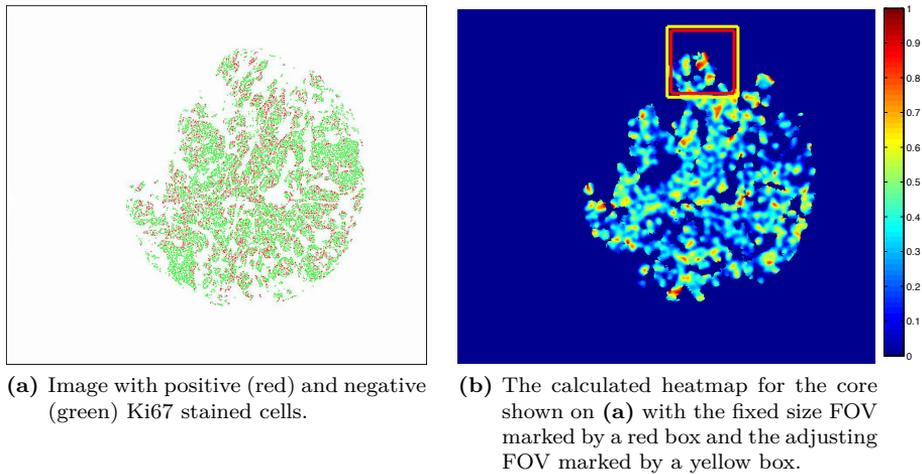
## 6.4 Hotspot detection

This section contains results from the hottest hotspot detection. The results present are chosen heatmaps and their adjusting FOV placed in the hottest hotspot together with the calculated score and the pathologist score. The results are mainly for the Ki67 stained images since the heatmaps from the ER images are either positive or negative and the hotspot detection therefore doesn't make much sense.

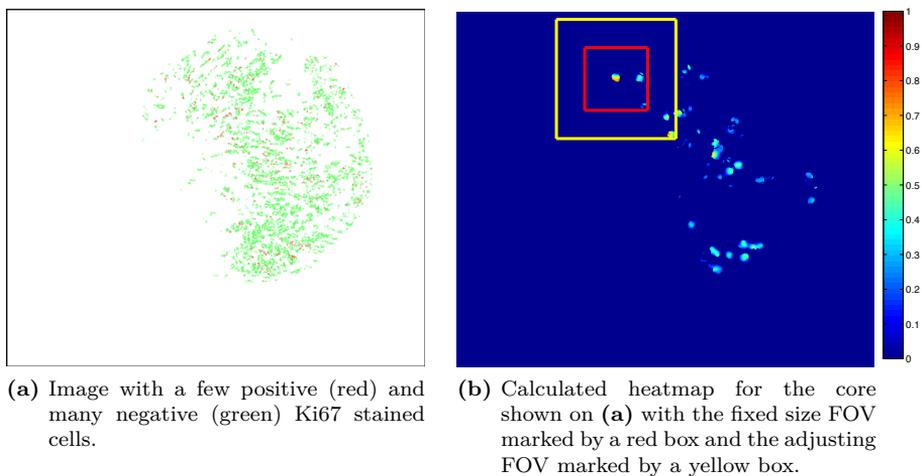
### 6.4.1 Ki67 Heatmaps with fixed size FOV and adjusting FOV

Figure 6.13 shows the image and the calculated heatmap with the fixed size FOV (red box) and the adjusting FOV (yellow box). The fixed size FOV contains 345 cells and the score is 34.2 % while the adjusting FOV contains 400 cells and the score is 34.25 %. The score provided by the pathologist is 17.3. The cell density in the core seems uniform and the explanation of the difference between the calculated scores and the pathologist score may be the manual selection of the hotspot by the pathologist.

Figure 6.14 shows the image and the calculated heatmap with the fixed size FOV (red box) and the adjusting FOV (yellow box) together. The image only contains a few positive cells. The fixed size FOV contains 156 cells and the score is 14.1 while the adjusting FOV contains 403 cells and the score is 11.56. The score provided by the pathologist is 8.24. The explanation of the deviation in the scores might be explained by the variation in the density of the negative cells.



**Figure 6.13:** (a) Ki67 stained TMA core with positive (red) and negative (green) cells. (b) The calculated heatmap for the TMA core on (a) with the fixed size FOV marked by a red box and the adjusting FOV marked by a yellow box.



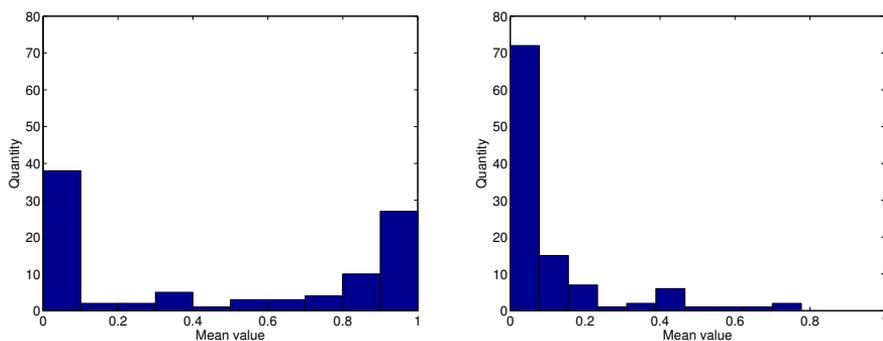
**Figure 6.14:** (a) Ki67 stained TMA core with a few positive (red) and many negative cells (green). (b) The calculated heatmap for the TMA core on (a) with the fixed size FOV marked by a red box and the adjusting FOV marked by a yellow box.

## 6.5 Mean, variance and histogram for heatmaps

This section contains the mean value, variance and some histograms for the calculated heatmaps. The mean and variance are calculated inside the tumor tissue as described in section 5.5.5.

Figure 6.15 shows the distribution of the mean value in the heatmaps for the ER stained and Ki67 stained images respectively.

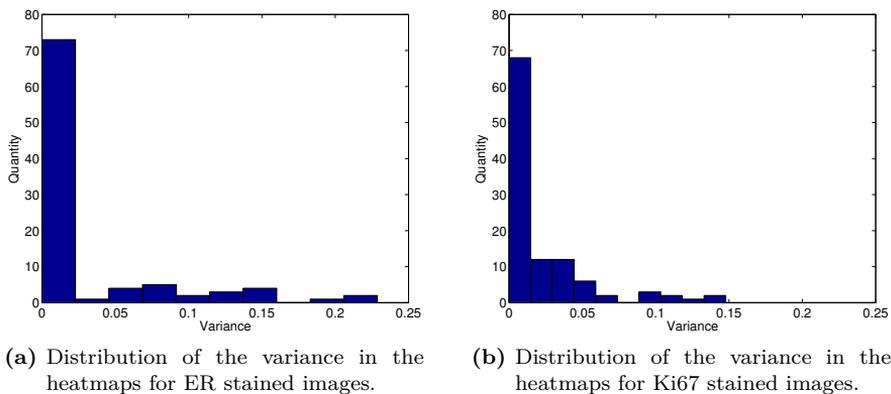
Figure 6.16 shows the distribution of the variance in the heatmaps for the ER stained and Ki67 stained images respectively.



(a) Distribution of the mean values in the heatmaps for ER stained images. The mean value is either high or low which matches the fact that almost all ER stained images are either 100 % or negative.

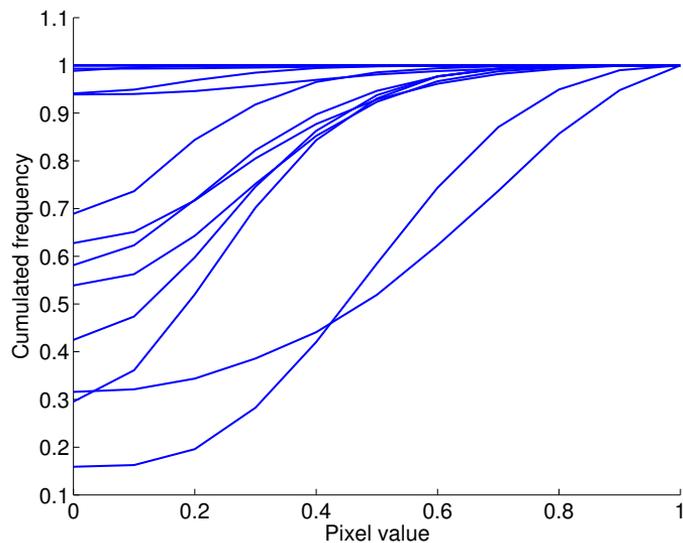
(b) Distribution of the mean values in the heatmaps for Ki67 stained images.

**Figure 6.15:** (a) Distribution of mean values in the heatmaps for the ER stained images and (b) distribution of mean values in the heatmaps for the Ki67 stained images.

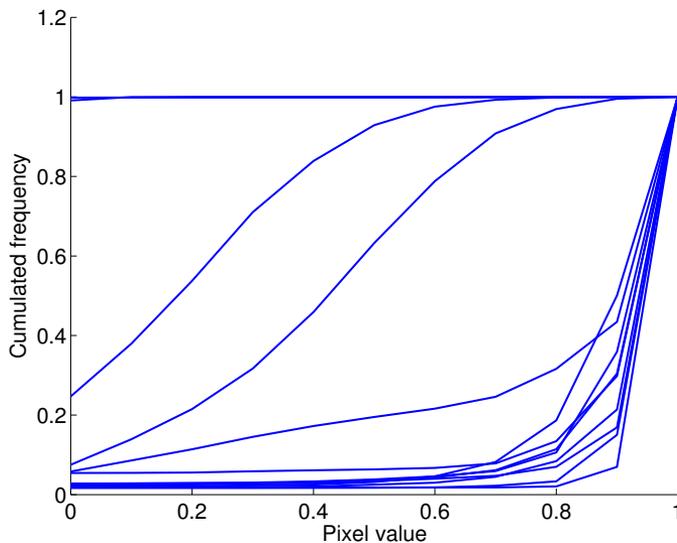


**Figure 6.16:** (a) Distribution of variance in the heatmaps for the ER stained images and (b) distribution of variance in the heatmaps for the Ki67 stained images.

The cumulated histograms for all heatmaps were also calculated to get an idea of the distribution of values in the individual heatmap. Some of the cumulated histograms for Ki67 stained images can be seen on figure 6.17 and for ER stained images on figure 6.18.



**Figure 6.17:** Cumulated histograms for some heatmaps based on Ki67 stained images.



**Figure 6.18:** Cumulated histograms for some heatmaps based on ER stained images.

As it can be seen from both the distribution of the mean values and the cumulated histograms for the heatmaps, almost all ER stained images are either 100% positive or negative. The mean values and the cumulated histograms for the Ki67 stained images' heatmaps show that some heatmaps contain few hotspots and many areas with only negatives and some heatmaps is a great mixture.

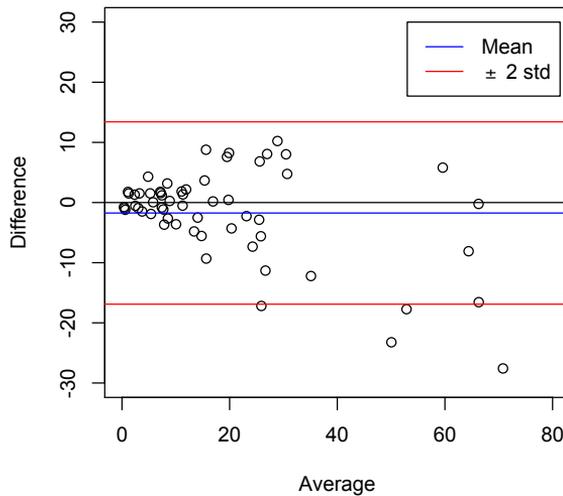
## 6.6 Ki67 scores from hottest hotspot

This section contains the results for the calculated Ki67 scores in the hottest hotspot inside a FOV of adjusting size containing between 400 and 450 cells. The correlation between the calculated scores and the pathologist scores is investigated and discussed using linear regression, statistical tests and other plots for visualization.

### 6.6.1 Bland Altman plot

Bland Altman plot is often used when two methods are compared [23]. In a Bland Altman plot the difference is plotted against the average of the scores from two methods.

Figure 6.19 shows the Bland Altman plot for the pathologist scores and the calculated Ki67 scores.



**Figure 6.19:** Bland Altman plot for the Ki67 scores found with the adjusting FOV. The difference of the pathologist scores and the calculated scores is plotted against the average of them. The mean and  $\pm 2$  standard deviations is marked by a blue and two red lines. The black line indicates a difference of 0. It is clear to see that the absolute difference increases with increasing scores.

The increasing variance for increasing scores is discussed in section 7.4.1. The next section contains the fitted regression model where the increasing variance also is taken into account.

## 6.6.2 Linear regression models

This section contains the fitted regression model used between the pathologist scores and the calculated Ki67 scores. The model is used to investigate if the scores are different from each other.

To find the best suitable regression model a variety of models were fitted to the data. The model used is presented in this section and the other fitted models are presented in the appendix C and appendix D.

An overview of the fitted models and where they can be found is listed in table 6.1.

Model	section/appendix
$y = \beta x + \alpha$	appendix C
$y = \beta x$	appendix C
$\text{Log}(y) = \beta \text{Log}(x) + \alpha$	appendix D
$\sqrt{y} = \beta \sqrt{x} + \alpha$	section 6.6.2.2
$\sqrt{y} = \beta \sqrt{x}$	section 6.6.2.2

**Table 6.1:** Overview of fitted regression models for the Ki67 scores and where the description and plots are located.

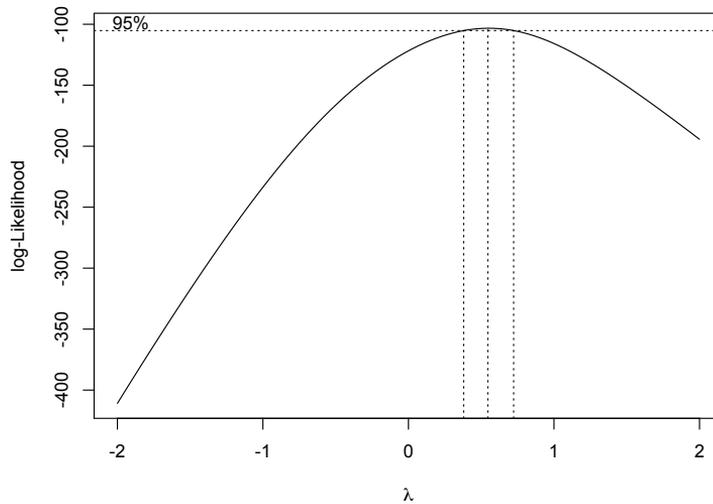
In the model  $y = \beta x + \alpha$ , it was found that the intercept ( $\alpha$ ) wasn't significant for the model and therefore the model  $y = \beta x$  was fitted to the data. The variance was increasing with increasing scores and therefore a Log transformation was performed.

The variance was decreased for large scores but increased for small scores in the model and therefore a Box Cox transformation was performed to find the optimal transformation of the data to make the variance constant and the residuals normally distributed.

The Box Cox transformation is described in the next section and the transformed and final model used is presented in the following section.

### 6.6.2.1 Box Cox transformation

A Box Cox transformation based on the model  $y = \beta x$  was performed. The Log-Likelihood as a function of  $\lambda$  was estimated and can be seen on figure 6.25.

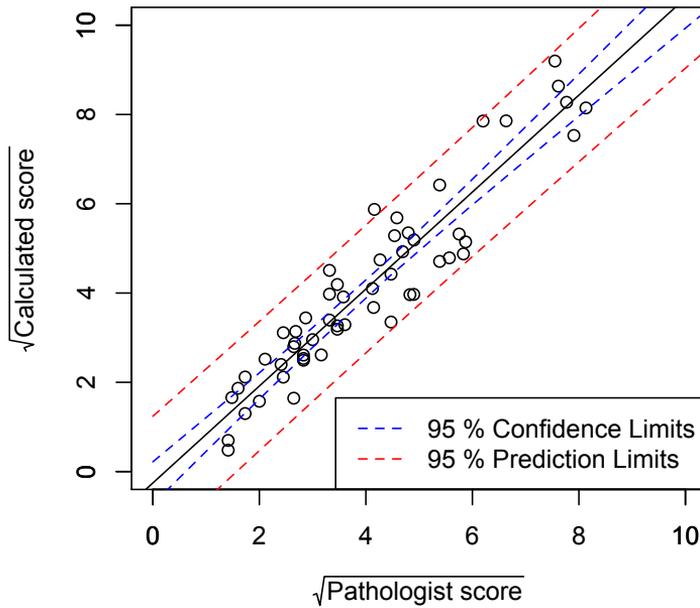


**Figure 6.20:** The Log-Likelihood as a function of  $\lambda$  for the model  $y = \beta x$ , where  $y$  is the calculated scores and  $x$  is the pathologist scores. The 95 % confidence interval is marked by dotted lines and the suitable  $\lambda$  for the transformation lies in this interval.  $\lambda = 0.5$  is chosen.

The chosen  $\lambda$  is 0.5 which is the most preferable transformation and it lies in the 95 % confidence interval (marked by dotted lines on the Log-Likelihood plot). The transformation with  $\lambda = 0.5$  is the same as taken the square root of the data. The fitted model is described in the following section.

### 6.6.2.2 Model: $\sqrt{y}=\beta\sqrt{x} + \alpha$ and $\sqrt{y}=\beta\sqrt{x}$

The regression model  $\sqrt{y}=\beta\sqrt{x} + \alpha$  and its confidence bounds can be seen on figure 6.26.



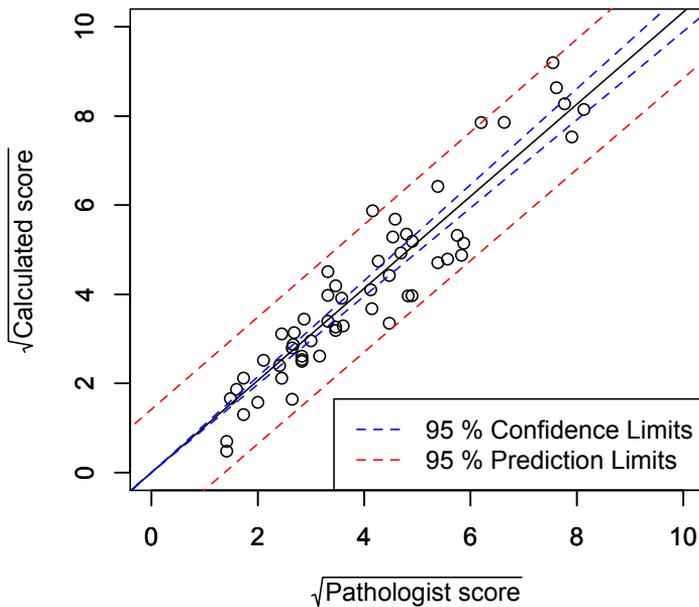
**Figure 6.21:** Plot of the regression model  $\sqrt{y}=\beta\sqrt{x} + \alpha$ .

The results of the two t-test of the model's coefficients can be found in table 6.2 and as it can be seen the intercept isn't significant for the model.

	Estimate	Std. Error	t value	Pr(> t )
<b>Model:</b> $\sqrt{y}=\beta\sqrt{x}+\alpha$				
Intercept ( $\alpha$ )	-0.255	0.236	-1.081	0.285
Slope ( $\beta$ )	1.0862	0.0538	20.181	<2e-16
<b>Model:</b> $\sqrt{y}=\beta\sqrt{x}$				
Slope ( $\beta$ )	1.0330	0.0217	47.52	<2e-16

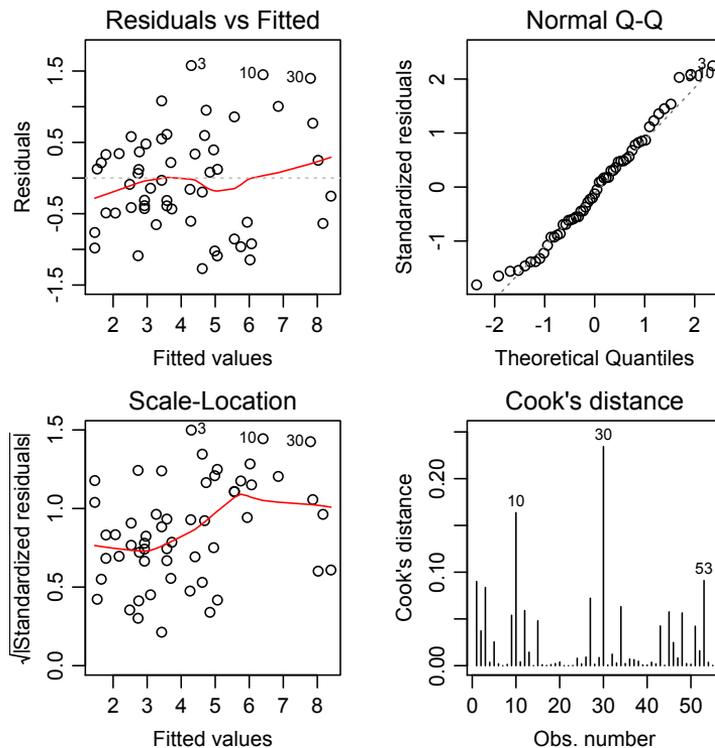
**Table 6.2:** t-test for  $\alpha$ 's and  $\beta$ 's significance in the two regression models:  $\sqrt{y}=\beta\sqrt{x}+\alpha$  and  $\sqrt{y}=\beta\sqrt{x}$ .

Therefore a new model of the form  $\sqrt{y}=\beta\sqrt{x}$  was fitted and can be seen on figure 6.27. The test value of  $\beta$  in the reduced model is also given in table 6.2.



**Figure 6.22:** Plot of the regression model  $\sqrt{y}=\beta\sqrt{x}$ .

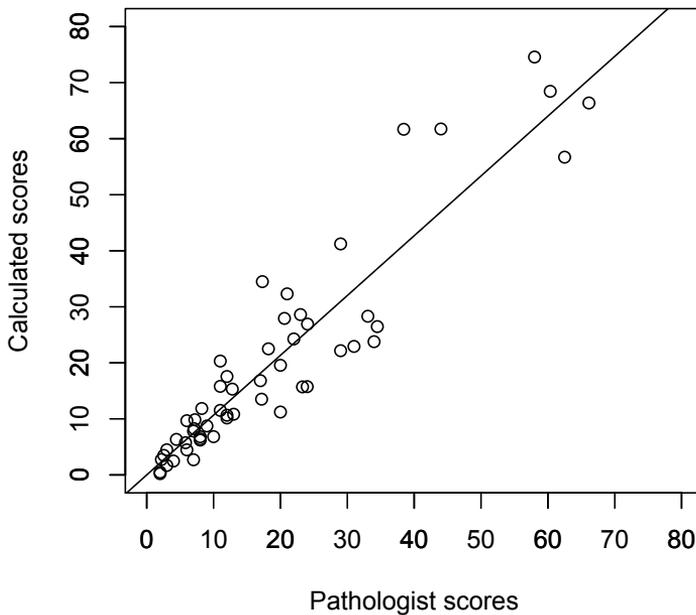
The diagnostic plot for the model can be seen on figure 6.28.



**Figure 6.23:** Diagnostic plot for the regression model  $\sqrt{y} = \beta\sqrt{x}$ .

The diagnostic plot shows that the variance is constant and the residuals are normally distributed which was the intention with the transformation of the data.

The model plotted on the original scale can be seen on figure 6.29.



**Figure 6.24:** Plot of the regression model  $\sqrt{y} = \beta\sqrt{x}$  on the original scale. The model is given as  $y = 1.067x$ .

The model on the original scale is  $y = 1.067x$ . This means that the calculated scores generally is around 7 % higher than the pathologist scores.

### 6.6.2.3 Linear hypothesis test: Scores calculated in the hottest hotspot compared with the pathologist scores

To test if the found model and thereby the method for calculating the Ki67 scores in the hottest hotspot is significantly different from the pathologist scores a linear hypothesis test (F-test) with the hypothesis

$$H_0: \beta = 1$$

$$H_1: \beta \neq 1$$

is performed. The calculated F-value and the p-value can be found in table 6.3.

---

Res.Df	RSS	Df	Sum of Sq	F	Pr(>F)
54	27.048	1	1.1521	2.3	0.1352

---

**Table 6.3:** Test of significant difference between model  $\sqrt{y} = 1.0330\sqrt{x}$  and  $y = x$  with test value  $\alpha = 0.05$ .

The p-value is higher than 0.05 which means that no significant difference between the estimated model and  $y = x$  can be proved.

This means that the score calculated in the hottest hotspot for Ki67 stained images can't be proved significantly different from the pathologist scores.

But what if the scores were calculated outside the hottest hotspot? The next section contains a similar analysis with scores measured randomly outside the hottest hotspot to investigate and perhaps give an idea of the meaning of the hottest hotspot.

## 6.7 Ki67 scores outside the hottest hotspot

To investigate and understand the use of the hottest hotspot to calculate the score in, the Ki67 score was calculate in the same heatmaps inside an adjusting FOV containing between 400 and 450 cells placed randomly outside the hottest hotspot.

### 6.7.1 Linear regression models

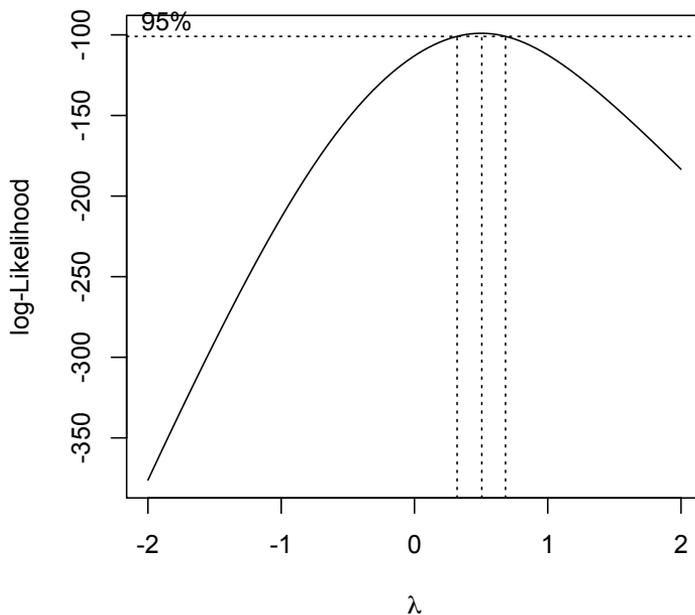
The procedure for fitting a linear regression model is exactly the same as in section 6.6.2. The fitted model used is presented in this section and the other fitted models and where they are found are listed in table 6.4.

Model	section/appendix
$y = \beta x + \alpha$	appendix E
$y = \beta x$	appendix E
$\text{Log}(y) = \beta \text{Log}(x) + \alpha$	appendix F
$\text{Log}(y) = \beta \text{Log}(x)$	appendix F
$\sqrt{y} = \beta \sqrt{x} + \alpha$	section 6.7.1.2
$\sqrt{y} = \beta \sqrt{x}$	section 6.7.1.2

**Table 6.4:** Overview of the fitted regression models for the Ki67 scores and where the description and plots are located.

### 6.7.1.1 Box Cox Transformation

The Box Cox transformation for the model  $y = \beta x$  was calculated. The used  $\lambda$  is again 0.5. The value lies in the confidence interval which can be seen on figure 6.25, where the Log-Likelihood as a function of  $\lambda$  is plotted.

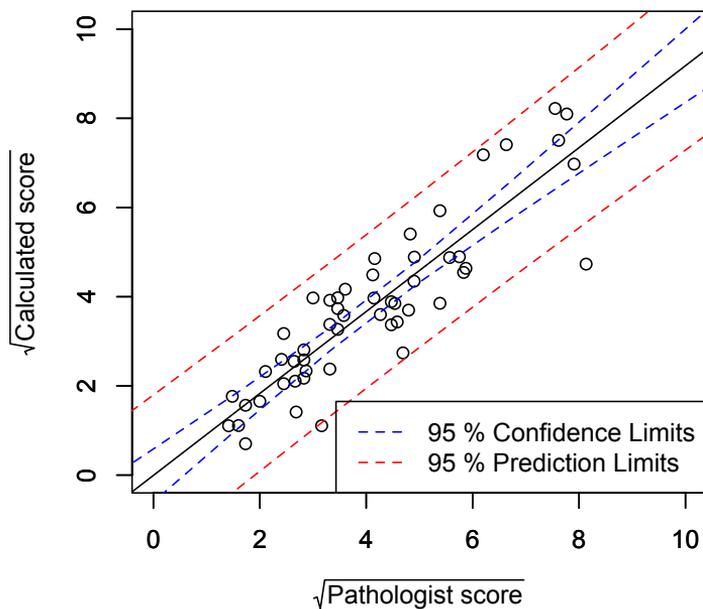


**Figure 6.25:** Log-likelihood as a function of  $\lambda$  to find a suitable  $\lambda$  for the transformation.

The fitted model with the Box Cox transformed data is presented in the next section.

### 6.7.1.2 Model: $\sqrt{y}=\alpha+\beta\sqrt{x}$

The fitted regression model  $\sqrt{y}=\alpha+\beta\sqrt{x}$  with the transformed data can be seen on figure 6.26. Table 6.5 contains the estimated coefficients and a t-test value for a test of their significance in the model. The intercept can be removed since it isn't significant. The model  $\sqrt{y}=\beta\sqrt{x}$  was fitted and the coefficient estimate is also presented in table 6.5.

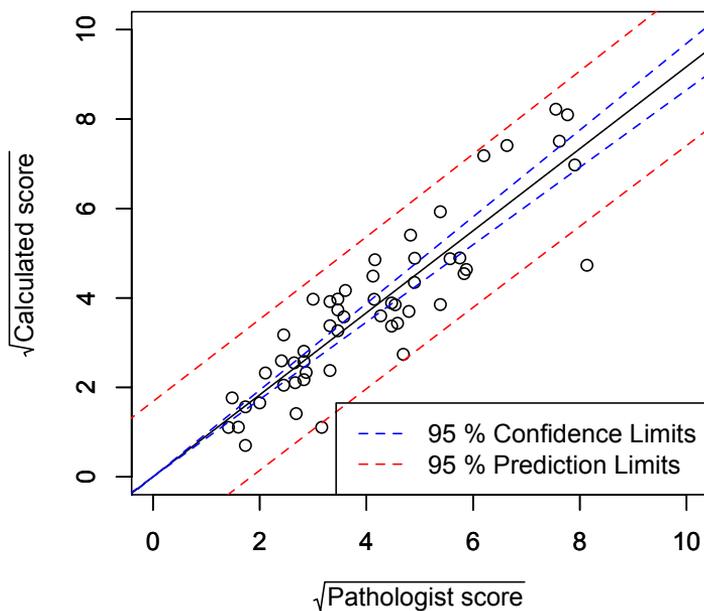


**Figure 6.26:** Plot of the regression model  $\sqrt{y}=\alpha+\beta\sqrt{x}$ .

	Estimate	Std. Error	t value	Pr(> t )
<b>Model:</b> $\sqrt{y}=\beta\sqrt{x}+\alpha$				
Intercept ( $\alpha$ )	-0.00717	0.296	-0.024	0.981
Slope ( $\beta$ )	0.919	0.0664	13.845	<2e-16
<b>Model:</b> $\sqrt{y}=\beta\sqrt{x}$				
Slope ( $\beta$ )	0.917	0.0259	35.39	<2e-16

**Table 6.5:** t-test for the estimated  $\alpha$ 's and  $\beta$ 's significance in the two regression models:  $\sqrt{y}=\beta\sqrt{x}+\alpha$  and  $\sqrt{y}=\beta\sqrt{x}$ .

Figure 6.27 shows the reduced fitted model and figure 6.28 contains some diagnostic plots for the model. The diagnostic plots shows that the residuals are normally distributed and the variance is more stable.



**Figure 6.27:** Plot of the regression model  $\sqrt{y}=\beta\sqrt{x}$  with 95 % confidence bounds.

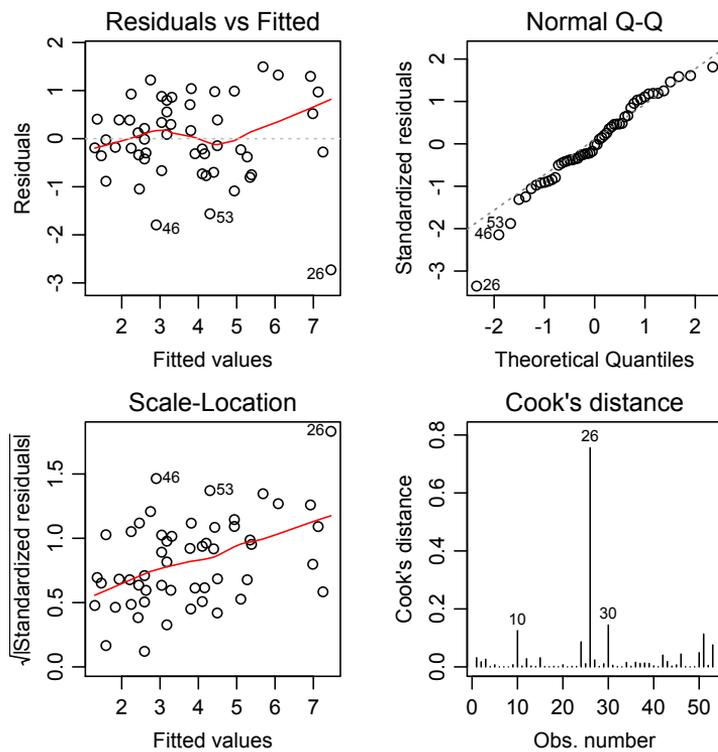
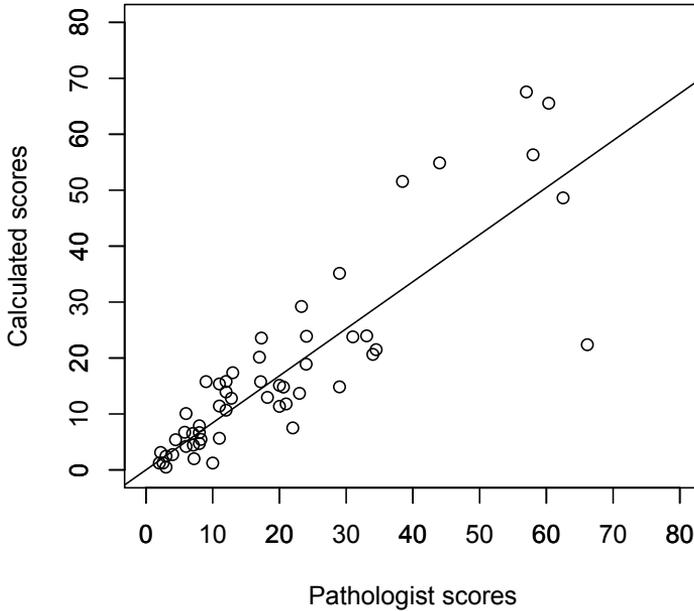


Figure 6.28: Diagnostic plots for the regression model  $\sqrt{y} = \beta\sqrt{x}$ .

Figure 6.29 shows the model on the original scale. The model on the original scale is  $y = 0.841x$ .



**Figure 6.29:** Plot of the regression model  $\sqrt{y} = \beta\sqrt{x}$  on the original scale. The model is  $y = 0.841x$ .

### 6.7.1.3 Linear Hypothesis Test: Scores calculated outside the hottest hotspot compared with the pathologist scores

The fitted model for the scores calculated randomly outside the hottest hotspot is tested to illustrate if a significant difference from the pathologist scores can be proved. This is done by a F-test with the hypothesis:

$$\begin{aligned} H_0: \beta &= 1 \\ H_1: \beta &\neq 1 \end{aligned}$$

The calculated F-value and the p-value can be found in table 6.6.

Res.Df	RSS	Df	Sum of Sq	F	Pr(>F)
52	36.697	1	7.2077	10.213	0.002372

**Table 6.6:** Test of significant difference between model  $\sqrt{y} = 0.917\sqrt{x}$  and  $y = x$  with test value  $\alpha = 0.05$ .

The p-value is below 0.05 and the  $H_0$  hypothesis is rejected. Thus the scores calculated randomly outside the hottest hotspot are significantly different from the pathologist scores.

#### 6.7.1.4 Linear Hypothesis Test: Scores calculated outside the hottest hotspot compared with scores calculated at the hottest hotspot

The fitted model for the scores calculated outside the hottest hotspot is compared to the model fitted for the scores calculated at the hottest hotspot by testing if the model is significantly different using an F-test. The hypothesis is:

$$\begin{aligned} H_0: \beta &= 1.0330 \\ H_1: \beta &\neq 1.0330 \end{aligned}$$

The calculated F-value and the p-value can be found in table 6.7.

Res.Df	RSS	Df	Sum of Sq	F	Pr(>F)
52	36.697	1	14.089	19.965	4.289e-05

**Table 6.7:** Test of significant difference between the models  $\sqrt{y} = 1.0330\sqrt{x}$  and  $\sqrt{y} = 0.917\sqrt{x}$ .

It can be seen that there is a significant difference between the two models which means that a significant difference between calculating the score in the hottest hotspot and another place than the hottest hotspot can be proved.

## 6.8 Cut off at 14 % for Ki67 scores

Ki67 scores above 14 % are Ki67 positive and thereby denoted aggressive cancers that will be treated with chemotherapy. The number of calculated scores and pathologist scores above 14 % was determined. A discussion of the results can be found in section 7.4.2.

26 out of 58 TMA cores have Ki67 score higher than 14 % for the pathologist scores.

The number of calculated scores above the cut off value is 28 out of 58. But is it the same TMA cores that are scored above the cut off value?

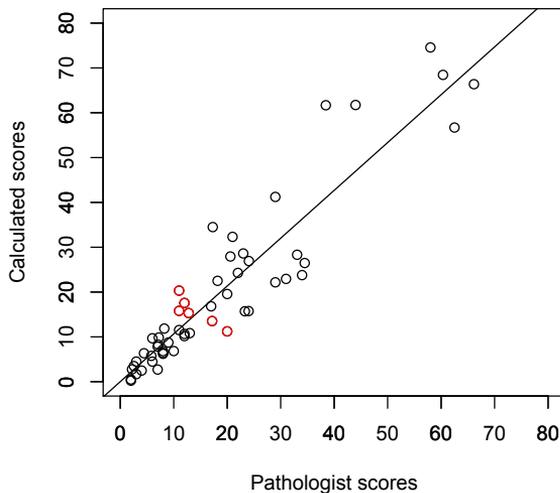
The calculated scores have 4 TMA cores scored above the cut off value that aren't scored above by the pathologist and the pathologist scores 2 TMA core above the cut off value that aren't scored above for the calculated scores.

Table 6.8 lists the scores for the 6 TMA cores that aren't classified as the same (either above or below the cut-off value) by the calculated score and the pathologist score.

Core coordinates	Score by pathologist	Score in hotspot	Deviation (%) between scores	Deviation (%) from cut-off 14% pathologist/hotspot
C7 <sub>TMA3</sub>	12.8	15.31	19.61	8.57/-9.36
E6 <sub>TMA3</sub>	12	17.56	46.33	14.30/-25.43
A10 <sub>TMA4</sub>	11	15.8	43.64	21.43/-12.86
D3 <sub>TMA4</sub>	11	20.32	84.73	21.43/-45.14
A6 <sub>TMA4</sub>	20	11.22	-78.25	-42.86/19.86
F7 <sub>TMA3</sub>	17.17	13.54	-26.81	-22.64/3.3

**Table 6.8:** Table of scores for the 6 TMA cores that aren't classified as the same by the pathologist and the calculated score. The first column lists the coordinates for the TMA core, the second column lists the score of the core provided by the pathologist, the third column lists the score calculated in the hottest hotspot, the fourth column lists the deviation between the two scores in % and the last column lists the deviation between the two scores and the cut-off value in %.

Figure 6.30 shows the calculated scores as a function of the pathologist scores together with the fitted regression model on the original scale. The outliers are marked by red.



**Figure 6.30:** The calculated Ki67 scores as a function of the pathologist scores together with the fitted regression model on the original scale. The outliers are marked by red.

The results are discussed in section 7.4.2.

## 6.9 Ki67 scores in the hottest, second, third and fourth hottest hotspot

The Ki67 score in the hottest, second, third and fourth hottest hotspot has been calculated to try to investigate the impact of the chosen hotspot and compare the scores with the pathologist scores.

The models have been fitted in the same way as the found models in section 6.6 and 6.7. The data has also been transformed using the Box Cox transformation, all with  $\lambda=0.5$ .

All four models have been tested for significant difference from the pathologist scores, with the hypothesis:

$$H_0: \beta = 1$$

$$H_1: \beta \neq 1$$

## 6.9 Ki67 scores in the hottest, second, third and fourth hottest hotspot 85

and the F-values and p-values can be found in table 6.9. The estimated models are also listed in the table and a plot of the model and a diagnostic plot for the second, third and fourth hottest hotspot can be found in appendix H. The model for the hottest hotspot is the same as the model presented in section 6.6.

	Res.Df	RSS	Df	Sum of Sq	F	Pr(>F)
<b>Hottest hotspot</b> $y = 1.067x$	54	27.048	1	1.152	2.3	0.135
<b>Second hottest hotspot</b> $y = 1.0235x$	54	29.353	1	0.145	0.267	0.608
<b>Third hottest hotspot</b> $y = 0.999x$	54	26.528	1	0.0296	0.0602	0.807
<b>Fourth hottest hotspot</b> $y = 0.998x$	54	29.456	1	0.00104	0.0019	0.965

**Table 6.9:** F- and p-values for the hypothesis testing of the four models' slopes different from 1. All tests show that non of the four models could be shown significantly different from the model  $y = x$ .

From the hypothesis test it was found that none of the models based on the scores calculated in the first, second, third and fourth hottest hotspot could be proved significantly different from the pathologist scores.

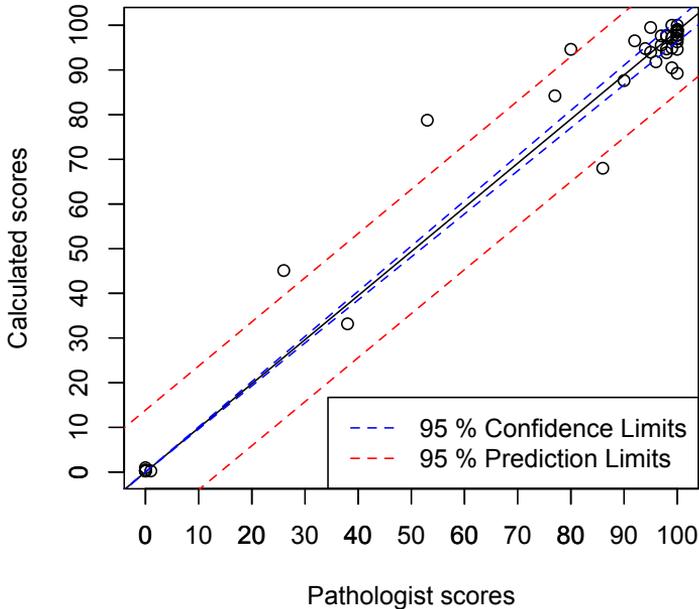
## 6.10 ER scores from hottest hotspot

To investigate whether the calculated ER scores were significantly different from the pathologist scores and to investigate the correlation between them a linear regression model was fitted. This section contains the fitted regression model used and a test of significant difference between the fitted model and the pathologist scores.

### 6.10.1 Model $y = \beta x$

The model  $y = \beta x$  is the fitted regression model used. The model  $y = \beta x + \alpha$  was fitted first and the intercept  $\alpha$  wasn't significant for the model. The model  $y = \beta x + \alpha$  and the table for the t-test of  $\alpha$  and  $\beta$  in the model can be found in appendix G.

Figure 6.31 shows the fitted regression model  $y = \beta x$ .



**Figure 6.31:** The fitted regression model  $y = \beta x$  with 95 % confidence bounds.

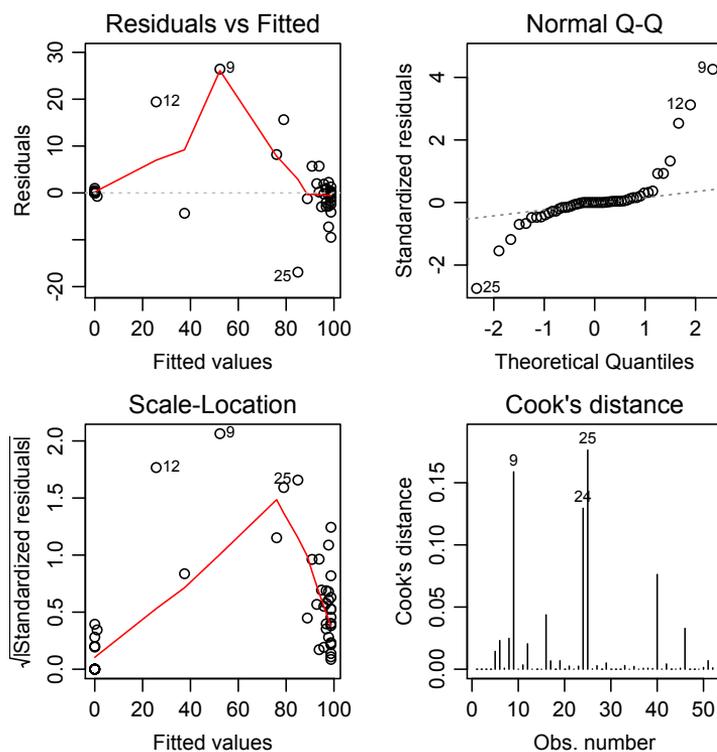
The estimated slope and the t-test of it can be found in table 6.11.

	Estimate	Std. Error	t value	$\Pr(> t )$
Slope ( $\beta$ )	0.98733	0.01206	81.83	$<2e-16$

**Table 6.10:** Test for  $\beta$ 's significant in the regression model.

Figure 6.32 shows some diagnostic plots for the fitted model. As it can be seen the residuals are high for the intermediate scores. The number of observations in this interval is lower than the number of observations in the ends. The reason is that ER scores often is 0 % or 100 % - either negative or positive. A few of the ER stained TMA cores in the data set have a mixture of negative and positive cells and the hotspot detection was used in these cores to calculate the score. A hotspot in the cores could easily be found manually but it might not be the hottest hotspot. Most of the calculated scores between 0 and 100 % that

lies outside the confidence boundaries are estimated higher than the pathologist score which might be explained by the use of the hottest hotspot. Only one score outside the confidence boundary is estimated lower than the pathologist score. The adjusting FOV is in this case increased to cover half of the core to contain above 400 cells and an area with high density of negative cells is included. The pathologist score might be based on significantly less cells than 400 due to the fixed size FOV used and the area with high density of negative cells will not be included.



**Figure 6.32:** Diagnostic plot for the model  $y = \beta x$ .

### 6.10.1.1 Linear Hypothesis Test: ER scores compared with the pathologist scores

To test if the found model and thereby the method for calculating the ER scores is significantly different from the pathologist scores a linear hypothesis test (F-test) with the hypothesis

$$\begin{aligned}H_0: \beta &= 1 \\H_1: \beta &\neq 1\end{aligned}$$

is performed. The calculated F-value and the p-value can be found in table 6.11.

Res.Df	RSS	Df	Sum of Sq	F	Pr(>F)
51	1979.3	1	52.013	1.3402	0.2524

**Table 6.11:** Test of significant difference between model  $y = 0.987x$  and  $y = x$  with test value  $\alpha = 0.05$ .



# Discussion

---

## 7.1 Amount of tumor tissue

This section contains a discussion of the amount of tumor tissue in the TMA cores as presented in section 6.2. If the amount of tumor tissue in a core is extremely low it might have an impact on the score since it will be based on few cells. Most of the TMA cores contained between 20-30 % tumor tissue and in general 10 % of the tumor tissue lied in the adjusted FOV. The tumor tissue might not contain many cells but over 70 % of all used FOV contained less than 15 % of the total amount of tumor tissue in the core and still contained a minimum of 400 cells. This indicates that no cores used contain too few cells and too low amount of tumor tissue that would affect the score.

## 7.2 Heatmaps and the spatial correlation

The heatmaps were calculated using a gauss weighted sliding window function. The size of the window and the standard deviation of the gauss function were estimated using the spatial correlation found by semivariogram analyses of the Ki67 and ER stained images. The found spatial correlation for the ER stained images was twice the spatial correlation for the Ki67 stained images.

The amount of tumor tissue in the cores could have an impact on the spatial correlation since small wide spread areas with few cells would have a larger

range than cores with large tumor areas. The distribution of tumor tissue was practically the same for the ER stained images and the Ki67 stained images, both having a large number of cores with 20-30 % tumor tissue.

The difference between the found standard deviation therefore might just be explained by the density of cells that are stained with the two different biomarkers. This was also the reason why two semivariogram analyses were performed instead of just one with all cores independent of biomarker.

### 7.3 Hotspot detection and FOV

The hotspot was detected as the largest area of pixels with value of the relative maximum in the heatmap. The found area was expanded using seed growing iteratively with 100 iterations and a threshold value of 60 % of the maximum value. A maximum number of iterations was set to avoid that the hotspot became too large and perhaps consisted of more than one hotspot. The reason why the seed growing was used, was to make sure that the neighborhood of the found largest area was taken into account when the FOV was placed at the center of the hotspot. The FOV was chosen to adjust in size depending on the number of cells and made sure that the FOV contained between 400 and 450 cells. The pathologist uses a FOV of fixed size. The amount of cells inside the fixed size FOV varied too much and the score therefore wouldn't be based on at least 400 cells. The size was tried varied but no size seemed to be able to fulfill the criteria of the 400 cells and the adjusting FOV was therefore chosen.

### 7.4 Quantative measured scores

This section contains a discussion of the quantative measured scores for the Ki67 and ER stained images.

#### 7.4.1 Increasing variance for increasing Ki67 scores

The Ki67 scores calculated in the hottest hotspot had a larger variance for larger scores. Several things can explain this

- **Densely packed cells in hotspot** - if the density of cells in the hotspot

is high the adjusting FOV will decrease in size compared to the FOV with defined size used by the pathologist. This might result in a higher estimate of the score if the FOV contains more positive cells, which depends on the hotspot size and the cell density and distribution around the hotspot.

- **Variance in the manual counting by the pathologist** - the manual count method used by the pathologist contains a high variance and the cells are only counted by one pathologist.
- **Few cells in the TMA core** - Some TMA cores have a low density of tumor cells and the fixed size FOV the pathologist uses can't cover 400 cells. The developed method expands the FOV to contain 400 cells if the TMA core in total contains more than 400 cells. The FOV might be expanded to include areas with higher density of negative cells. In this case the score will be lower.
- **Hottest hotspot not used** - One other explanation is simply that the pathologist didn't score in the hottest hotspot. The hotspot is manually selected by the pathologist and as found in [19] the hottest hotspot is only selected as the hottest hotspot in 50 % of the cases.

#### 7.4.2 Classification of patients based on Ki67 scores above and below cut off value

Section 6.8 contained the results for classification of patients above the cut off value of 14 % that will be treated with chemotherapy. 6 TMA cores weren't classified as the same based on the calculated scores and the pathologist scores. 4 of them weren't classified as above the cut off value by the pathologist but were classified as above based on the calculated score. 2 of them weren't classified as below the cut off value by the pathologist but was below in the calculated score.

A simple explanation why the 4 TMA cores weren't classified as above cut off by the pathologist and classified as above by the developed method is that the pathologist scores might not be calculated in the hottest hotspot and therefore the scores are lower. The size of the FOV might also have an impact on the score.

The two cores classified as above the cut off value by the pathologist and below by the method are cores with low density of cells across the whole core. If a fixed size of FOV is used the score will be based on an significantly lower number of cells than using the adjusted FOV as done in the developed method.

### 7.4.3 Evaluation of the score results and the fitted regression models

The calculated Ki67 scores inside the hottest hotspot were generally 7 % higher than the scores estimated by the pathologist. The pathologist scores were calculated in a manual selected hotspot and as found in [19] the hottest hotspot is only chosen in half of the cases. The explanation of why the pathologist scores are lower might simply be that the hottest hotspot wasn't chosen as the hotspot and the scores are thereby lower than the calculated scores where an automatic detection of the hottest hotspot was used.

It was also seen that the scores calculated in the second, third and fourth hottest hotspot generally were closer to the pathologist scores with slopes of 1.02, 0.99 and 0.99 for the fitted models. This might also indicate that the second, third or fourth hottest hotspot was used in the manual selection of the hottest hotspot.

The automatic detection of the hottest hotspot using the heatmap takes the cell density into account which is a thing that can be difficult to interpret at the manual selection.

The calculated ER scores had a few outliers which were the scores not being 0 and 100 %. The adjusting FOV was increased extremely for one core where the cell density was low in the whole core except for one area with many negative cells. The FOV was therefore increased in size and the large area of negative cells was included which made the score lower than the score provided by the pathologist.

The explanation for the overestimated scores could be the cell density in the images which made a manual hotspot detection of the hottest hotspot difficult and therefore the pathologist scores might not be calculated in the hottest hotspot.

## Conclusion

---

A visualization of the heterogeneity described by the response to the two biomarkers Ki67 and ER was performed in the form of a heatmap. The heatmap made it possible to get an idea of the individual tumor's population of cells and was used as a tool to find the hottest hotspot that should be used to calculate an individual response score.

An automatic hotspot detection in the heatmap was made and two types of FOV was used to calculate the score in. The FOV with adjusting size was chosen and several regression models were fitted to the scores for investigation of difference between the calculated scores and the pathologist scores.

No significant difference between the scores calculated in the hottest hotspot with the generated heatmaps used as guidance and the pathologist scores could be proven. The method might therefore be possible to use as an automatic method of measuring a quantitative measure of the tumor heterogeneity based on the nuclear population.



# List of Figures

---

2.1	Illustration of originate sites for Ductal Carcinoma In Situ (DCIS), Invasive Ductal Carcinoma (IDC) and Invasive Lobular Carcinoma (ILC). Image originates from [5]. . . . .	5
2.2	An example of a histology image stained with the biomarker Ki67. The brown cells are cells that reacts positive to the stain and therefore is Ki67 positive. The blue cells are Ki67 negative cells. The histology image is a TMA core. . . . .	7
2.3	An illustration of cell division for cancer cells. The circles represent cells and the symbols represent different kind of genes. For some divisions the cancer cells' genes change slightly which in the end causes different types of cancer cells resulting in cancer heterogeneity. . . . .	9
2.4	An illustration of the construction of a semivariogram. The variogram value as a function of the size of the displacement vector $h$ between points is shown. The bins are marked by red vertical lines and the red dot in each bin is the mean value. The mean value in the first bin is divided by 2 to minimize the nugget effect. To transform the values into a semivariogram the mean values are divided by 2. . . . .	12
2.5	Range, sill and nugget on a spherical model. . . . .	13

- 
- 3.1 An example of a virtual slide and the coordinate system. It contains 59 TMA cores stained with Ki67. Core E10 is missing. . . . 16
- 3.2 An ER stained core that is folded. The fold is marked with arrows. Cores that are folded, torn apart or stretched too much are removed since the alignment with PCK doesn't make sense then. . . . . 17
- 3.3 Two neighbor slices that aren't sufficient similar since the tumor regions found in the PCK stained core **(a)** doesn't match the cells areas in the ER stained core **(b)**. . . . . 18
- 5.1 **(a)** An ER stained TMA core with positive colored brown and **(b)** the neighbor slice stained with PCK where the epithelial tissue is stained dark brown. The two cores have been aligned. . 26
- 5.2 Examples of exported images from VIS with segmented cells, **(a)** an ER stained image with positive (red) cells and a few negative (green) cells and **(b)** a Ki67 stained image with a mixture of positive (red) and negative (green) cells. . . . . 29
- 5.3 Example of **(a)** a TMA core where all cells in the core are segmented and **(b)** the same TMA core where cells inside tumor regions have been segmented. . . . . 30
- 5.4 The same TMA core as on figure 5.3 with the fitted circle (red), the four used points (blue) and the center of the circle (pink). . . 31
- 5.5 Example of tumor mask with regions outside the TMA core. **(a)** the raw tumor mask, **(b)** the TMA core mask and **(c)** the tumor mask and the TMA core mask multiplied to eliminate small objects labeled as tumor tissue but lying outside the TMA core. 32
- 5.6 Visualization of the blob detection of the positive and the negative cells. **(a)** the image from VIS with segmented positive (red) and negative (green) cells. The red square is a zoom box used to get a better visualization of the small cells. **(b)** The image on **(a)** in the zoom box. **(c)** the positive cells and their center from the blob detection and **(d)** the negative cells and their center from the blob detection. . . . . 35

- 5.7 Zoom on image where 5000 cells are random sampled between 9000 cells in total. The random sampled cells are marked by a blue star and as it can be seen the sampling well represent the cells. Almost all images contain less than 10.000 cells. The estimated range for this image is 6.54 pixels. . . . . 37
- 5.8 An example of the sliding window function where a ROI of the image with positive cells on **(a)** is multiplied with the circular gauss filter on **(b)** and the remaining cells on **(c)** have values corresponding to their distance to the center of the circle. The values in **(c)** are summed and corresponds to the value for positive cells in that ROI. The same procedure is done for the negative cell image and the percentage of the positive cells is assigned to the center of the ROI. . . . . 39
- 5.9 The two distribution of range from the two semivariogram analyses. **(a)** the distribution of range from the semivariogram analysis of ER stained images and **(b)** the distribution of range from the semivariogram analysis of Ki67 stained images. . . . . 41
- 5.10 The two circular gauss weighted window functions used for the heatmaps for the Ki67 and ER stained images respectively. . . . 42
- 5.11 **(a)** Calculated heatmap for a Ki67 stained image and **(b)** the same heatmap but after multiplication with the tumor mask and TMA core mask done before calculation of mean value and variance. 44
- 5.12 Illustration of the process with detecting and dilating the hottest hotspot. **(a)** the heatmap, **(b)** relative maximum value on the heatmap marked by black, **(c)** result of threshold with the relative maximum value, **(d)** blob detection of the threshold image to find largest blob and **(e)** dilated hottest hotspot using seed growing with max 100 iterations. . . . . 48
- 5.13 The distribution of cells in the FOV for the two types of FOV, **(a)** the defined size FOV and **(b)** the adjusting FOV with lower limit of 400 cells and upper limit of 450 cells. . . . . 49
- 5.14 The two types of FOV, **(a)** a FOV of defined size and **(b)** a FOV with adjusting size until it contains minimum 400 cells. In this case the FOV is decreased a bit compared to the FOV of defined size. . . . . 50

6.1	(a) The semivariogram (red squares) with a fitted spherical model (blue line) for the Ki67 stained image on (b). . . . .	52
6.2	(a) The semivariogram (red squares) with a fitted spherical model (blue line) for the Ki67 stained image on (b). . . . .	52
6.3	(a) The semivariogram (red squares) with a fitted spherical model (blue line) for the ER stained image on (b). . . . .	53
6.4	(a) The semivariogram (red squares) with a fitted spherical model (blue line) for the ER stained image on (b). . . . .	53
6.5	Distribution of percentage tumor tissue inside TMA core for all stained cores. . . . .	54
6.6	Distribution of tumor tissue inside the FOV as percentage of the total amount of tumor tissue in the TMA core. . . . .	55
6.7	(a) A Ki67 stained image with a great mixture of positive (red) and negative (green) cells and (b) the calculated heatmap. The hotspots are clearly seen on the heatmap but hard to identify on (a). . . . .	57
6.8	(a) A Ki67 stained image with large areas with only positive (red) or negative (green) cells and (b) the calculated heatmap. The hotspots are obvious on the cell image but the cell densities are hard to incorporate in a manual selection of the hotspot. The heatmap clearly shows areas with high density of positive cells and eliminates areas with only negative cells or areas with many negative cells densely packed and few positives. . . . .	58
6.9	(a) A calculated heatmap for a Ki67 stained image (b) with few tumor cells and almost all negative (green) to the Ki67 stain. Areas with positive cells are easy to identify, but the density of positive cells versus the density of negative cells can be hard to measure by eye and the heatmap clearly points out areas with more positive cells than negative in the few areas where positive cells occur. . . . .	59

---

6.10	(a) An untypical ER stained image with a great mixture of positive (red) and negative (green) cells and (b) its heatmap. Most ER stained images have either positive cells or negative cells. In this case a mixture of positive and negative cells exists and the hotspot is easy to identify. The heatmap can nevertheless be used for the exact finding of the hottest hotspot and not just a hotspot. . . . .	61
6.11	(a) A typical positive ER stained image with almost only positive (red) cells and a few negative (green) cells and (b) its heatmap. A manual selection of hotspot is easy and the heatmap can be used to support the choice. . . . .	62
6.12	(a) A typical negative ER stained image with almost only negative (green) cells and only one positive (red) cell and (b) its heatmap. The positive cell is hard to identify on the image and the heatmap therefore nicely illustrates its presence. The impact on the total score is not substantial and the heatmap can only be used as an illustration. . . . .	63
6.13	(a) Ki67 stained TMA core with positive (red) and negative (green) cells. (b) The calculated heatmap for the TMA core on (a) with the fixed size FOV marked by a red box and the adjusting FOV marked by a yellow box. . . . .	65
6.14	(a) Ki67 stained TMA core with a few positive (red) and many negative cells (green). (b) The calculated heatmap for the TMA core on (a) with the fixed size FOV marked by a red box and the adjusting FOV marked by a yellow box. . . . .	65
6.15	(a) Distribution of mean values in the heatmaps for the ER stained images and (b) distribution of mean values in the heatmaps for the Ki67 stained images. . . . .	66
6.16	(a) Distribution of variance in the heatmaps for the ER stained images and (b) distribution of variance in the heatmaps for the Ki67 stained images. . . . .	67
6.17	Cumulated histograms for some heatmaps based on Ki67 stained images. . . . .	67
6.18	Cumulated histograms for some heatmaps based on ER stained images. . . . .	68

6.19	Bland Altman plot for the Ki67 scores found with the adjusting FOV. The difference of the pathologist scores and the calculated scores is plotted against the average of them. The mean and $\pm 2$ standard deviations is marked by a blue and two red lines. The black line indicates a difference of 0. It is clear to see that the absolute difference increases with increasing scores. . . . .	69
6.20	The Log-Likelihood as a function of $\lambda$ for the model $y = \beta x$ , where $y$ is the calculated scores and $x$ is the pathologist scores. The 95 % confidence interval is marked by dotted lines and the suitable $\lambda$ for the transformation lies in this interval. $\lambda = 0.5$ is chosen. . . . .	71
6.21	Plot of the regression model $\sqrt{y} = \beta\sqrt{x} + \alpha$ . . . . .	72
6.22	Plot of the regression model $\sqrt{y} = \beta\sqrt{x}$ . . . . .	73
6.23	Diagnostic plot for the regression model $\sqrt{y} = \beta\sqrt{x}$ . . . . .	74
6.24	Plot of the regression model $\sqrt{y} = \beta\sqrt{x}$ on the original scale. The model is given as $y = 1.067x$ . . . . .	75
6.25	Log-likelihood as a function of $\lambda$ to find a suitable $\lambda$ for the transformation. . . . .	77
6.26	Plot of the regression model $\sqrt{y} = \alpha + \beta\sqrt{x}$ . . . . .	78
6.27	Plot of the regression model $\sqrt{y} = \beta\sqrt{x}$ with 95 % confidence bounds. . . . .	79
6.28	Diagnostic plots for the regression model $\sqrt{y} = \beta\sqrt{x}$ . . . . .	80
6.29	Plot of the regression model $\sqrt{y} = \beta\sqrt{x}$ on the original scale. The model is $y = 0.841x$ . . . . .	81
6.30	The calculated Ki67 scores as a function of the pathologist scores together with the fitted regression model on the original scale. The outliers are marked by red. . . . .	84
6.31	The fitted regression model $y = \beta x$ with 95 % confidence bounds. . . . .	87
6.32	Diagnostic plot for the model $y = \beta x$ . . . . .	88

B.1	Distribution of percentage tumor tissue inside TMA core for Ki67 stained cores. . . . .	110
B.2	Distribution of percentage tumor tissue inside TMA core for ER stained cores. . . . .	110
C.1	The fitted regression model $y = \beta x + \alpha$ with 95 % confidence bounds. . . . .	112
C.2	The fitted regression model $y = \beta x$ with 95 % confidence bounds.	113
C.3	Diagnostic plot for the model: $y = \beta x$ . . . . .	114
D.1	Linear regression model $\text{Log}(y)=\beta\text{Log}(x)+\alpha$ . . . . .	116
D.2	Diagnostic plots for the model $\text{Log}(y)=\beta\text{Log}(x)+\alpha$ . . . . .	117
E.1	Linear regression model of form $y = \beta x + \alpha$ fitted to scores calculated outside the hottest hotspot as a function of the pathologist scores. . . . .	120
E.2	Linear regression model $y = \beta x$ fitted to scores calculated outside the hottest hotspot. . . . .	121
E.3	Diagnostic plot for the model $y = \beta x$ . . . . .	122
F.1	Linear regression model $\text{Log}(y)=\beta\text{Log}(x)+\alpha$ . . . . .	124
F.2	Linear regression model $\text{Log}(y)=\beta\text{Log}(x)$ . . . . .	125
F.3	Diagnostic plots for the linear regression model $\text{Log}(y)=\beta\text{Log}(x)+\alpha$ .	126
G.1	Linear regression model fitted to scores from method 1 as a function of the pathologist scores. . . . .	128
H.1	(a) the fitted model $\sqrt{y} = \beta\sqrt{x}$ and on (b) the fitted model transformed to the original scale . . . . .	130
H.2	Diagnostic plot for the model shown on H.1. . . . .	131

---

H.3	(a) the fitted model $\sqrt{y} = \beta\sqrt{x}$ and on (b) the fitted model transformed to the original scale . . . . .	132
H.4	Diagnostic plot for the model shown on H.3. . . . .	133
H.5	(a) the fitted model $\sqrt{y} = \beta\sqrt{x}$ and on (b) the fitted model transformed to the original scale . . . . .	134
H.6	Diagnostic plot for the model shown on H.5. . . . .	135

# List of Tables

---

5.1	The median of the ranges for Ki67 stained images and ER stained images from the semivariogram analyses. . . . .	40
6.1	Overview of fitted regression models for the Ki67 scores and where the description and plots are located. . . . .	70
6.2	t-test for $\alpha$ 's and $\beta$ 's significance in the two regression models: $\sqrt{y}=\beta\sqrt{x}+\alpha$ and $\sqrt{y}=\beta\sqrt{x}$ . . . . .	73
6.3	Test of significant difference between model $\sqrt{y} = 1.0330\sqrt{x}$ and $y = x$ with test value $\alpha = 0.05$ . . . . .	76
6.4	Overview of the fitted regression models for the Ki67 scores and where the description and plots are located. . . . .	77
6.5	t-test for the estimated $\alpha$ 's and $\beta$ 's significance in the two regression models: $\sqrt{y}=\beta\sqrt{x}+\alpha$ and $\sqrt{y}=\beta\sqrt{x}$ . . . . .	79
6.6	Test of significant difference between model $\sqrt{y} = 0.917\sqrt{x}$ and $y = x$ with test value $\alpha = 0.05$ . . . . .	82
6.7	Test of significant difference between the models $\sqrt{y} = 1.0330\sqrt{x}$ and $\sqrt{y} = 0.917\sqrt{x}$ . . . . .	82

6.8	Table of scores for the 6 TMA cores that aren't classified as the same by the pathologist and the calculated score. The first column lists the coordinates for the TMA core, the second column lists the score of the core provided by the pathologist, the third column lists the score calculated in the hottest hotspot, the fourth column lists the deviation between the two scores in % and the last column lists the deviation between the two scores and the cut-off value in %.	83
6.9	F- and p-values for the hypothesis testing of the four models' slopes different from 1. All tests show that non of the four models could be shown significantly different from the model $y = x$ .	85
6.10	Test for $\beta$ 's significant in the regression model.	87
6.11		89
C.1	Test for $\alpha$ and $\beta$ 's significant in the regression model $y = \beta x + \alpha$ .	112
C.2	Test for $\beta$ 's significant in the regression model $y = \beta x$ .	113
E.1	Test for $\alpha$ and $\beta$ 's significant in the regression model $y = \beta x + \alpha$ .	120
E.2	Test for $\alpha$ 's significant in the regression model $y = \beta x$ .	121
F.1	Test for $\alpha$ and $\beta$ 's significant in the regression model $\text{Log}(y) = \beta \text{Log}(x) + \alpha$	124
F.2	Test for $\alpha$ and $\beta$ 's significant in the regression model.	125
G.1	Test for $\alpha$ and $\beta$ 's significant in the regression model $y = \beta x + \alpha$ .	128

## APPENDIX A

# Semivariogram function by Allan Aabjerg Nielsen

---

---

```
function [sv]=semivariogram(X)

dist = pdist(X(:,1:2));
pc = pdist(X(:,3)).^2;
sv = nan(nlag+2,3);
idx = find(dist==0);
sv(1,:) = [size(idx,2) mean(dist(idx)) 0.5*mean(pc(idx))];
idx = find(0<dist & dist<=(lagdist/2));
sv(2,:) = [size(idx,2) mean(dist(idx)) 0.5*mean(pc(idx))];

for ii=3:(nlag+2)
idx = find((2*ii-5)*(lagdist/2)<dist & dist<=(2*ii-3)*
lagdist/2);
sv(ii,:) = [size(idx,2) mean(dist(idx)) 0.5*mean(pc(idx))];
end

%Plot semivariogram
figure; plot(sv(:,2),sv(:,3),'o')
end
```

---



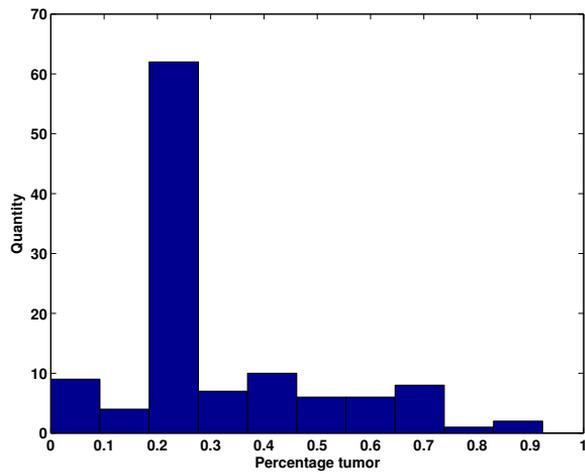
## APPENDIX B

# Amount of tumor tissue in ER and Ki67 stained images separately

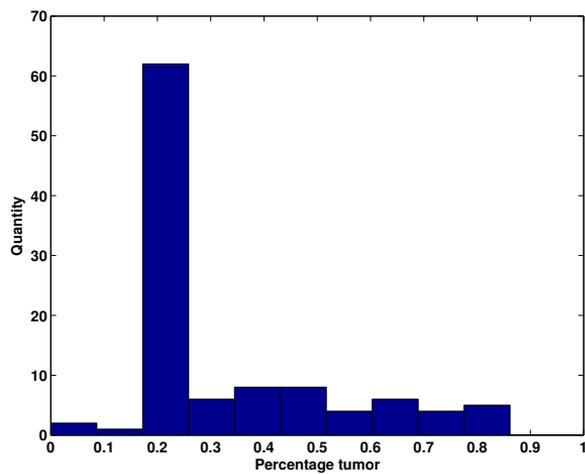
---

This appendix contains the distribution of the percentage of tumor tissue in all Ki67 stained cores and all ER stained cores separately.

Figure B.1 shows the distribution for the Ki67 stained cores and figure B.2 for the ER stained cores.



**Figure B.1:** Distribution of percentage tumor tissue inside TMA core for Ki67 stained cores.



**Figure B.2:** Distribution of percentage tumor tissue inside TMA core for ER stained cores.

## APPENDIX C

# Regression model for Ki67 scores from the hottest hotspot

---

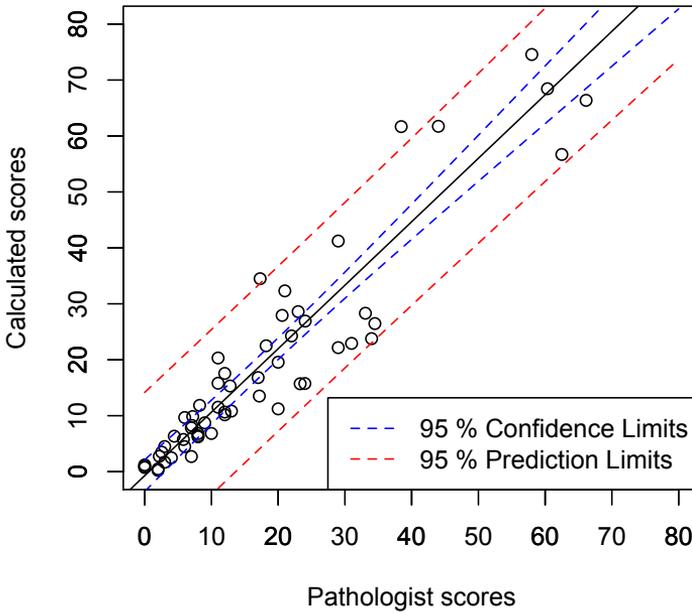
This appendix contains the first fitted regression models  $y = \beta x + \alpha$  and  $y = \beta x$  for the Ki67 scores calculated in the hottest hotspot and some diagnostic plots for the models.

Figure C.1 shows the fitted regression model  $y = \beta x + \alpha$  and table C.1 shows the test values of the slope and intercept of the model. Two t-tests, one for the intercept ( $\alpha$ ) and one for the slope ( $\beta$ ) of the model are performed to test for significance for the model. The t-test of the intercept  $\alpha$  with the hypothesis:

$$\begin{aligned} H_0: \alpha &= 0 \\ H_1: \alpha &\neq 0 \end{aligned}$$

and the t-test of the slope  $\beta$  with the hypothesis:

$$\begin{aligned} H_0: \beta &= 0 \\ H_1: \beta &\neq 0 \end{aligned}$$



**Figure C.1:** The fitted regression model  $y = \beta x + \alpha$  with 95 % confidence bounds.

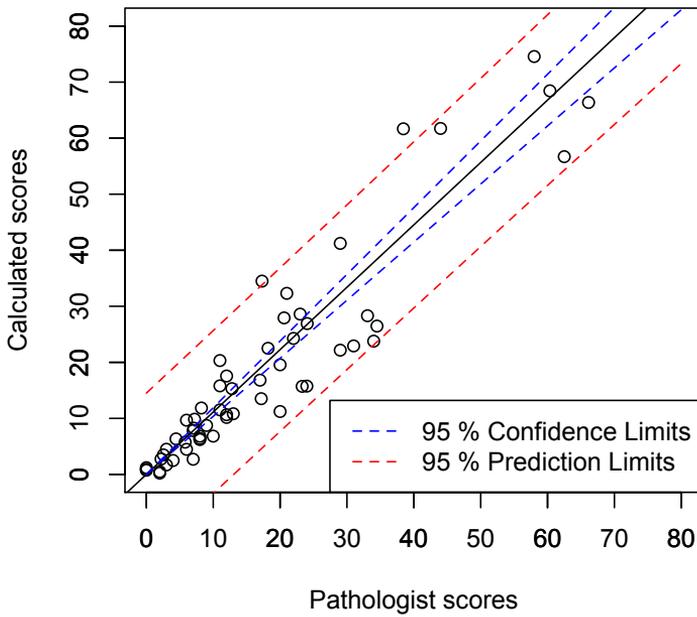
	Estimate	Std. Error	t value	Pr(> t )
Intercept ( $\alpha$ )	-0.93495	1.54675	-0.604	0.548
Slope ( $\beta$ )	1.14106	0.06082	18.760	<2e-16

**Table C.1:** Test for  $\alpha$  and  $\beta$ 's significant in the regression model  $y = \beta x + \alpha$ .

Since the intercept wasn't significant, the model was reduced to the model  $y = \beta x$ .

Figure C.2 shows the fitted model  $y = \beta x$  and figure C.3 shows some diagnostic plots for the model.

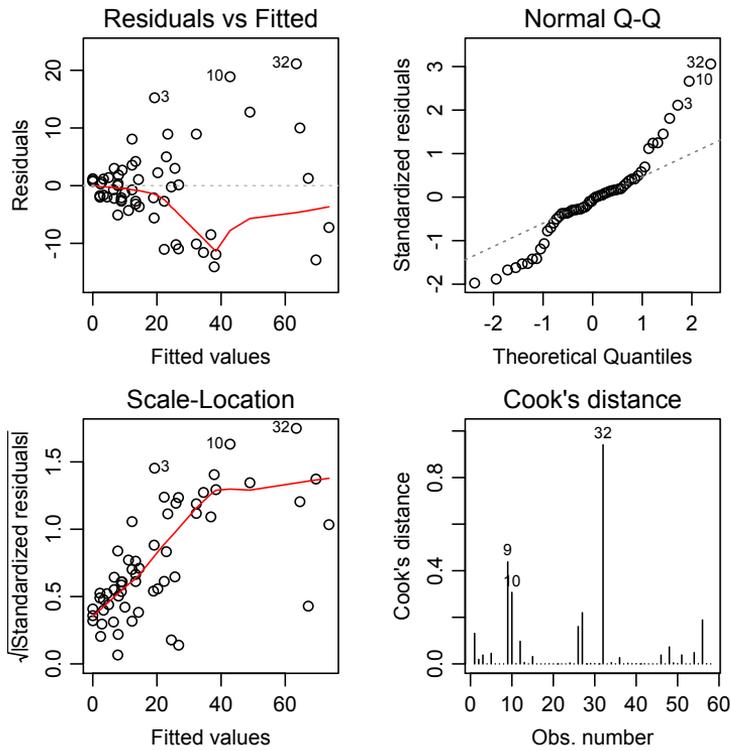
Table C.2 contain the test value for the t-test of the slope different from 0.



**Figure C.2:** The fitted regression model  $y = \beta x$  with 95 % confidence bounds.

	Estimate	Std. Error	t value	$\Pr(> t )$
Slope	1.11319	0.03842	28.98	$<2e-16$

**Table C.2:** Test for  $\beta$ 's significant in the regression model  $y = \beta x$ .



**Figure C.3:** Diagnostic plot for the model:  $y = \beta x$ .

As it can be seen from the QQ-plot tails exist in each end which indicates that the residuals aren't normally distributed and the residual plot also shows an increasing variance for higher score values. A transformation of the data was therefore necessary and the obvious choice was a Log transformation of the data, a regression model of the form  $\text{Log}(y) = \beta \text{Log}(x) + \alpha$ .

The fitted Log model and its diagnostic plot can be found in appendix D.

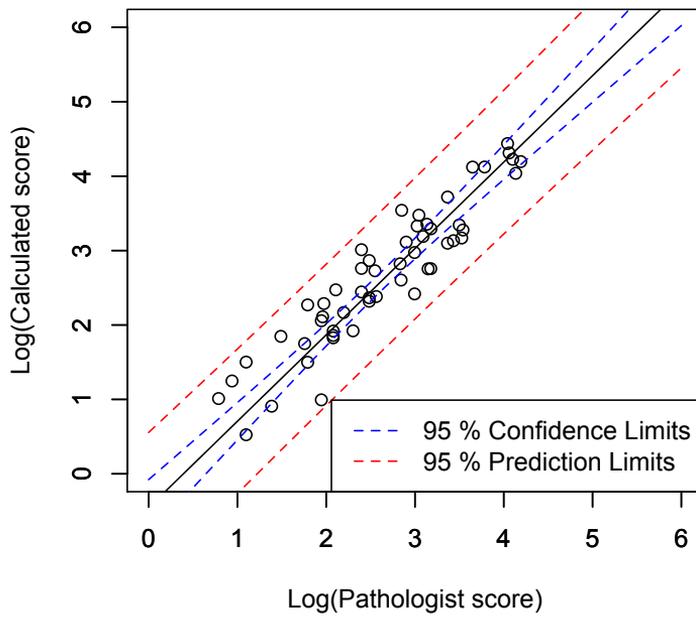
## APPENDIX D

# Log model for Ki67 scores from the hottest hotspot

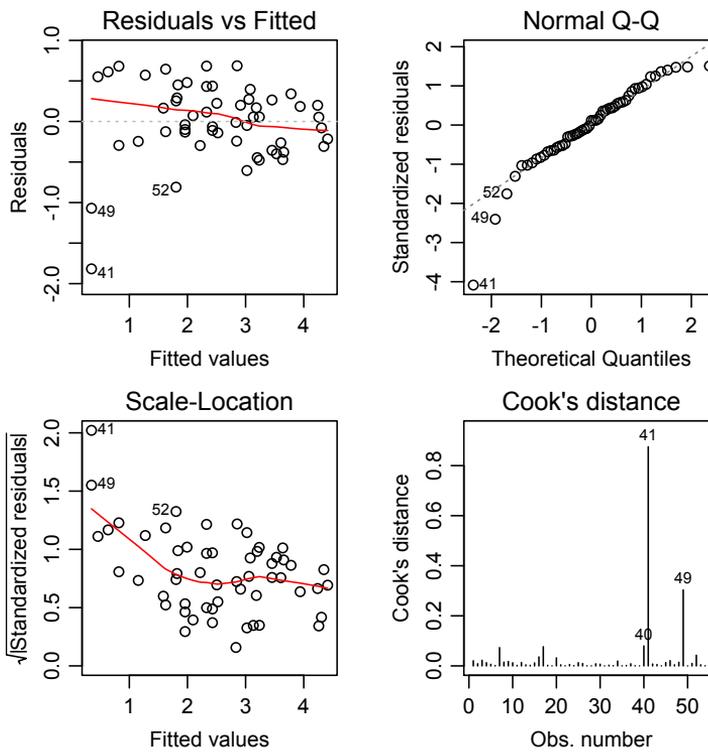
---

This appendix contain the Log transformed model for the Ki67 scores calculated at the hottest hotspot as a function of the pathologist scores.

Figure D.1 shows the fitted model and the transformed data and figure D.2 shows some diagnostic plots for the model.



**Figure D.1:** Linear regression model  $\text{Log}(y)=\beta\text{Log}(x)+\alpha$ .



**Figure D.2:** Diagnostic plots for the model  $\text{Log}(y) = \beta \text{Log}(x) + \alpha$ .

The increasing variance for higher scores are eliminated but the variance isn't constant and increasing for the low scores.



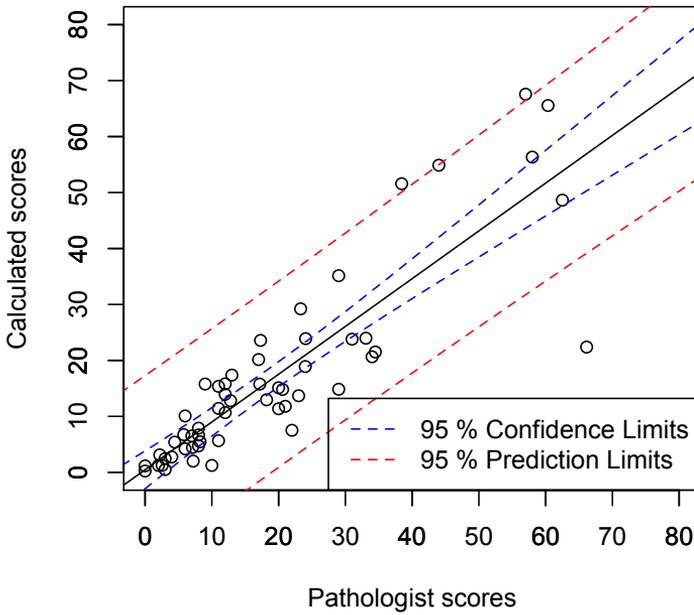
## APPENDIX E

# Regression model for Ki67 scores outside the hottest hotspot

---

This appendix contains the first fitted regression models  $y = \beta x + \alpha$  and  $y = \beta x$  for the Ki67 scores calculated randomly outside the hottest hotspot and some diagnostic plots for the models.

Figure E.1 shows the fitted regression model  $y = \beta x + \alpha$  and the table E.1 shows the test values from the t-test of the slope and intercept in the model.

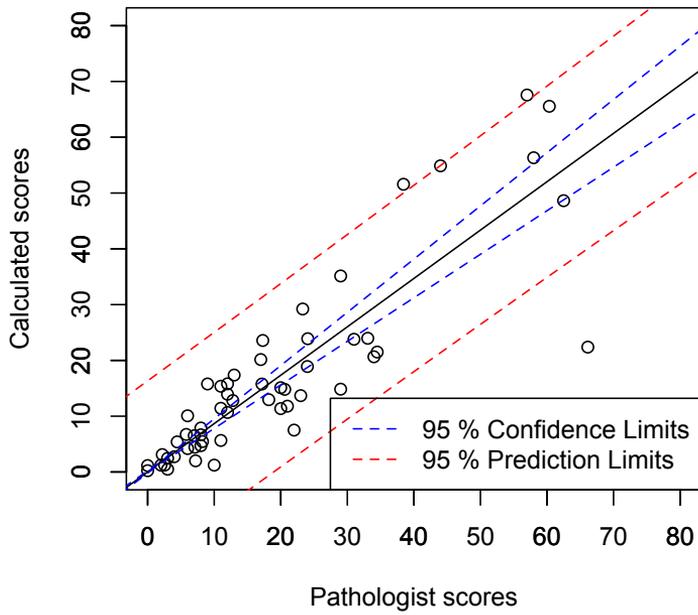


**Figure E.1:** Linear regression model of form  $y = \beta x + \alpha$  fitted to scores calculated outside the hottest hotspot as a function of the pathologist scores.

	Estimate	Std. Error	t value	$\Pr(> t )$
Intercept	0.45351	1.68108	0.27	0.788
Slope	0.85372	0.06615	12.90	<2e-16

**Table E.1:** Test for  $\alpha$  and  $\beta$ 's significant in the regression model  $y = \beta x + \alpha$ .

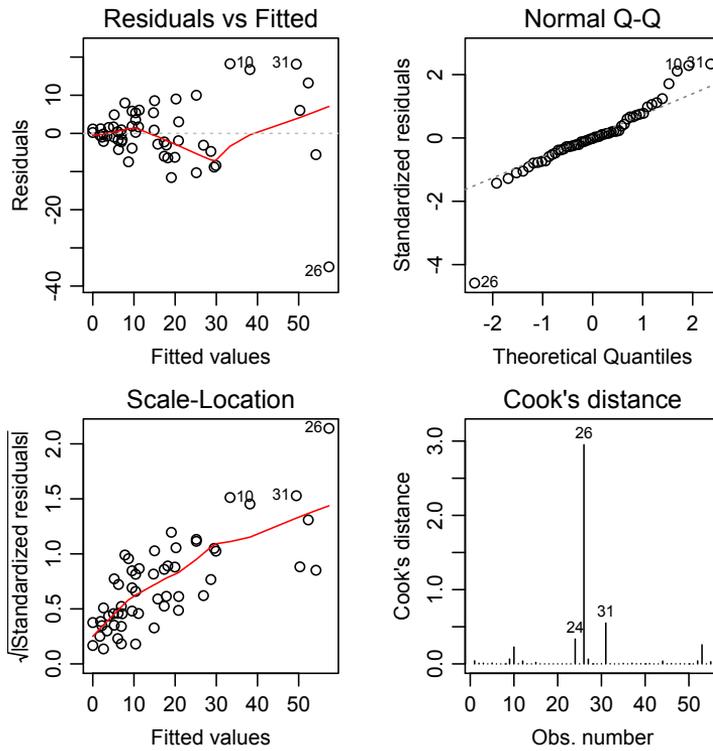
The intercept wasn't significantly different from 0 and the model was reduced to  $y = \beta x$ . Figure E.2 shows the reduced fitted model and table E.2 contain the test values for the slope ( $\beta$ ) different from 0 in the model. Figure E.3 shows the diagnostic plot for the model.



**Figure E.2:** Linear regression model  $y = \beta x$  fitted to scores calculated outside the hottest hotspot.

	Estimate	Std. Error	t value	$\Pr(> t )$
Slope	0.86714	0.04323	20.06	$<2e-16$

**Table E.2:** Test for  $\alpha$ 's significant in the regression model  $y = \beta x$ .



**Figure E.3:** Diagnostic plot for the model  $y = \beta x$ .

As it can be seen from the QQ-plot tails exist in each end which indicates that the residuals aren't normally distributed and the residual plot also shows an increasing variance for higher score values.

A transformation of the data was therefore necessary and the obvious choice was a Log transformation of the data, a regression model of the form  $\text{Log}(y) = \beta \text{Log}(x) + \alpha$ . The fitted Log model and its diagnostic plot can be found in appendix F.

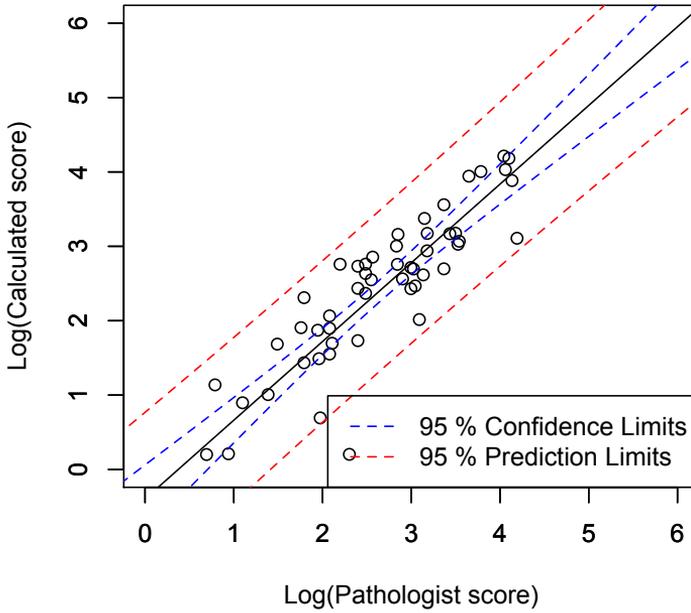
## APPENDIX F

# Log model for Ki67 scores outside the hottest hotspot

---

This appendix contains the fitted log transformed model for the Ki67 scores calculated randomly outside the hottest hotspot.

Figure F.1 shows the fitted model  $\text{Log}(y)=\beta\text{Log}(x)+\alpha$  and table F.1 contains the test value for the t-test of the slope and intercept different from 0.



**Figure F.1:** Linear regression model  $\text{Log}(y)=\beta\text{Log}(x)+\alpha$ .

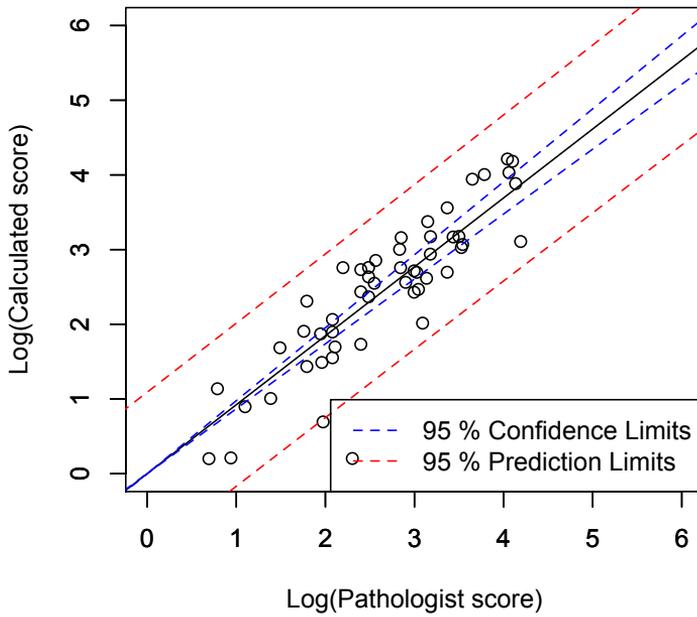
	Estimate	Std. Error	t value	$\text{Pr}(> t )$
Intercept	-0.39590	0.22747	-1.74	0.0878
Log(Slope)	1.05805	0.08197	12.91	$<2e-16$

**Table F.1:** Test for  $\alpha$  and  $\beta$ 's significant in the regression model  $\text{Log}(y)=\beta\text{Log}(x)+\alpha$

The intercept wasn't significant for the model and a reduced model of the form  $\text{Log}(y)=\beta\text{Log}(x)$  was fitted.

Figure F.2 shows the fitted model  $\text{Log}(y)=\beta\text{Log}(x)$  and table F.2 contains for the test value of the slope in the model.

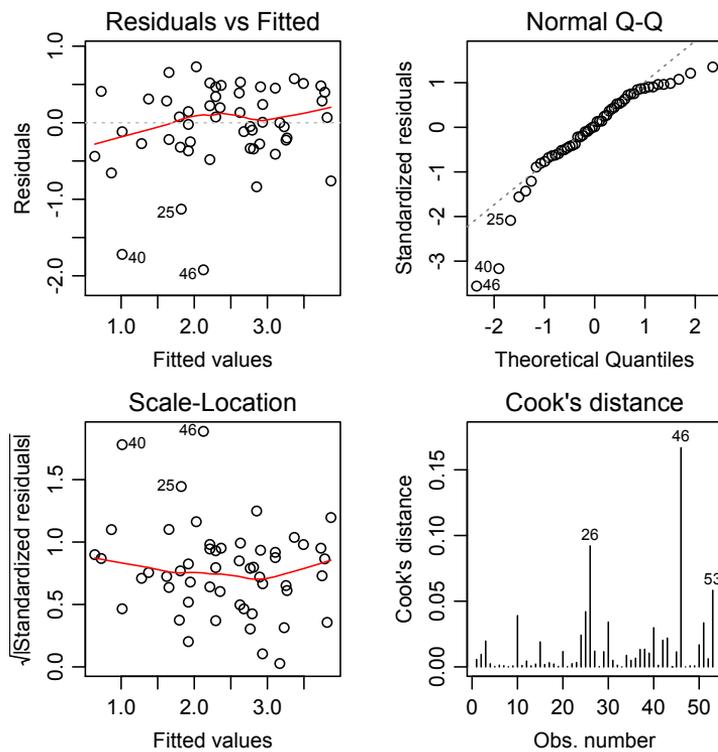
Figure F.3 shows the diagnostic plot for the model.



**Figure F.2:** Linear regression model  $\text{Log}(y)=\beta\text{Log}(x)$ .

	Estimate	Std. Error	t value	$\text{Pr}(> t )$
$\text{Log}(\text{Slope})$	0.9230	0.0269	34.31	$<2e-16$

**Table F.2:** Test for  $\alpha$  and  $\beta$ 's significant in the regression model.



**Figure F.3:** Diagnostic plots for the linear regression model  $\text{Log}(y)=\beta\text{Log}(x)+\alpha$ .

The increasing variance for higher scores are eliminated but the variance isn't constant and high for the low scores.

## APPENDIX G

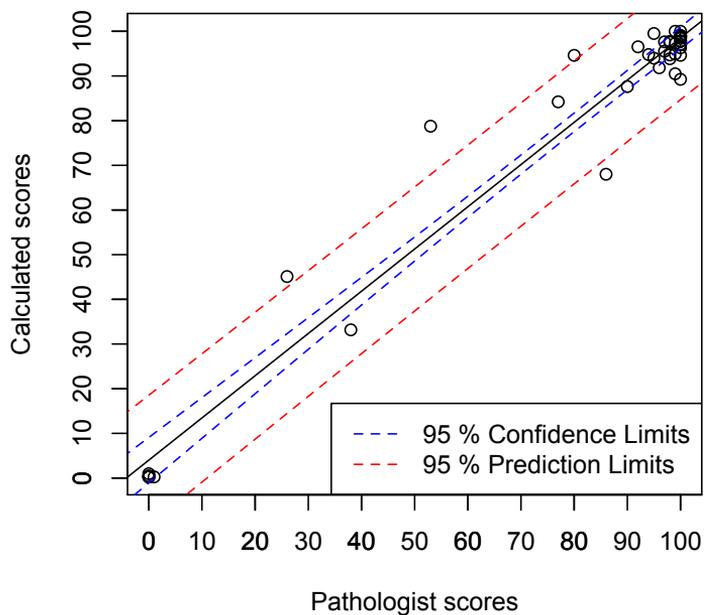
# Regression model for the ER scores

---

This appendix contain the first fitted model to the ER scores of the form  $y = \beta x + \alpha$  and the table for the test of the coefficients in the model.

Figure G.1 shows the fitted model and table G.1 contains the test values for the t-test of the slope and intercept.

As it can be seen from the table the intercept isn't significant for the model and therefore the model was reduced as shown and described in section 6.10.



**Figure G.1:** Linear regression model fitted to scores from method 1 as a function of the pathologist scores.

	Estimate	Std. Error	t value	Pr(> t )
Intercept	1.77435	1.54346	1.15	0.256
Slope	0.96866	0.01956	49.52	<2e-16

**Table G.1:** Test for  $\alpha$  and  $\beta$ 's significant in the regression model  $y = \beta x + \alpha$ .

## APPENDIX H

# Regression models for Ki67 scores in the second, third and fourth hottest hotspot

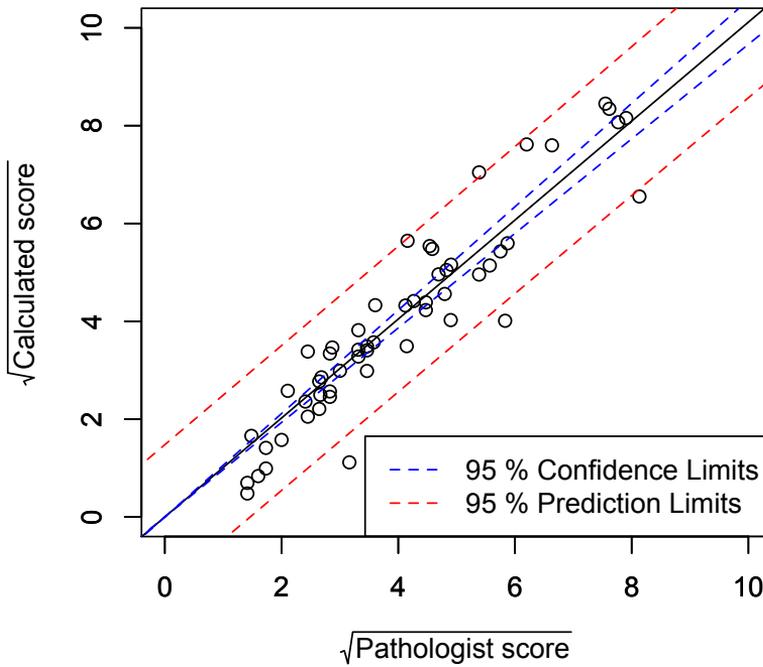
---

This appendix contain the fitted models for the Ki67 scores calculated in the second, third and fourth hottest hotspot.

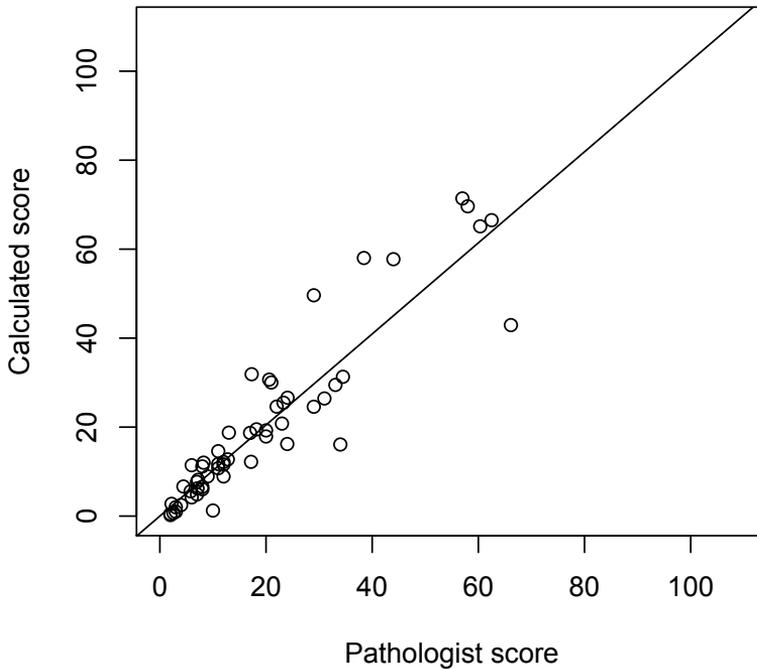
Figure H.1 shows the fitted model with transformed data and the model transformed to the original scale for the second hottest hotspot. The diagnostic plot for the model can be seen on figure H.2.

Figure H.3 shows the fitted model with transformed data and the model transformed to the original scale for the third hottest hotspot. The diagnostic plot for the model can be seen on figure H.4.

Figure H.5 shows the fitted model with transformed data and the model transformed to the original scale for the fourth hottest hotspot. The diagnostic plot for the model can be seen on figure H.6.



(a) The fitted model  $\sqrt{y} = \beta\sqrt{x}$  for the scores calculated in the second hottest hotspot.



(b) The fitted model on (a) transformed to the original scale.

**Figure H.1:** (a) the fitted model  $\sqrt{y} = \beta\sqrt{x}$  and on (b) the fitted model transformed to the original scale

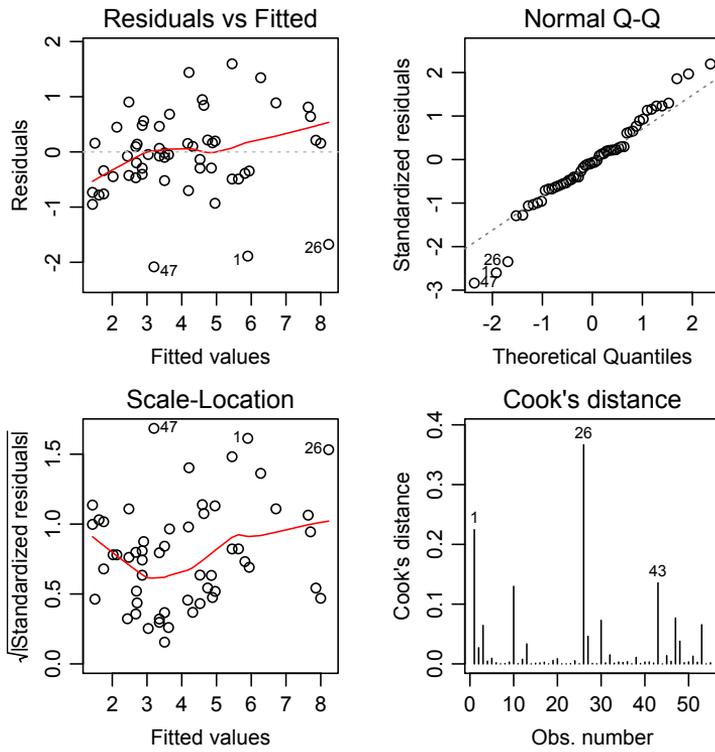
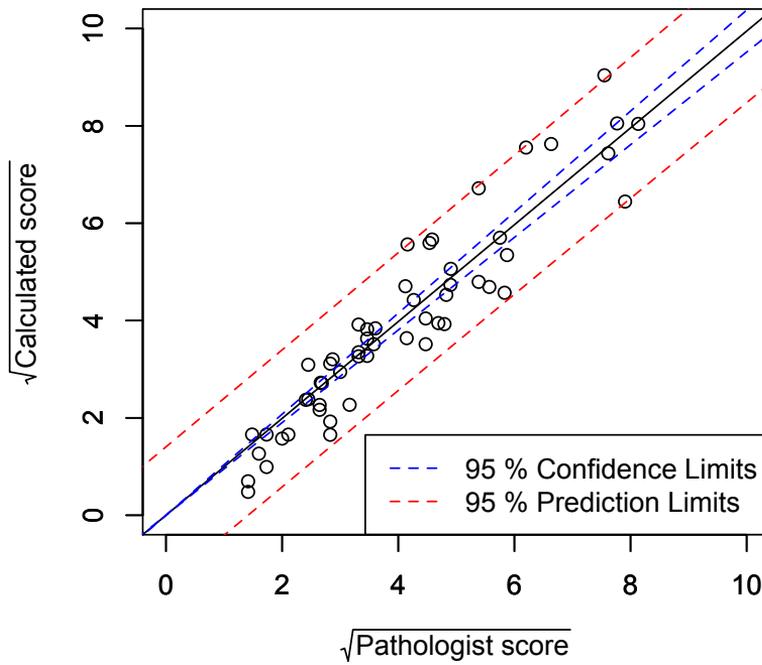
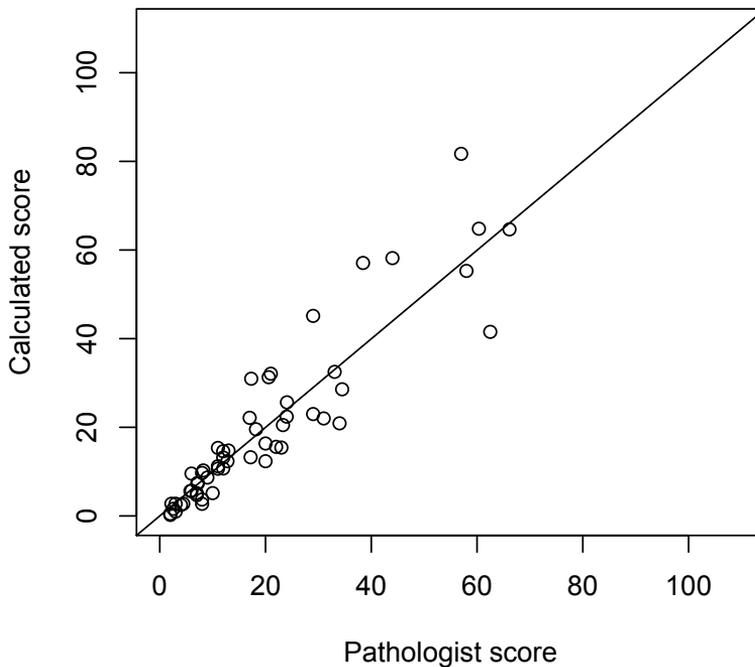


Figure H.2: Diagnostic plot for the model shown on H.1.



(a) The fitted model  $\sqrt{y} = \beta\sqrt{x}$  for the scores calculated in the third hottest hotspot.



(b) The fitted model on (a) transformed to the original scale.

**Figure H.3:** (a) the fitted model  $\sqrt{y} = \beta\sqrt{x}$  and on (b) the fitted model transformed to the original scale

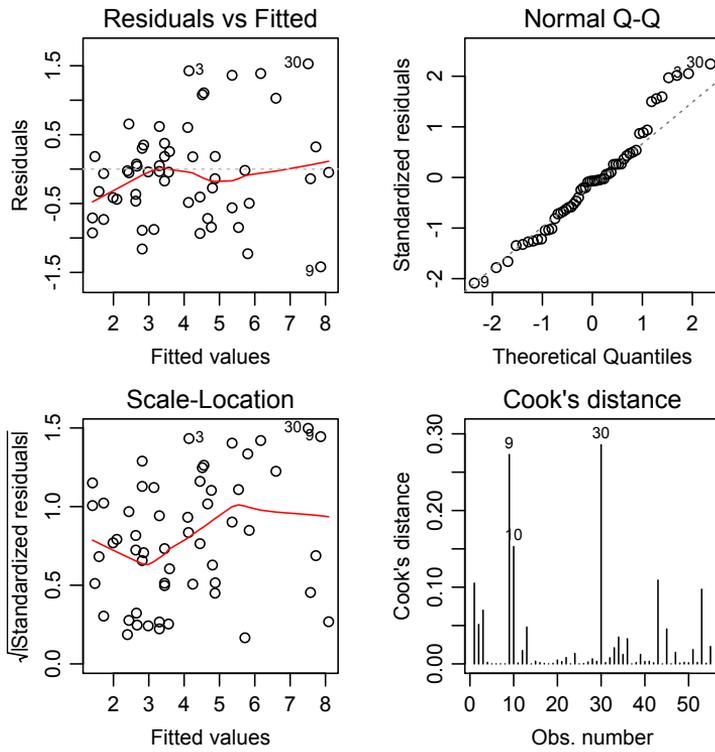
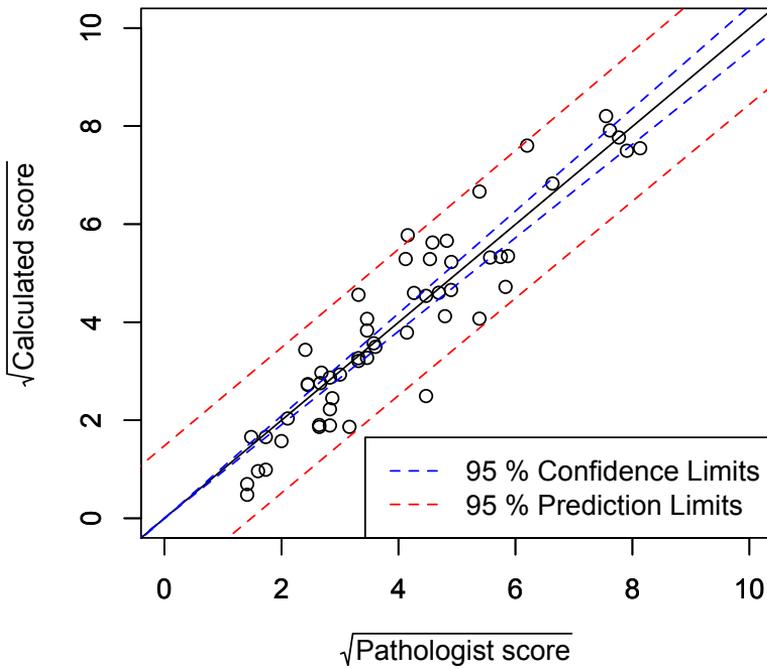
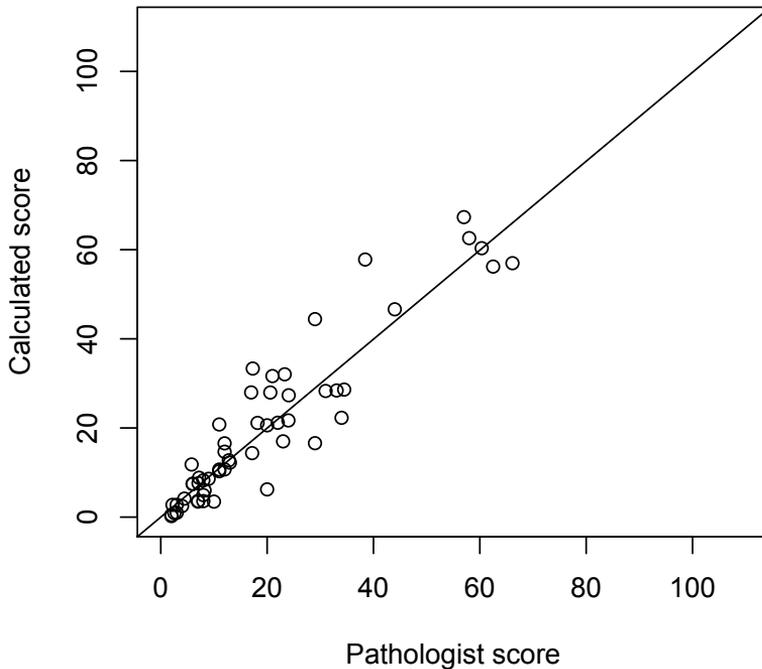


Figure H.4: Diagnostic plot for the model shown on H.3.



(a) The fitted model  $\sqrt{y} = \beta\sqrt{x}$  for the scores calculated in the fourth hottest hotspot.



(b) The fitted model on (a) transformed to the original scale.

**Figure H.5:** (a) the fitted model  $\sqrt{y} = \beta\sqrt{x}$  and on (b) the fitted model transformed to the original scale

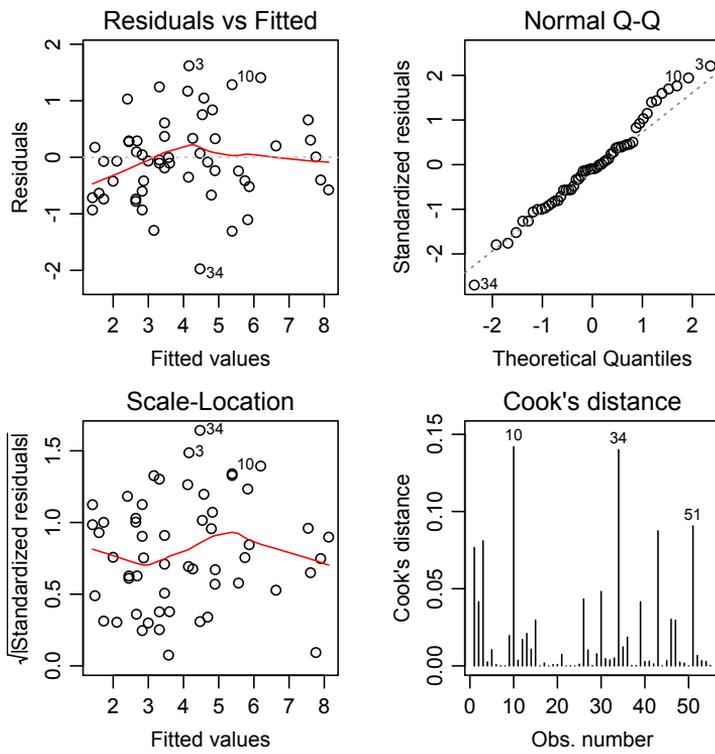


Figure H.6: Diagnostic plot for the model shown on H.5.



# Bibliography

---

- [1] <http://www.who.int/mediacentre/factsheets/fs297/en/>.
- [2] WHO. <http://www.who.int/cancer/detection/breastcancer/en/index1.html>.
- [3] Ulrik Baandrup; Per Prætorius Clausen; Grete Krag Jacobsen; Niels Græm; Claus Fenger. *Klinisk patologi*. FADL's Forlag, 2004.
- [4] John Hopkins Medicine. [http://www.hopkinsmedicine.org/avon\\_foundation\\_breast\\_center/breast\\_cancers\\_other\\_conditions/invasive\\_ductal\\_carcinoma.html](http://www.hopkinsmedicine.org/avon_foundation_breast_center/breast_cancers_other_conditions/invasive_ductal_carcinoma.html).
- [5] Breast Cancer Care. <http://www.breastcancercare.org.uk/breast-cancer-information/about-breast-cancer/diagnosis/questions-you-may-want-ask-your-doctor/what-type-breast-cancer-do-i-have>.
- [6] J.A. Ramos-Vara. Technical aspects of immunohistochemistry. *Vet Pathol*, 2005.
- [7] M. K. Heatley. Cytokeratins and cytokeratin staining in diagnostic histopathology. *Histopathology*, pages 479–483, 1996.
- [8] Rodney T. Miller. Cytokeratin ae1/ae3. *The Focus - Immunohistochemistry*, November 2003.
- [9] Werner W. Franke; Erika Schmid; Mary Osborn; Klaus Weber. Intermediate-sized filaments of human endothelial cells. *The Rockefeller University Press*, 1979.
- [10] The Danish Cancer Society. <http://www.cancer.dk/hjaelp+viden/ordbog/k/ki67.htm>.

- [11] Elisabeth Luporsi; Fabrice Andre; Frederique Spyrtatos; Pierre-Marie Martion; Jocelyne Jacquemier; Frederique Penault-Llorca; Nicole Tubiana-Mathieu; Brigitte Sigal-Zafrani; Laurent Arnould; Anne Gompel; Caroline Egele; Bruno Poulet; Krishna B. Clough; Hubert Crouet; Alain Fourquet; Jean-Pierre Lefranc; Carole Mathelin; Nicolas Rouyer; Daniel Serin; Marc Spielmann; Margaret Haugh; Marie-Pierre Chenard; Etienne Brain; Patricia de Cremoux; Jean-Pierre Bellocq. Ki-67: level of evidence and methodological considerations for its role in the clinical management of breast cancer: analytical and critical review. *Breast Cancer Res Treat*, 2012.
- [12] Ander Urruticoechea; Ian E. Smith; Mitch Dowsett. Proliferation marker ki-67 in early breast cancer. *Journal of clinical oncology*, 2005.
- [13] Mehrdad Nadji; Carmen Gomez-Fernandez; Parvin Ganjei-Azar; Azorides R. Morales. Immunohistochemistry of estrogen and progesterone receptors reconsidered. *American Society for Clinical Pathology*, 2005.
- [14] Corbin E. Meacham; Sean J. Morrison. Tumour heterogeneity and cancer cell plasticity. *Nature*, 2013.
- [15] S Detre; G Sacconi Jotti; M Dowsett. A "quickscore" method for immunohistochemical semiquantitation: validation for oestrogen receptor in breast carcinomas. *J Clin Pathol*, 1995.
- [16] Michael Gnant; Nadia Harbeck; Christoph Thomssen. St. gallen 2011: Summary of the consensus discussion. *BreastCare*, 2011.
- [17] Guido Sauter. Representativity of tma studies. *Tissue Microarrays; Methods and Protocols, Methods in Molecular Biology*, 664:27–35, 2010.
- [18] Makiko Ono; Hitoshi Tsuda; Mayu Yunokawa; Kan Yonemori; Chikako Shimizu; Kenji Tamura; Takayuki Kinoshita; Yasuhiro Faujiwara. Prognostic impact of ki67 labeling indices with 3 different cutoff values, histological grade, and nuclear grade in hormone-receptor-positive, her2-negative, node-negative invasive breast cancer. *Breast Cancer*, 2013.
- [19] Einar Gudlaugsson; Ivar Skaland; Emiel A M Janssen; Rune Smaaland; Zhiming Shao; Anaïs Malpica; Feja Voorhorst; Jan P A Baak. Comparison of the effect of different techniques for measurement of ki67 proliferation on reproducibility and prognosis prediction accuracy in breast cancer. *Histopathology*, 2012.
- [20] <http://www.visiopharm.com>.
- [21] Steven J Potts; Joseph S Krueger; Nicholas D Landis; David A Eberhard; G David Young; Steven C Schmechel; Holger Lange. Evaluating tumor heterogeneity in immunohistochemistry-stained breast cancer tissue. *Laboratory Investigation*, 2012.

- [22] Henri P. Gavin. The levenberg-marquardt method for nonlinear least squares curve-fitting problems. Technical report, Department of Civil and Environmental Engineering Duke University, 2013.
- [23] D. G. Altman; J. M. Bland. Measurement in medicine: the analysis of method comparison studies. *The Statistician*, 1983.