



REDEFINING
CIRCULATING TUMOR CELLS
BY IMAGE PROCESSING

SJOERD T. LIGTHART

REDEFINING CIRCULATING TUMOR CELLS
BY IMAGE PROCESSING

Sjoerd Theodorus Ligthart

Samenstelling promotiecommissie:

Prof. dr.	G. van der Steenhoven	Universiteit Twente (voorzitter)
Prof. dr.	L.W.M.M. Terstappen MD	Universiteit Twente (promotor)
Prof. dr.	W. Steenbergen	Universiteit Twente
Prof. dr.	C.H. Slump	Universiteit Twente
Prof. dr.	J.S. de Bono MD	Royal Marsden NHS foundation trust
Prof. dr.	I.T. Young	TU Delft
Dr.	F.-C. Bidard MD	Institut Curie
Dr.	J.F. Keij	Veridex LLC

This work was financially supported by Veridex LLC.

Copyright © 2012 by S.T. Ligthart, Enschede, the Netherlands.

All rights reserved. No part of this book may be reproduced or transmitted, in any form or by any means, electronically or mechanically, including photocopying, microfilming, and recording, or by any information storage or retrieval system, without prior written permission of the author.

ISBN 978-90-365-3364-5
DOI 10.3990/1.9789036533645

REDEFINING CIRCULATING TUMOR CELLS
BY IMAGE PROCESSING

PROEFSCHRIFT

ter verkrijging van
de graad van doctor aan de Universiteit Twente,
op gezag van de rector magnificus,
prof. dr. H. Brinksma,
volgens besluit van het College voor Promoties
in het openbaar te verdedigen
op donderdag 10 mei 2012 om 16:45 uur

door

Sjoerd Theodorus Ligthart

geboren op 4 februari 1980
te Hoogkarspel

Dit proefschrift is goedgekeurd door:

Prof. dr. L.W.M.M. Terstappen MD (promotor)

CONTENTS

Thesis motivation and outline	ix
Motivation	ix
Outline	x
1 Introduction	1
1.1 Cancer is currently the deadliest disease in the western world	1
1.2 Diagnosing and treating cancer	3
1.3 Frequency and clinical relevance of CTC in metastatic cancer	4
1.4 CTC enrichment and staining with the CellTracks Autoprep [®]	7
1.5 Enumeration of CTC with the Celltracks Analyzer II [®]	8
1.6 Expression of treatment targets on CTC	10
1.7 Next generation imaging systems	13
1.7.1 CellTracks FISH [™]	13
1.7.2 CellTracks TDI [™]	13
1.8 Challenges for the CellSearch system	14
1.9 References	15
2 Simulation and calibration of spectral imaging methods	23
2.1 Introduction	24
2.2 Materials and methods	25
2.2.1 Introduction	25
2.2.2 Testing the spectral imaging methods: the challenge	25
2.2.3 Fluorochrome combinations	25
2.2.4 Fluorescence excitation and emission theory	26
2.2.5 Test image for simulation	29
2.2.6 Inserting a ND filter for using the DM method	29
2.2.7 Optimizing the LCTF method	30
2.2.8 Optimizing the Fourier method	30
2.2.9 Optimizing the Prism method	31
2.2.10 Classification of the test image after linear unmixing	32
2.2.11 Measuring cameras and setups using a calibration board	33
2.3 Results	34
2.3.1 Optimizing the spectral imaging methods	34
2.3.2 Total integration time for simulation of the test image	36

2.3.3	Measuring throughput in microscopes	36
2.4	Discussion	38
2.5	References	40
3	FISH probe counting in CTC	43
3.1	Introduction	43
3.2	Method	44
3.2.1	Patient samples	44
3.2.2	Sample preparation for CTC enumeration	44
3.2.3	Data acquisition for CTC enumeration	45
3.2.4	Samples used for FISH counting algorithm development	46
3.2.5	Sample preparation for FISH probes on CTC	46
3.2.6	Data acquisition for FISH probe detection in CTC	46
3.2.7	Algorithm for counting FISH signals	47
3.2.8	Expert reviewing of samples	49
3.3	Results	50
3.3.1	Counting of the leukocyte training samples	51
3.3.2	Counting of samples containing CTC	54
3.4	Discussion	54
3.4.1	Automated counting is necessary and feasible	54
3.4.2	Sources of error for human and PC	56
3.4.3	Future research	58
3.5	References	58
4	Image analysis algorithm for CTC recognition	61
4.1	Introduction	62
4.2	Materials and methods	63
4.2.1	Patients	63
4.2.2	Isolation of EpCAM+ objects and fluorescence imaging	64
4.2.3	Counting of manual CTC by human reviewers	64
4.2.4	Algorithm development for classifying automated CTC	66
4.3	Results	70
4.3.1	Processing of images	70
4.3.2	Optimal channel for segmentation	71
4.3.3	Features with highest impact on HR	71
4.3.4	Rules to assign a CTC	72
4.3.5	Automatic CTC count versus manual CTC count in patients and controls	74
4.4	Discussion	74
4.5	References	76
5	Automated CTC identification	79
5.1	Introduction	80
5.2	Methods	81
5.2.1	Ethics statement	81

5.2.2	Participants	81
5.2.3	Manual Counting of Circulating Tumour Cells (mCTC)	82
5.2.4	Automated counting of EpCAM+ objects using a computer algorithm (aCTC)	84
5.2.5	Statistical analysis	84
5.3	Results	85
5.3.1	Choosing the optimal classifier and processing of samples	85
5.3.2	Automated CTC count compared to manual CTC count in patients and controls	85
5.3.3	Defining cut-off values for aCTC AND mCTC	86
5.3.4	Validation of automated CTC count	88
5.4	Discussion	90
5.5	References	93
6	Morphology of CTC	97
6.1	Introduction	98
6.2	Materials and Methods	99
6.2.1	Patients	99
6.2.2	Manual enumeration of Circulating Tumor Cells	100
6.2.3	Automated enumeration of Circulating Tumor Cells	100
6.2.4	Additional morphological measurements on aCTC	101
6.2.5	Identification and morphological measurements of leukocytes	101
6.2.6	Statistical analysis	101
6.3	Results	103
6.3.1	Frequency of aCTC versus mCTC	103
6.3.2	Measurements of morphological parameters of aCTC	104
6.3.3	mCTC, aCTC and morphological parameters versus survival in metastatic breast, colorectal and prostate cancer	107
6.4	Discussion	110
6.5	References	113
7	Interpretation of changes in CTC counts	119
7.1	Introduction	120
7.2	Materials and methods	120
7.2.1	Patient data	120
7.2.2	CTC enumeration	121
7.2.3	Poisson model for reduction in CTC count	122
7.2.4	CTC reduction criteria	123
7.2.5	Statistical analysis	124
7.3	Results	124
7.3.1	Overall survival as a function of CTC number	124
7.3.2	Changes in the number of CTC and overall survival	124
7.3.3	True CTC changes determined using a Poisson model	126

7.3.4	Relation between CTC definitions and clinical outcome	127
7.3.5	Correlation of reduction criteria with survival	128
7.3.6	Applying criteria for mCTC change to multiple time points	129
7.4	Discussion	131
7.5	References	134
8	Unbiased quantitative CTC Her-2 assessment	137
8.1	Introduction	138
8.2	Materials and methods	139
8.2.1	Patients	139
8.2.2	Her-2 assessment of the primary tumor	139
8.2.3	Manual CTC enumeration and Her-2 assessment . .	140
8.2.4	Automated CTC enumeration and Her-2 assessment	140
8.2.5	Determination of a threshold for automated Her-2 assessment	141
8.3	Results	142
8.3.1	Identification of CTC and leukocytes in M1 breast- cancer patients	142
8.3.2	Her-2 expression breast-cancer cell lines	143
8.3.3	Her-2 staining of leukocytes as internal control . . .	143
8.3.4	Her-2 expression of CTC from M1 breast-cancer patients	146
8.3.5	Primary tissue Her-2 expression versus CTC Her-2 expression in M1 patients	147
8.3.6	Validation in M0 breast cancer patients	148
8.3.7	Processing time and reproducibility of the algorithm	149
8.4	Discussion	150
8.5	References	152
	Summary	157
	Conclusions	157
	Outlook	159
	Samenvatting	161
	Conclusie	161
	Vooruitzichten	163
	Publications	165
	Journal articles and book chapters	165
	Conference contributions (oral)	166
	Conference contributions (poster)	166
	Acknowledgments	169
	About the author	171

THESIS MOTIVATION AND OUTLINE

MOTIVATION

As of 2008, more people in the Netherlands died of cancer than of cardiovascular disease (CBS statline, 2011). The hope is that cancer can be contained and may become a chronic disease, but this aim can certainly not be achieved without a tremendous amount of research and financial investments. One of the major difficulties in fighting cancer lies in the fact that cancer is extremely heterogeneous and can alter during the disease. Furthermore, certain cancer cells can enter the blood stream, travel through the body and create distant metastases. If the cancer cells have spread, treatment options are limited and patient prognosis is very unfavorable. The cells that give rise to metastases have to travel through the blood and are termed circulating tumor cells (CTC). These CTC are the subject of this thesis.

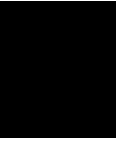
Assessing the presence of CTC in the blood of cancer patients may improve the staging of the patient's cancer and indicate whether or not the cancer is actively spreading throughout the body. In addition, characterization of CTC offers the opportunity of a "real time liquid biopsy" that can help to select the most appropriate therapy. Monitoring the number of CTC after administration of treatment may indicate whether or not the therapy is effective. Current techniques to characterize CTC rely on (i) isolation of CTC from the bloodstream by making use of differences in phenotypical and or physical characteristics between CTC and blood cells, and (ii) labeling them with for instance fluorescent markers to distinguish them from blood cells and other contaminants. In this last step, human interpretation of recorded images is used to identify the CTC. This review by trained experts is laborious, time consuming, and introduces intra- en inter-reviewer variations. Furthermore, the assessment of the reviewers is a qualitative one, as reviewers cannot extract numerical data from the images easily. They have to rely on comparison with CTC they have seen before during the training. The definition of CTC that is currently used by reviewers, is prognostic for patient survival, it is however currently unknown if this definition is optimal. Characterization of CTC by computer algorithms may tackle these issues and the development of such algorithms is explored in this thesis.

OUTLINE

The outline of this thesis is as follows: **chapter 1** is an introduction to cancer, CTC, and research that is currently performed to isolate and characterize CTC. It states in more detail which challenges are faced in this thesis. **Chapter 2** shows simulations and measurements on equipment for recording images of CTC. It shows a method to compare different microscopic setups in a quantitative way, which improves comparison of samples that were recorded by these setups. **Chapter 3** shows how computer algorithms may help in the enumeration of chromosomes in CTC that have been labeled by Fluorescence In-Situ Hybridization, a technique that is commonly used for CTC characterization.

Chapters 4 and 5 show details of the creation of new definitions for CTC from prostate cancer patients by a computer algorithm, which was trained by using survival data of these patients. The optimal CTC definition was validated on an independent data set from a different group of prostate cancer patients. We compared the results of this algorithm with results from the current CellSearch CTC method, which is cleared by the Food and Drug administration of the U.S. for monitoring the disease of metastatic breast, colorectal and prostate cancer patients. In **chapter 6** we explored if this new definition is also valid for indentifying CTC in samples from breast and colon cancer patients. We use our newly discovered definition to measure more parameters of CTC and investigate if these parameters correlate with patient survival.

Training of the algorithm described in chapters 4 and 5 showed a wide range of CTC definitions that correlated well with survival. Using definitions that included small tumor particles resulted in the counting of much more objects as compared to the CellSearch definition. In **chapter 7** we investigated what definitions are best suited for measuring a relevant change in CTC number after administration of therapy. Finally, we show in **chapter 8** that our computerized algorithm is able to characterize the expression of the biomarker Her-2 on CTC in a quantitative way. This type of biomarker is relevant for the choice of treatment, paving the way for personalized treatment.

INTRODUCTION¹

Sjoerd T. Ligthart

1.1 CANCER IS CURRENTLY THE DEADLIEST DISEASE IN THE
WESTERN WORLD

As of 2008, more people in the Netherlands died of cancer than of cardiovascular disease [1]. There are a number of causes why cancer is such a difficult disease to detect, fight, and ultimately control. For a healthy individual, there are several risk factors that have been proven to cause cancer such as genetic factors, tobacco use, infection, radiation, lack of physical activity, poor diet, obesity, and environmental pollutants. As human beings are constantly confronted with these risk factors, it is very difficult to establish when and where in the body a tumor may arise. When cells are affected by one or more of these risk factors, their genes may be altered and become oncogenes, causing behavior that is not seen in healthy cells: they may divide and grow uncontrollably, programmed cell-death (apoptosis) may be deregulated, cells may induce blood-vessel growth (angiogenesis) and may invade these blood vessels or the lymphatic system. When they have entered the blood, these cells are termed circulating tumor cells (CTC).

The existence of CTC was already recognized in the 19th century [2]. The majority of CTC will be destroyed by the reticuloendothelial system, but some do survive and may cause distant metastasis, as was already

¹Part of this chapter will be published in "Identification of Circulating Tumor Cells" by Coumans F.A.W., Ligthart S.T., and Terstappen L.W.M.M., Biofunctional Surface Engineering in Medicine "Nanobiotechnology" series by Pan Stanford Publishing, editor: Martin Schol.

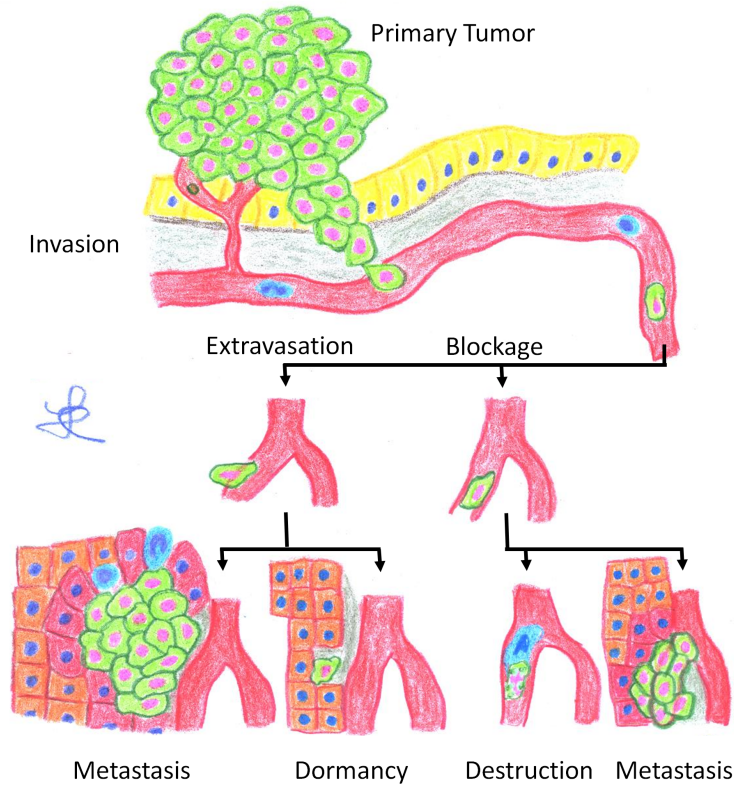


FIGURE 1.1 : A schematic representation of possible routes CTC may follow in the body (drawing by Leon Terstappen).

described by Stephen Paget in 1889 [3]. He proposed that CTC may have a bigger affinity with certain organs: the "seed and soil" hypothesis. About 100 years later, it was proposed that CTC may have the biggest affinity with the primary tumor from which the CTC originated and thus settle near the primary tumor [4]. Nevertheless, these CTC pose a big threat to a patient as is illustrated in figure 1.1.

When CTC invade the bloodstream, they could either get stuck in a blood vessel or find a place where they can exit the blood stream. Blockage could lead to destruction of the CTC by macrophages or to division of CTC to form a (micro) metastasis. Extravasation of CTC could lead to metastasis or cell dormancy. Dormant cells may be inactive up to a few years and can suddenly be re-activated to cause a recurrence of the disease at a -for cancer researchers still- random moment. When distant metastases are formed, it is exponentially more difficult to treat the de-localized disease.

1.2 DIAGNOSING AND TREATING CANCER

Carcinomas are usually detected by physical examination or medical imaging modalities such as a X-ray CT, PET, or MRI scans. Definite diagnosis requires a trained pathologist who must judge if a small piece of tissue -a biopsy- is malignant or not, i.e. if the tumor is spreading and invading nearby tissue. In case of a malignancy, the most important question is whether or not the tumor has spread beyond the primary tumor. Imaging modalities, a genetic profile of the tumor cells and serum tumor markers are utilized to verify if spreading has occurred and to set up a treatment plan. The first choice is usually surgical removal of the tumor and in some cases the choice is made to administer neo-adjuvant therapy in an attempt to reduce tumor size and to enhance the chance that it can be removed completely. Neo-adjuvant therapy may consist of chemotherapy (inhibits division of cells globally), radiation therapy (damages DNA of cells locally), immunotherapy (stimulates the immune system to destroy the tumor), or targeted therapy (deregulates processes in the tumor cells specifically). After surgery the risk profile for recurrence of the disease is made by T (tumor size), N (lymphnode involvement), and M (metastasis) staging [5].

Based on the risk profile it is decided whether or not the patient will receive adjuvant therapy; the type of therapy depends on the characteristics of the tumor. Frequently such treatments consist of chemotherapy, but expression of specific receptors on the tumor offers the potential for targeted therapy, which in general is less toxic. For example Trastuzumab -an antibody that recognizes human epidermal growth factor receptor 2 (Her-2)- can be administered when the tumor cells express Her-2 [6]. Likewise expression of the estrogen receptor (ER) and progesterone receptor (PR) in breast cancer tumors permits the administration of therapies targeting these receptors. The risk profile determined by classical means needs significant improvements, as a proportion of patients with a low risk profile that did not receive adjuvant therapy currently still develops a recurrence; a significant proportion of patients that do receive adjuvant therapy would not have needed the therapy. Genetic profiling of the tumor [7, 8], detection of presence of micrometastases [9] or detection of CTC [10, 11] can improve the accuracy of the risk profile.

In those cases in which metastatic disease has been established, treatment is not curative for most cancer types and treatments are intended to prolong survival while maintaining a reasonable quality of life. Monitoring of the treatment is performed by imaging modalities in those cases that the disease is "measurable", other approaches are the use of tumor markers and clinical signs and symptoms. These methods are not sufficiently sensitive and specific, or cannot be performed in a timely fashion for the determination of treatment effects. Tumor markers are restricted for those tumors that produce proteins such as PSA for the majority of prostate cancer patients and MUC-1 for a subset of breast cancer patients. Other causes can however

also give rise to an increase in these tumor markers [12]. Image scans are usually performed before initiation and 3–6 months after initiation of therapy. However, one can only follow some of the lesions in time and frequently the majority of the disease is not measurable. In addition, there exists high variability in interpretation of scan images [13], scans may impose a treat to the health of the patient [14] and the costs of scans are substantial [15]. Measurement of the number of CTC could help improve the management of the therapy of patients with metastatic disease as it reflects the activity of the disease in the whole body. In addition CTC may serve as a liquid biopsy to determine the presence of treatment targets and guide towards the most optimal therapy for the individual patient.

1.3 FREQUENCY AND CLINICAL RELEVANCE OF CTC IN METASTATIC CANCER

CTC are very rare events in the blood of cancer patients and were usually only observed in blood smears of patients with extensive metastatic disease [16, 17, 18, 19, 20]. The CellSearch[®] system identifies CTC in 7.5 ml of blood and has been extensively validated for patients with metastatic carcinoma [21]. Modeling of the CTC distribution in 7.5 ml of blood from patients with metastatic breast, colorectal, and prostate cancer was used to arrive at the CTC frequency distribution in all 5 liters of blood [22]. Figure 1.2 depicts the cumulative probability in which CTC can be detected as a function of blood volume in patients with metastatic carcinomas. The figure also shows the frequency of erythrocytes, platelets and leukocytes in blood and highlights the difficulty of detecting CTC in all patients. Ten CTC per ml of blood can only be detected in $\sim 20\%$ of patients, 1 CTC per ml of blood in $\sim 40\%$ of patients and 100 CTC per liter of blood in $\sim 80\%$ of patients.

The CellSearch system only detects CTC that express both the epithelial cell adhesion molecule (EpCAM) [23, 24, 25] and cytokeratins 8, 18 or 19 [26]. Their frequency may therefore be underestimated using the CellSearch system. Whether or not CTC with alternative phenotypes have a similar relation with clinical outcome remains to be determined. A variety of different technologies are currently being explored to identify CTC by other means and should further improve our understanding of CTC [27, 28, 29, 30, 31, 32, 33, 34]. Furthermore, the CellSearch system may also miss some EpCAM⁺,CK 8, 18 or 19⁺ cells. Enumeration of EpCAM⁺ CTC in 100 μ l of whole blood by flow-cytometry showed that the CTC yield can only be increased by ~ 6.5 fold [22, 24]. However, the CTC definition used in the flow-cytometric analysis is less strict, as it does not include CKs, which is the most likely explanation for the largest portion of this discrepancy.

The definition of a CTC in the CellSearch system was set in a series of preclinical studies and was tested in system validation studies [21, 35], as well as in prospective multicenter studies for breast, colon, and prostate cancer [36, 37, 38, 39]. These studies showed that metastatic patients that had equal or more CTC than a certain cut-off (five CTC for breast and prostate cancer, three for colon cancer) had significant lower probability of overall survival and thus a worse prognosis than the group that was below this cut-off. Re-analysis of prostate cancer data furthermore showed that there exists a continuous relationship between the number of CTC and survival [22, 40], and that fragments of tumor cells termed tumor micro particles (TMPs) -CK positive, CD45 negative objects that are $<4 \mu\text{m}$ - are present at a much higher frequency and that their presence also indicates a worse prognosis [41].

Next to these three major types of carcinomas, CTC were enumerated in patients suffering from lung cancer [42, 43], neuroendocrine tumors [44], gastric cancer [45], bladder cancer [46], and ovarian cancer [47]. Specific targets on CTC were investigated to provide tailored treatment in the future: Urokinase receptor (uPAR) and Her-2 analysis [48, 49, 50, 51], IGF-1 receptor [52], and numerous genes using multiplex PCR [53, 54, 55, 56]. Currently, clinical trials are underway to investigate if CTC can be used to guide therapy (see for example ClinicalTrials.gov identifier: NCT00382018).

The relation between CTC and survival is illustrated in figure 1.3 by Kaplan-Meier plots of the probability of overall survival for 296 metastatic breast and prostate cancer patients. For this analysis CTC were identified by an automated algorithm developed as part of this thesis in the images stored by the CellSearch system. Panel A shows the Kaplan-Meier of patients before initiation of therapy. Patients with 0 CTC (N=96, 32%) had a median survival of 33.1 months, patients with 1-3 CTC (N=61, 21%) had a median survival of 21.9 months, patients with 4-19 CTC (N=71, 24%) had a median survival of 15.8 months and patients with ≥ 20 CTC (N=68, 23%) had a median survival of only 9.5 months. Panel B shows the Kaplan-Meier plot of patients at first follow-up after the initiation of

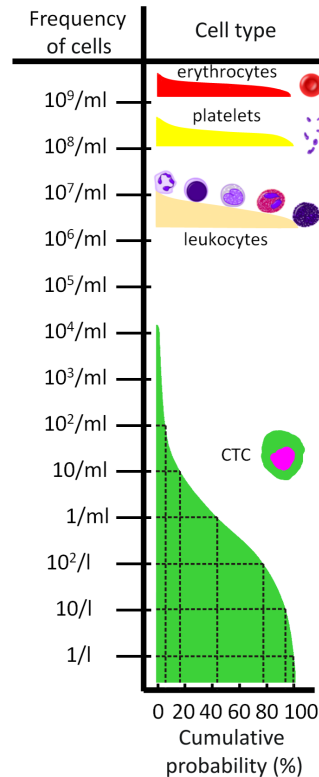


FIGURE 1.2 : Frequency of erythrocytes, platelets, leukocytes and circulating tumor cells in blood of metastatic carcinoma patients

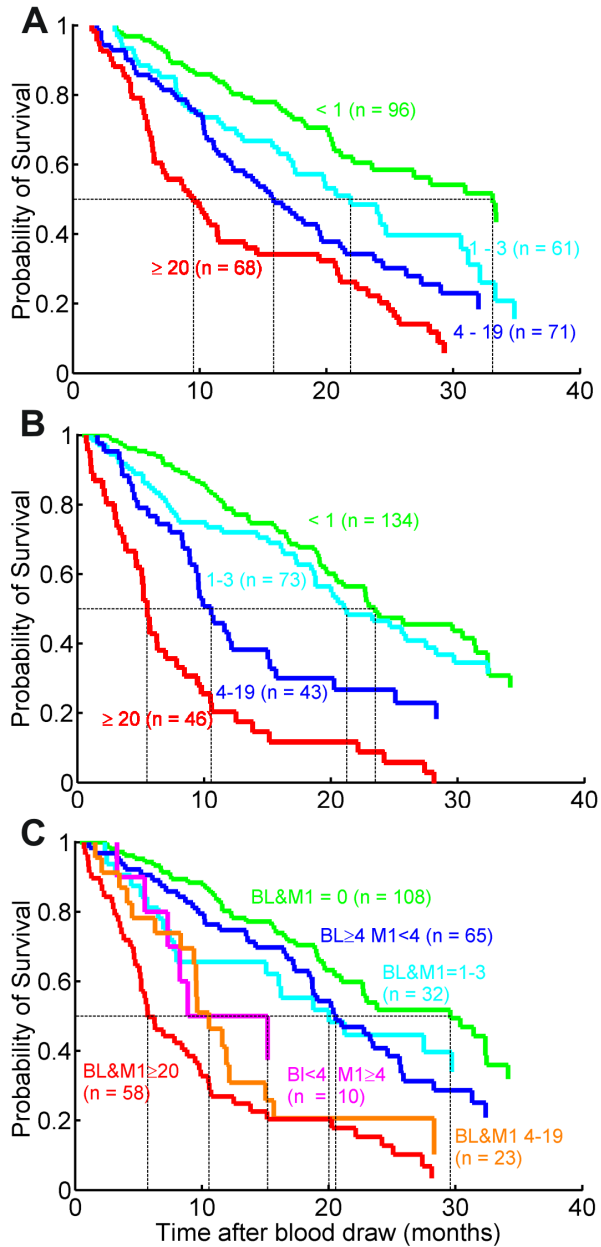


FIGURE 1.3 : CTC and the probability of overall survival for 296 metastatic breast and prostate cancer patients. Panel A shows patients before therapy, Panel B 2-5 weeks after initiation of therapy. Panel C shows the influence of changes in the number of CTC after initiation of therapy. BL = baseline measurement, M1 = first follow-up measurement.

therapy. The number of patients with 0 CTC increased (N=134, 45%) and had a median survival of 23.5 months, patients with 1–3 CTC also increased (N=73, 25%) with a median survival of 21.3 months, patients with 4–19 CTC decreased (N=43, 15%) with a median survival of 10.6 months and patients with ≥ 20 CTC also decreased (N=46, 16%) with an even shorter median survival 5.5 months. Panel C shows the Kaplan-Meier of alterations of CTC counts in patients upon treatment. CTC remained above 20 in 58 of the 68 patients with a median survival of 5.7 months (red line) indicating that therapy did not result in a beneficial effect. Survival did also not improve for those patients with lower CTC or rising number of CTC during therapy with a median survival of 10.6 months, (orange) and of 15.2 months (purple). The group of patients with 0 CTC (green) before and after initiation of therapy increased from 96 to 108 and had a median survival of 29.6 months. Survival of the 65 patients with a CTC reduction (blue) to below 4 clearly improved and patients that remained with low counts (light blue) did not significantly alter.

The low numbers of CTC detected urges the need for elimination of the error in the assignment of CTC as is achieved by the automation of the image analysis algorithm as described in this thesis. A guideline for the interpretation of changes in CTC counts is provided in chapter 7 of this thesis. Patients in these studies in whom 0 CTC were detected in 7.5 ml of blood had metastatic disease and the question arises whether this is a distinct group of patients, if CTC are missed by the CellSearch system or if the volume of blood examined is simply too low. Extrapolation of the sample volume to 5 liters of blood predicted that 99% of patients had at least 1 CTC before initiation of therapy, which decreased to 97% after the first cycles of therapy. Survival chances of patients with EpCAM+ cytokeratin+ nucleated CTC are reduced by 6.6 months for each tenfold CTC increase [22]. These results suggest that a technological leap is needed to identify CTC in all patients with metastatic disease and those patients with primary disease that are at risk for disease recurrence.

1.4 CTC ENRICHMENT AND STAINING WITH THE CELLTRACKS AUTOPREP[®]

The CellTracks Autoprep is an automated sample preparation device that is part of the CellSearch System. Blood is collected from a patient by venipuncture or from a venous port into a CellSave Preservative Tube[®]. These tubes contain EDTA as anticoagulant and a cellular preservative to avoid degradation of the blood sample up to 96 hours while it is being transported to a facility where an Autoprep system is present. In a first step the blood is diluted, mixed by inversion, and centrifuged after which it is transferred to the Autoprep station. The plasma separated from the cells by the centrifugation step is aspirated and discarded. Next, ferrofluid

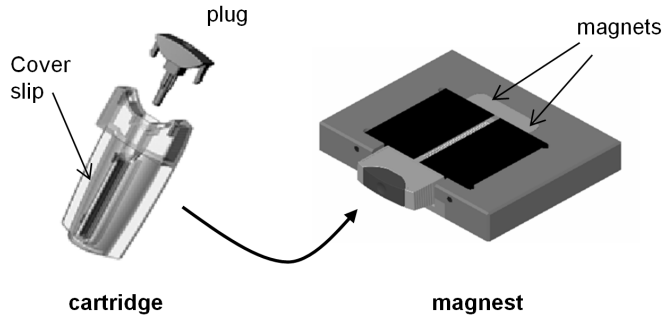


FIGURE 1.4 : Analysis cartridge to which the enriched sample is transferred and magnet in which cells are magnetically pulled towards the cover slip.

conjugated to EpCAM is added as well as dilution and system buffers [57]. The sample is placed between magnets which causes the cells which are labeled with ferrofluid to travel to the area within the tube that has the highest magnetic gradient. After the magnetically labeled cells and the free fluid are captured at the wall of the tube, the remaining blood is aspirated and discarded. Buffers are added and the magnetic separation is repeated. Next, fluorescent markers for DNA (4',6-diamidino-2-phenylindole: DAPI), cytokeratins 8, 18, and 19 (conjugated to phycoerythrin: PE), and CD45 (conjugated to allophycocyanin: APC) are added and the sample is left to incubate. After another magnetic separation step and more aspiration steps, the remaining 300 μl is transferred to the analysis cartridge, which is placed in a presentation device termed CellTracks Magnest® (see figure 1.4).

This magnest consists of two magnets that create an upward magnetic force, pulling the ferrofluid labeled cells to a cover slip within the cartridge. A simulated image of the magnetic force lines produced by the magnets in the magnest is shown in figure 1.5 [58]. Finally, cells are left to rise to the analysis surface of the cartridge. As can be seen in the figure, the magnets are designed in such a way that cells will move straight up; their distribution across the analysis surface is therefore homogeneous.

1.5 ENUMERATION OF CTC WITH THE CELLTRACKS ANALYZER II®

After the cells are settled, the magnest is placed in the CellTracks Analyzer II, a semi-automatic epi-fluorescence microscope. Employing a mercury arc lamp and a $10\times/0.45\text{NA}$ objective, the whole cartridge is scanned by the Celltracks Analyzer II in four fluorescence channels: channels for

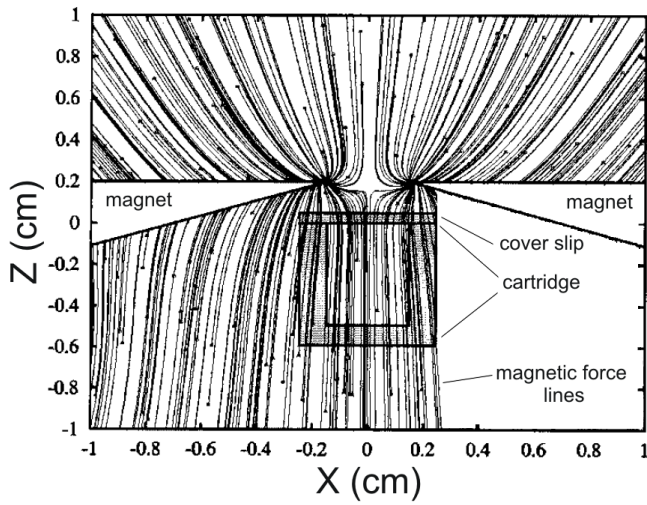


FIGURE 1.5 : Simulation of magnetic force lines as produced by the magnet (adapted from [58]). Cells are pulled straight up towards the cover slip inside the cartridge. The force lines in the lower right part of the figure were erased for viewing purposes.

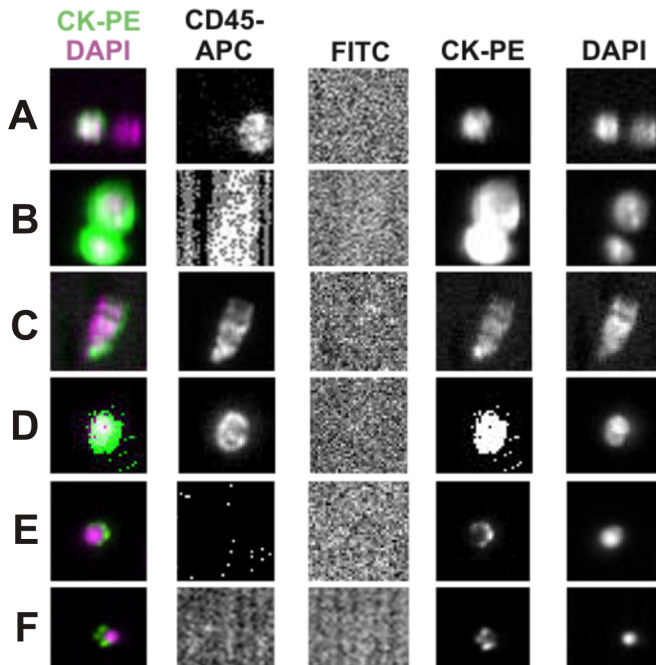


FIGURE 1.6 : CellSearch thumbnail gallery. The CellSearch software presents all objects that are both positive for CK and DAPI to an operator for review.

the aforementioned fluorochromes DNA-DAPI, CK-PE, and CD45-APC, and a fourth channel termed "FITC". This fourth channel may be used for control cells or an extra biomarker, but is generally used to verify if objects are auto-fluorescent (and thus debris). Images are captured using a charge-coupled device (CCD) camera employing pixels of $6.7\ \mu\text{m}$ by $6.7\ \mu\text{m}$. When a scan is complete, the CellSearch software searches for objects that are positively stained for DNA and CK, and creates a thumbnail gallery showing these objects. Figure 1.6 shows an example of such a gallery, in which next to the four fluorescence channels also an overlay of the DAPI and PE channels is shown.

A trained reviewer must now distinguish CTC from debris and leukocytes that were carried over during the enrichment procedure. He or she has a set of rules in determining if an object is a CTC or not, which are shown using a decision tree in figure 1.7. These rules were set and tested by means of preclinical studies [59, 60, 61, 62, 63, 64]. By scoring the cells according to this set of rules, object A from figure 1.6 is a CTC next to a leukocyte. Object B are two bright CTC close together that have some spill-over signal in the CD45-APC channel. Object C fails rule 2 (and also has questionable morphology), and object D fails rule 6. Although these examples are relatively straightforward, not all images are, as is exemplified by the two objects in E and F. These seem to be small cells and are a bit speckled suggesting that they are undergoing apoptosis [65]. When shown to reviewers, it was found that these objects give rise to the highest inter-reviewer variability. It was measured that there exists variability of 4% to 31% for classifying CTC between reviewers (median 14%), and a 7.5% variability between laboratories [66]. The rules for classifying objects are mostly qualitative, because a reviewer cannot easily verify the number of grey levels in the image. Reviewers may therefore be biased by the auto-scaling of these images, which is done purely for viewing purposes. If a bright object is located near a dim object, this dim object may be classified wrongly due to this effect.

1.6 EXPRESSION OF TREATMENT TARGETS ON CTC

Before initiation of a therapy one would like to know whether or not a specific therapy or combination of therapies is going to be effective. The number of treatment options that target specific sites on or in the tumor cells is rapidly increasing. Usually the expression of targets for a therapy is assessed on the primary tumor, because in patients with metastatic disease the tumor cells may have altered and no longer be representative for the tumor. Tumor biopsies are cumbersome for the patient and are not always available. The ability to determine treatment targets on CTC could solve this problem. Fluorescently labeled antibodies that identify treatment targets can be used in the CellSearch system to identify these targets. An

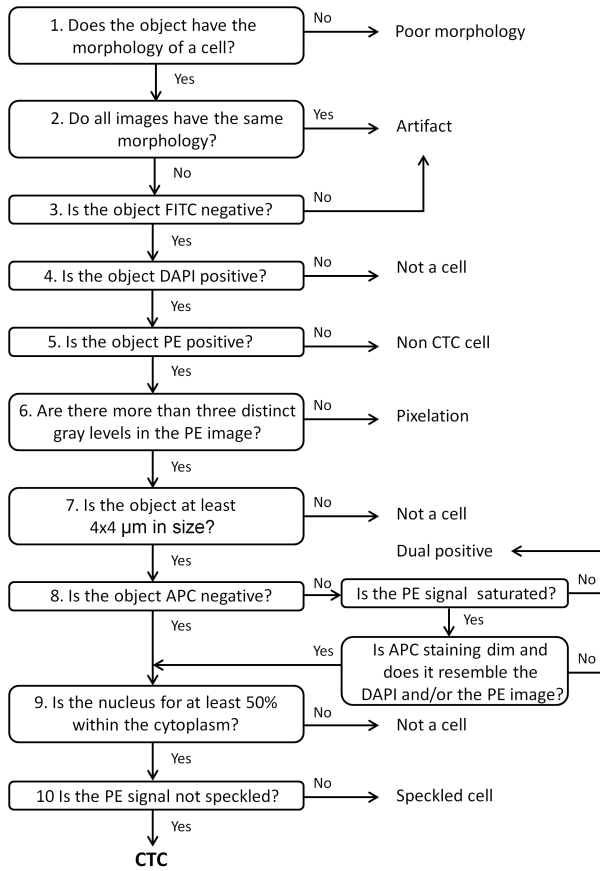


FIGURE 1.7 : Decision tree showing how the trained reviewers classify whether or not an object is a CTC.

example is shown in figure 1.8. Seven cells were detected in the blood sample of this breast cancer patient of which 4 (57%) expressed the Her-2 receptor. These results suggest that treatments targeting Her-2 would only be effective on a portion of the tumor cells. Tumor cells within a tumor are heterogeneous, which holds also true for CTC. To relate the expression of treatment targets on CTC with response to therapy, a quantitative, accurate and reproducible assessment of the expression is needed. This is not feasible by manual review of the images illustrated in Figure 1.8. However, it can be achieved by using a computer algorithm as described in chapter 8 of this thesis.

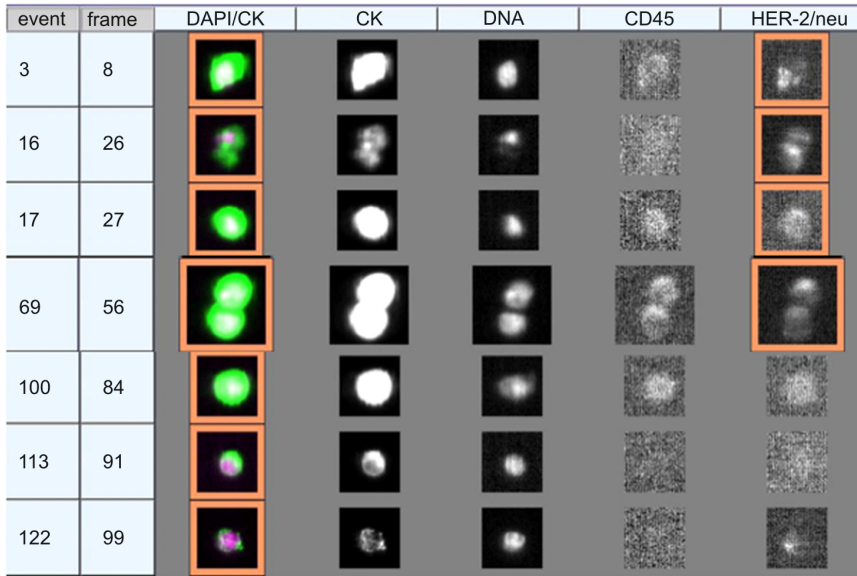


FIGURE 1.8 : Screen shot of the CellTracks Analyzer II showing an image gallery of 7 CTC of a metastatic breast cancer patient. Her-2 labeled with FITC was used in the CellTracks Autoprep in addition to DAPI, CK-PE and CD45-APC. The last column shows the Her 2 expression. The top 4 CTC in the gallery are marked as positive.

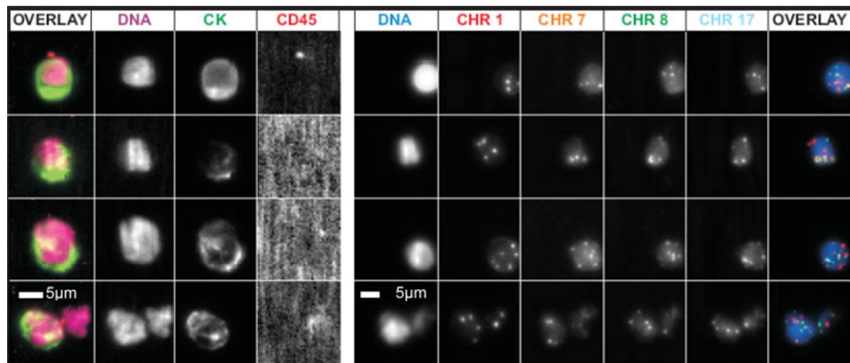


FIGURE 1.9 : Image gallery of CTC after scanning at 10 \times (left four columns), and after subsequent imaging of chromosomes 1, 7, 8, 17, and DAPI at 40 \times . Every row depicts one object, of which an overlay of all channels is created in the first and last column.

1.7 NEXT GENERATION IMAGING SYSTEMS

1.7.1 CELLTRACKS FISHTM

Using the current CellSearch system, is it possible to isolate and enumerate CTC, but at a relatively low resolution: using the $10\times/0.45$ NA objective, a scan of the total surface area can be performed resulting in 144-180 images of four fluorescence channels. The resolution of the objective is according to Abbe's law $\frac{\lambda}{2NA}$, and thus 556 nm using light with a wavelength of 500 nm. However, the sampling density of the system using $6.7\ \mu\text{m}$ pixels and a $10\times$ objective in x and y is 670 nm. Hence, applying the Nyquist-Shannon sampling criterion, the smallest details that can be resolved in an image of this system in x and y are $\sim 1.3\ \mu\text{m}$, which is high enough for counting CTC. If targets, such as chromosomes visualized by fluorescent in situ hybridization (FISH), within cells are to be visualized and characterized, a higher resolution is needed.

For this purpose, a modified version of the CellTracks Analyzer was build, employing a $40\times/0.63\text{NA}$ objective that increased light collection and thus improved resolution. This CellTracks FISH system is able to load $10\times$ magnification images from a regular CTC scan, locate cells that were hybridized with fluorochromes against certain chromosomes and image them at multiple focal depths at $40\times$ magnification. This configuration allowed for successful imaging of centromeres on chromosome 1, 7, 8, and 17 and showed discrepancies in chromosomes copy number between leukocytes and CTC from prostate cancer patients [67]. Figure 1.9 shows CTC imaged with this system in a gallery, next to the original images from the $10\times$ scan. This system was also used to successfully determine ERG, AR, and PTEN gene status in cells from prostate cancer patients [68].

1.7.2 CELLTRACKS TDITM

To scan a complete analysis cartridge at $40\times$ magnification, a new system was built that incorporates three lasers, beam homogenizing optics using micro-lens arrays, a $40\times/0.6\text{NA}$ objective mounted in a piezo z-stage, and a time delay integration (TDI) CCD camera [69, 70]. The setup is depicted schematically in figure 1.10. The TDI camera allows for continuous image acquisition, as the electrons in the camera are shifted from row to row at a speed that matches the speed of the x-y stage in system. It thereby eliminates the time it takes to transfer from one field to another in regular start-stop based systems. Next to decreasing the total imaging time, this system also improves imaging of CD45-APC labeled leukocytes by using a red laser instead of the Hg-lamp that is installed in the regular CellSearch system, which has weak excitation power in the red part of spectrum. Next to the fluorescence channels, bright-field images are recorded which could further identify the imaged objects.

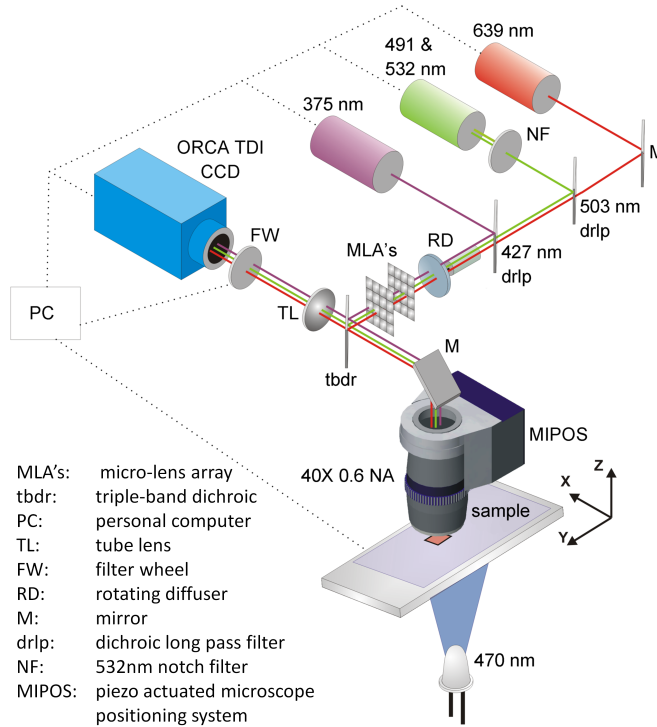


FIGURE 1.10 : Schematic representation of the optical pathway and components of the CellTracks TDI setup. The three lasers produce a bundle that is homogenized into a square mean profile using the MLAs and reflected using the triple band dichroic. The bundle is focused on the sample, which is placed on a x-y translation stage by an objective mounted in z translation piezo system. Emitted fluorescence passes the triple band dichroic and is focused on the TDI CCD camera. This camera is triggered by an encoder signal from the x-y stage to synchronize the time delay integration with sample movement.

1.8 CHALLENGES FOR THE CELLSEARCH SYSTEM

1. The definition of what constitutes a CTC was set before clinical studies were started. This qualitative definition proved to be good enough for discriminating patients based on CTC counts and correlated very well with survival, but may not be the optimal definition. A computer algorithm may find a better, quantitative, definition of objects most dangerous for patients. Furthermore, it is unknown if there is one optimal definition for all CTC or that different definitions should be used to classify objects coming from tumors in the prostate, breast, or colon.

2. Trained reviewers have to classify objects, in the case of CellSearch CTC, CellTracks FISH, and for the measurement of extra markers such as Her-2. Manual classification causes intra- and inter-reviewer variability, as well as inter-laboratory variability. Furthermore, it is time consuming, laborious and costly. This task could be performed by a computer algorithm with higher reproducibility and provide quantitative results.
3. Measuring changes in CTC counts over time within patients is difficult because of the low frequency of CTC and because it is currently unknown what the best method is to measure an increase or decrease in the number of CTC. It is furthermore unknown which definition of a CTC is best suited for measuring significant changes in CTC counts.

The above challenges are addressed in this thesis.

ACKNOWLEDGEMENTS

The author wishes to thank Edzo Klawer for contributions to this chapter.

1.9 REFERENCES

- [1] CBS, “Statline,” 2011.
- [2] T. Ashworth, “A case of cancer in which cells similar to those in the tumours were seen in the blood after death,” *Australian Medical Journal*, vol. 14, pp. 146–147, 1869.
- [3] S. Paget, “The distribution of secondary growths in cancer of the breast,” *The Lancet*, vol. 133, no. 3421, pp. 571–573, 1889.
- [4] L. Norton, “A gompertzian model of human-breast cancer growth,” *Cancer Research*, vol. 48, no. 24, pp. 7067–7071, 1988.
- [5] *AJCC Cancer Staging Manual*. New York: Springer, 7 ed., 2010.
- [6] D. J. Slamon, B. Leyland-Jones, S. Shak, H. Fuchs, V. Paton, A. Bajamonde, T. Fleming, W. Eiermann, J. Wolter, M. Pegram, J. Baselga, and L. Norton, “Use of chemotherapy plus a monoclonal antibody against her2 for metastatic breast cancer that overexpresses her2,” *New England Journal of Medicine*, vol. 344, no. 11, pp. 783–792, 2001.
- [7] M. J. van de Vijver, Y. D. He, L. J. van ’t Veer, H. Dai, A. A. M. Hart, D. W. Voskuil, G. J. Schreiber, J. L. Peterse, C. Roberts, M. J. Marton, M. Parrish, D. Atsma, A. Witteveen, A. Glas, L. Delahaye, T. van der Velde, H. Bartelink, S. Rodenhuis, E. T. Rutgers, S. H. Friend, and R. Bernards, “A gene-expression signature as a predictor of survival in breast cancer,” *New England Journal of Medicine*, vol. 347, no. 25, pp. 1999–2009, 2002.

- [8] S. Paik, S. Shak, G. Tang, C. Kim, J. Baker, M. Cronin, F. L. Baehner, M. G. Walker, D. Watson, T. Park, W. Hiller, E. R. Fisher, D. L. Wickerham, J. Bryant, and N. Wolmark, "A multigene assay to predict recurrence of tamoxifen-treated, node-negative breast cancer," *New England Journal of Medicine*, vol. 351, no. 27, pp. 2817–2826, 2004.
- [9] L. J. van't Veer, H. Y. Dai, M. J. van de Vijver, Y. D. D. He, A. A. M. Hart, M. Mao, H. L. Peterse, K. van der Kooy, M. J. Marton, A. T. Witteveen, G. J. Schreiber, R. M. Kerkhoven, C. Roberts, P. S. Linsley, R. Bernards, and S. H. Friend, "Gene expression profiling predicts clinical outcome of breast cancer," *Nature*, vol. 415, no. 6871, pp. 530–536, 2002.
- [10] J. K. Juckstock, B. K. Rack, T. Zwingers, P. G. M. Hepp, A. Schneewiess, M. W. Beckmann, W. Lichtenegger, H. L. Sommer, K. Pantel, H. Tesch, H. Forstbauer, R. Lorenz, M. Rezai, J. K. Neugebauer, U. Andergassen, K. Friese, and W. Janni, "Prognostic relevance of circulating tumor cells (ctc) before adjuvant chemotherapy in patients with breast cancer: Results of the german success trial," in *ASCO*, (Chicago, IL), American Society of Clinical Oncology, 2011.
- [11] B. Franken, M. R. De Groot, W. J. B. Mastboom, I. Vermes, J. Van der Palen, A. G. J. Tibbe, and L. W. M. M. Terstappen, "Circulating tumor cells, disease recurrence and survival in newly diagnosed breast cancer," in *SABC*, (San Antonio, TX), 2011.
- [12] M. J. Duffy, "Serum tumor markers in breast cancer: Are they of clinical value?," *Clinical Chemistry*, vol. 52, no. 3, pp. 345–351, 2006.
- [13] G. T. Budd, M. Cristofanilli, M. J. Ellis, A. Stopeck, E. Borden, M. C. Miller, J. Matera, M. Repollet, G. V. Doyle, L. Terstappen, and D. F. Hayes, "Circulating tumor cells versus imaging - predicting overall survival in metastatic breast cancer," *Clinical Cancer Research*, vol. 12, no. 21, pp. 6403–6409, 2006.
- [14] D. J. Brenner and E. J. Hall, "Current concepts - computed tomography - an increasing source of radiation exposure," *New England Journal of Medicine*, vol. 357, no. 22, pp. 2277–2284, 2007.
- [15] M. A. Dinan, L. H. Curtis, B. G. Hammill, E. F. Patz, A. P. Abernethy, A. M. Shea, and K. A. Schulman, "Changes in the use and costs of diagnostic imaging among medicare beneficiaries with cancer, 1999-2006," *Jama-Journal of the American Medical Association*, vol. 303, no. 16, pp. 1625–1631, 2010.
- [16] R. W. Carey, P. D. Taft, J. M. Bennett, and S. Kaufman, "Carcinocythemia (carcinoma cell leukemia). an acute leukemia-like picture due to metastatic carcinoma cells," *American Journal of Medicine*, vol. 60, no. 2, pp. 273–8, 1976.
- [17] H. C. Engell, "Cancer cells in the circulating blood; a clinical study on the occurrence of cancer cells in the peripheral blood and in venous blood draining the tumour area at operation," *Acta Chirurgica Scandinavica. Supplementum*, vol. 201, pp. 1–70, 1955.
- [18] M. V. Gallivan and J. J. Lokich, "Carcinocythemia (carcinoma cell leukemia). report of two cases with english literature review," *Cancer*, vol. 53, no. 5, pp. 1100–2, 1984.
- [19] R. L. Myerowitz, P. A. Edwards, and G. P. Sartiano, "Carcinocythemia (carcinoma cell leukemia) due to metastatic carcinoma of breast - report of a case," *Cancer*, vol. 40, no. 6, pp. 3107–3111, 1977.
- [20] L. T. Yam and A. J. Janckila, "Immunocytodiagnosis of carcinocythemia in disseminated breast-cancer," *Acta Cytologica*, vol. 31, no. 1, pp. 68–72, 1987.

- [21] W. J. Allard, J. Matera, M. C. Miller, M. Repollet, M. C. Connelly, C. Rao, A. G. J. Tibbe, J. W. Uhr, and L. Terstappen, "Tumor cells circulate in the peripheral blood of all major carcinomas but not in healthy subjects or patients with nonmalignant diseases," *Clinical Cancer Research*, vol. 10, no. 20, pp. 6897–6904, 2004.
- [22] F. Coumans, S. T. Ligthart, and L. W. M. Terstappen, "All patients with metastatic carcinoma have circulating tumor cells," *submitted*, 2012.
- [23] D. Herlyn, M. Herlyn, A. H. Ross, C. Ernst, B. Atkinson, and H. Koprowski, "Efficient selection of human-tumor growth-inhibiting monoclonal-antibodies," *Journal of Immunological Methods*, vol. 73, no. 1, pp. 157–167, 1984.
- [24] C. G. Rao, D. Chianese, G. V. Doyle, M. C. Miller, T. Russell, R. A. Sanders, and L. Terstappen, "Expression of epithelial cell adhesion molecule in carcinoma cells present in blood and primary and metastatic tumors," *International Journal of Oncology*, vol. 27, no. 1, pp. 49–57, 2005.
- [25] B. Simon, D. K. Podolsky, G. Moldenhauer, K. J. Isselbacher, S. Gattonicelli, and S. J. Brand, "Epithelial glycoprotein is a member of a family of epithelial-cell surface-antigens homologous to nidogen, a matrix adhesion protein," *Proceedings Of The National Academy Of Sciences Of The United States Of America*, vol. 87, no. 7, pp. 2755–2759, 1990.
- [26] R. Moll, W. W. Franke, D. L. Schiller, B. Geiger, and R. Krepler, "The catalog of human cytokeratins - patterns of expression in normal epithelia, tumors and cultured-cells," *Cell*, vol. 31, no. 1, pp. 11–24, 1982. ISI Document Delivery No.: PQ461 Times Cited: 4408 Cited Reference Count: 89 Moll, r franke, ww schiller, dl geiger, b krepler, r Cell press Cambridge.
- [27] C. Alix-Panabieres, J. P. Brouillet, M. Fabbro, H. Yssel, T. Rousset, T. Maudelonde, G. Choquet-Kastylevsky, and J. P. Vendrell, "Characterization and enumeration of cells secreting tumor markers in the peripheral blood of breast cancer patients," *Journal of Immunological Methods*, vol. 299, no. 1-2, pp. 177–188, 2005.
- [28] P. R. C. Gascoyne, J. Noshari, T. J. Anderson, and F. F. Becker, "Isolation of rare cells from cell mixtures by dielectrophoresis," *Electrophoresis*, vol. 30, no. 8, pp. 1388–1398, 2009.
- [29] J. P. Gleghorn, E. D. Pratt, D. Denning, H. Liu, N. H. Bander, S. T. Tagawa, D. M. Nanus, P. A. Giannakakou, and B. J. Kirby, "Capture of circulating tumor cells from whole blood of prostate cancer patients using geometrically enhanced differential immunocapture (gedi) and a prostate-specific antibody," *Lab on a Chip*, vol. 10, no. 1, pp. 27–29, 2010.
- [30] H. J. Kahn, A. Presta, L. Y. Yang, J. Blondal, M. Trudeau, L. Lickley, C. Holloway, D. R. McCready, D. Maclean, and A. Marks, "Enumeration of circulating tumor cells in the blood of breast cancer patients after filtration enrichment: correlation with disease stage," *Breast Cancer Research and Treatment*, vol. 86, no. 3, pp. 237–247, 2004.
- [31] R. T. Krivacic, A. Ladanyi, D. N. Curry, H. B. Hsieh, P. Kuhn, D. E. Bergsrud, J. F. Kepros, T. Barbera, M. Y. Ho, L. B. Chen, R. A. Lerner, and R. H. Bruce, "A rare-cell detector for cancer," *Proceedings Of The National Academy Of Sciences Of The United States Of America*, vol. 101, no. 29, pp. 10501–10504, 2004.
- [32] G. Vona, A. Sabile, M. Louha, V. Sitruk, S. Romana, K. Schutze, F. Capron, D. Franco, M. Pazzagli, M. Vekemans, B. Lacour, C. Brechot, and P. Paterlini-Brechot, "Isolation by size of epithelial tumor cells - a new method for the immunomorphological and molecular characterization of circulating tumor cells," *American Journal of Pathology*, vol. 156, no. 1, pp. 57–63, 2000.

- [33] R. Konigsberg, E. Obermayr, G. Bises, G. Pfeiler, M. Gneist, F. Wrba, M. de Santis, R. Zeillinger, M. Hudec, and C. Dittrich, "Detection of epcam positive and negative circulating tumor cells in metastatic breast cancer patients," *Acta Oncologica*, vol. 50, no. 5, pp. 700–710, 2011.
- [34] S. J. Tan, R. L. Lakshmi, P. F. Chen, W. T. Lim, L. Yobas, and C. T. Lim, "Versatile label free biochip for the detection of circulating tumor cells from peripheral blood in cancer patients," *Biosensors & Bioelectronics*, vol. 26, no. 4, pp. 1701–1705, 2010.
- [35] S. Riethdorf, H. Fritsche, V. Muller, T. Rau, C. Schindibeck, B. Rack, W. Janni, C. Coith, K. Beck, F. Janicke, S. Jackson, T. Gornet, M. Cristofanilli, and K. Pantel, "Detection of circulating tumor cells in peripheral blood of patients with metastatic breast cancer: A validation study of the cellsearch system," *Clinical Cancer Research*, vol. 13, no. 3, pp. 920–928, 2007.
- [36] S. J. Cohen, C. J. A. Punt, N. Iannotti, B. H. Savidman, K. D. Sabbath, N. Y. Gabrail, J. Picus, M. Morse, E. Mitchell, M. C. Miller, G. V. Doyle, H. Tissing, L. Terstappen, and N. J. Meropol, "Relationship of circulating tumor cells to tumor response, progression-free survival, and overall survival in patients with metastatic colorectal cancer," *Journal of Clinical Oncology*, vol. 26, no. 19, pp. 3213–3221, 2008.
- [37] M. Cristofanilli, G. T. Budd, M. J. Ellis, A. Stopeck, J. Matera, M. C. Miller, J. M. Reuben, G. V. Doyle, W. J. Allard, L. Terstappen, and D. F. Hayes, "Circulating tumor cells, disease progression, and survival in metastatic breast cancer," *New England Journal of Medicine*, vol. 351, no. 8, pp. 781–791, 2004.
- [38] J. S. de Bono, H. I. Scher, R. B. Montgomery, C. Parker, M. C. Miller, H. Tissing, G. V. Doyle, L. Terstappen, K. J. Pienta, and D. Raghavan, "Circulating tumor cells predict survival benefit from treatment in metastatic castration-resistant prostate cancer," *Clinical Cancer Research*, vol. 14, no. 19, pp. 6302–6309, 2008.
- [39] J. Y. Pierga, D. Hajage, T. Bachelot, S. Delaloge, E. Brain, M. Campone, V. Dieras, E. Rolland, L. Mignot, C. Mathiot, and F. C. Bidard, "High independent prognostic and predictive value of circulating tumor cells compared with serum tumor markers in a large prospective trial in first-line chemotherapy for metastatic breast cancer patients," *Annals of Oncology*, 2011.
- [40] H. I. Scher, X. Y. Jia, J. S. de Bono, M. Fleisher, K. J. Pienta, D. Raghavan, and G. Heller, "Circulating tumour cells as prognostic markers in progressive, castration-resistant prostate cancer: a reanalysis of immc38 trial data," *Lancet Oncology*, vol. 10, no. 3, pp. 233–239, 2009.
- [41] F. A. W. Coumans, C. J. M. Doggen, G. Attard, J. S. de Bono, and L. W. M. M. Terstappen, "All circulating epcam+ck+cd45- objects predict overall survival in castration-resistant prostate cancer," *Annals of Oncology*, vol. 21, no. 9, pp. 1851–7, 2010.
- [42] M. G. Krebs, R. Sloane, L. Priest, L. Lancashire, J. M. Hou, A. Greystoke, T. H. Ward, R. Ferraldeschi, A. Hughes, G. Clack, M. Ranson, C. Dive, and F. H. Blackhall, "Evaluation and prognostic significance of circulating tumor cells in patients with non-small-cell lung cancer," *Journal of Clinical Oncology*, vol. 29, no. 12, pp. 1556–1563, 2011.
- [43] F. Tanaka, K. Yoneda, N. Kondo, M. Hashimoto, T. Takuwa, S. Matsumoto, Y. Okumura, S. Rahman, N. Tsubota, T. Tsujimura, K. Kuribayashi, K. Fukuoka, T. Nakano, and S. Hasegawa, "Circulating tumor cell as a diagnostic marker in primary lung cancer," *Clinical Cancer Research*, vol. 15, no. 22, pp. 6980–6986, 2009.

- [44] M. S. Khan, T. Tsigani, M. Rashid, J. S. Rabouhans, D. Yu, V. L. Tu, M. Caplin, and T. Meyer, "Circulating tumor cells and epcam expression in neuroendocrine tumors," *Clinical Cancer Research*, vol. 17, no. 2, pp. 337–345, 2011.
- [45] S. Matsusaka, K. Chin, M. Ogura, M. Suenaga, E. Shinozaki, Y. Mishima, Y. Terui, N. Mizunuma, and K. Hatake, "Circulating tumor cells as a surrogate marker for determining response to chemotherapy in patients with advanced gastric cancer," *Cancer Science*, vol. 101, no. 4, pp. 1067–1071, 2010.
- [46] M. Rink, F. K. H. Chun, S. Minner, M. Friedrich, O. Mauermann, H. Heinzer, H. Huland, M. Fisch, K. Pantel, and S. Riethdorf, "Detection of circulating tumour cells in peripheral blood of patients with advanced non-metastatic bladder cancer," *BJU International*, vol. 107, no. 10, pp. 1668–1675, 2011.
- [47] A. Poveda, S. B. Kaye, R. McCormack, S. B. Wang, T. Parekh, D. Ricci, C. A. Lebedinsky, J. C. Tercero, P. Zintl, and B. J. Monk, "Circulating tumor cells predict progression free survival and overall survival in patients with relapsed/recurrent advanced ovarian cancer," *Gynecologic Oncology*, vol. 122, no. 3, pp. 567–572, 2011.
- [48] F. C. Bidard, C. Mathiot, A. Degeorges, M. C. Etienne-Grimaldi, R. Delva, X. Pivot, C. Veyret, L. Bergougnoux, P. de Cremoux, G. Milano, and J. Y. Pierga, "Clinical value of circulating endothelial cells and circulating tumor cells in metastatic breast cancer patients treated first line with bevacizumab and chemotherapy," *Annals of Oncology*, vol. 21, no. 9, pp. 1765–1771, 2010.
- [49] S. Meng, D. Tripathy, S. Shete, R. Ashfaq, H. Saboorian, B. Haley, E. Frenkel, D. Euhus, M. Leitch, C. Osborne, E. Clifford, S. Perkins, P. Beitsch, A. Khan, L. Morrison, D. Herlyn, L. Terstappen, N. Lane, J. Wang, and J. Uhr, "upar and her-2 gene status in individual breast cancer cells from blood and tissues," *Proceedings Of The National Academy Of Sciences Of The United States Of America*, vol. 103, no. 46, pp. 17361–17365, 2006.
- [50] S. D. Meng, D. Tripathy, S. Shete, R. Ashfaq, B. Haley, S. Perkins, P. Beitsch, A. Khan, D. Euhus, C. Osborne, E. Frenkel, S. Hoover, M. Leitch, E. Clifford, E. Vitetta, L. Morrison, D. Herlyn, L. Terstappen, T. Fleming, T. Fehm, T. Tucker, N. Lane, J. Q. Wang, and J. Uhr, "Her-2 gene amplification can be acquired as breast cancer progresses," *Proceedings Of The National Academy Of Sciences Of The United States Of America*, vol. 101, no. 25, pp. 9393–9398, 2004.
- [51] S. Riethdorf, V. Muller, L. L. Zhang, T. Rau, S. Loibl, M. Komor, M. Roller, J. Huober, T. Fehm, I. Schrader, J. Hilfrich, F. Holms, H. Tesch, H. Eidtmann, M. Untch, G. von Minckwitz, and K. Pantel, "Detection and her2 expression of circulating tumor cells: Prospective monitoring in breast cancer patients treated in the neoadjuvant geparquattro trial," *Clinical Cancer Research*, vol. 16, no. 9, pp. 2634–2645, 2010.
- [52] J. S. de Bono, G. Attard, A. Adjei, M. N. Pollak, P. C. Fong, P. Haluska, L. Roberts, C. Melvin, M. Repollet, D. Chianese, M. Connely, L. Terstappen, and A. Gualberto, "Potential applications for circulating tumor cells expressing the insulin-like growth factor-i receptor," *Clinical Cancer Research*, vol. 13, no. 12, pp. 3611–3616, 2007.
- [53] A. M. Sieuwerts, J. Kraan, J. Bolt-de Vries, P. van der Spoel, B. Mostert, J. W. M. Martens, J. W. Gratama, S. Sleijfer, and J. A. Foekens, "Molecular characterization of circulating tumor cells in large quantities of contaminating leukocytes by a multiplex real-time pcr," *Breast Cancer Research and Treatment*, vol. 118, no. 3, pp. 455–468, 2009.
- [54] E. A. Punnoose, S. K. Atwal, J. M. Spoerke, H. Savage, A. Pandita, R. F. Yeh, A. Pirzkall, B. M. Fine, L. C. Amler, D. S. Chen, and M. R. Lackner, "Molecular biomarker analyses using circulating tumor cells," *Plos One*, vol. 5, no. 9, 2010.

- [55] A. M. Sieuwerts, B. Mostert, J. Bolt-de Vries, D. Peeters, F. E. de Jongh, J. M. L. Stouthard, L. Y. Dirix, P. A. van Dam, A. Van Galen, V. de Weerd, J. Kraan, P. van der Spoel, R. Ramirez-Moreno, C. H. M. van Deurzen, M. Smid, J. X. Yu, J. Jiang, Y. Wang, J. W. Gratama, S. Sleijfer, J. A. Foekens, and J. W. M. Martens, "mrna and microrna expression profiles in circulating tumor cells and primary tumors of metastatic breast cancer patients," *Clinical Cancer Research*, vol. 17, no. 11, pp. 3600–3618, 2011.
- [56] D. A. Smirnov, D. R. Zweitzig, B. W. Foulk, M. C. Miller, G. V. Doyle, K. J. Pienta, N. J. Meropol, L. M. Weiner, S. J. Cohen, J. G. Moreno, M. C. Connelly, L. Terstappen, and S. M. O'Hara, "Global gene expression profiling of circulating tumor cells," *Cancer Research*, vol. 65, no. 12, pp. 4993–4997, 2005.
- [57] P. A. Liberti, C. G. Rao, and L. Terstappen, "Optimization of ferrofluids and protocols for the enrichment of breast tumor cells in blood," *Journal of Magnetism and Magnetic Materials*, vol. 225, no. 1-2, pp. 301–307, 2001.
- [58] A. G. J. Tibbe, B. G. de Grooth, J. Greve, G. J. Dolan, C. Rao, and L. Terstappen, "Magnetic field design for selecting and aligning immunomagnetic labeled cells," *Cytometry*, vol. 47, no. 3, pp. 163–172, 2002.
- [59] D. F. Hayes, T. M. Walker, B. Singh, E. S. Vitetta, J. W. Uhr, S. Gross, C. Rao, G. V. Doyle, and L. Terstappen, "Monitoring expression of her-2 on circulating epithelial cells in patients with advanced breast cancer," *International Journal of Oncology*, vol. 21, no. 5, pp. 1111–1117, 2002.
- [60] M. Kagan, D. Howard, T. Bendele, J. Mayes, J. Silvia, M. Repollet, J. Doyle, J. Al-lard, N. Tu, T. Bui, T. Russell, C. Rao, M. Hermann, H. Rutner, and L. Terstappen, "A sample preparation and analysis system for identification of circulating tumor cells," *Journal of Clinical Ligand Assay*, vol. 25, no. 1, pp. 104–110, 2002.
- [61] J. G. Moreno, S. M. O'Hara, S. Gross, G. Doyle, H. Fritsche, L. G. Gomella, and L. Terstappen, "Changes in circulating carcinoma cells in patients with metastatic prostate cancer correlate with disease status," *Urology*, vol. 58, no. 3, pp. 386–392, 2001.
- [62] E. Racila, D. Euhus, A. J. Weiss, C. Rao, J. McConnell, L. Terstappen, and J. W. Uhr, "Detection and characterization of carcinoma cells in the blood," *Proceedings Of The National Academy Of Sciences Of The United States Of America*, vol. 95, no. 8, pp. 4589–4594, 1998.
- [63] L. Terstappen, C. Rao, S. Gross, V. Kotelnikov, E. Racilla, J. Uhr, and A. Weiss, "Flowcytometry - principles and feasibility in transfusion medicine. enumeration of epithelial derived tumor cells in peripheral blood," *Vox Sanguinis*, vol. 74, pp. 269–274, 1998.
- [64] L. Terstappen, C. Rao, S. Gross, and A. J. Weiss, "Peripheral blood tumor cell load reflects the clinical activity of the disease in patients with carcinoma of the breast," *International Journal of Oncology*, vol. 17, no. 3, pp. 573–578, 2000.
- [65] C. J. Larson, J. G. Moreno, K. J. Pienta, S. Gross, M. Repollet, S. M. O'Hara, T. Russell, and L. Terstappen, "Apoptosis of circulating tumor cells in prostate cancer patients," *Cytometry Part A*, vol. 62A, no. 1, pp. 46–53, 2004.
- [66] J. Kraan, S. Sleijfer, M. H. Strijbos, M. Ignatiadis, D. Peeters, J. Y. Pierga, F. Farace, S. Riethdorf, T. Fehm, L. Zorzino, A. G. Tibbe, M. Maestro, R. Gisbert-Criado, G. Denton, J. S. de Bono, C. Dive, J. A. Foekens, and J. W. Gratama, "External quality assurance of circulating tumor cell enumeration using the cellsearch((r)) system: A feasibility study," *Cytometry B Clin Cytom*, vol. 80B, no. 2, pp. 112–118, 2010.

- [67] J. F. Swennenhuis, A. G. Tibbe, R. Levink, R. C. Sipkema, and L. W. Terstappen, "Characterization of circulating tumor cells by fluorescence in situ hybridization," *Cytometry Part A*, vol. 75, no. 6, pp. 520–7, 2009.
- [68] G. Attard, J. F. Swennenhuis, D. Olmos, A. H. M. Reid, E. Vickers, R. A'Hern, R. Levink, F. Coumans, J. Moreira, R. Riisnaes, N. B. Oommen, G. Hawche, C. Jameson, E. Thompson, R. Sipkema, C. P. Carden, C. Parker, D. Dearnaley, S. B. Kaye, C. S. Cooper, A. Molina, M. E. Cox, L. Terstappen, and J. S. de Bono, "Characterization of *erg*, *ar* and *pten* gene status in circulating tumor cells from patients with castration-resistant prostate cancer," *Cancer Research*, vol. 69, no. 7, pp. 2912–2918, 2009.
- [69] T. M. Scholtens, F. Schreuder, S. T. Ligthart, J. F. Swennenhuis, A. G. J. Tibbe, J. Greve, and L. Terstappen, "Celltracks tdi: An image cytometer for cell characterization," *Cytometry Part A*, vol. 79A, no. 3, pp. 203–213, 2011.
- [70] T. M. Scholtens, F. Schreuder, S. T. Ligthart, J. F. Swennenhuis, J. Greve, and L. W. M. M. Terstappen, "Automated identification of circulating tumor cells by image cytometry," *Cytometry Part A*, vol. 81A, pp. 138–148, 2011.

SIMULATION AND CALIBRATION OF SPECTRAL IMAGING METHODS

Sjoerd T. Ligthart, Cees Otto, Jan Greve, and Leon W.M.M. Terstappen

ABSTRACT

Introduction: Multi- and hyper-spectral fluorescence imaging methods have been available for decades. It is however unknown how different methods compare quantitatively, i.e. which method offers the best trade-off between throughput in photons per second and resolution in nanometer per measured spectral band. We simulated four spectral imaging methods and calibrated real life spectral systems for quantitative comparison. **Materials and Methods:** Four spectral imaging methods, based on dichroic mirrors, a prism, a liquid crystal tunable filter (LCTF), and an interferometer were simulated for measuring centromere-like objects. Combinations of five quantum dots and DAPI were used in a labeling scheme to cover all 24 different chromosomes. Linear unmixing and classification of objects was performed on the simulated images. The total integration time was set such that the classification error was 5% for every method. Six real life systems were calibrated using a LED calibration board; LED current was plotted against camera output for various wavelength ranges. **Results:** Integration time needed for 5% classification error of objects was 0.24, 3.6, 7.8, and 11.5 s, for the dichroic mirror, prism, LCTF, and interferometer methods, respectively. Camera outputs showed a wide range of sensitivities, even those who employed the same CCD sensors. For measuring a large range of wavelengths with high spectral resolution, the interferometer outperformed the LCTF system by a factor of two. **Conclusions:** A designated set of dichroic mirrors is to be preferred for fast measurements of combination of quantum dots. However, a LCTF system offers great flexibility, while prism and interferometer-based setups are preferred if very high spectral resolution is required.

2.1 INTRODUCTION

Multi- and hyper-spectral fluorescence imaging have been around for several decades after the discovery of fluorescent specific probes coupled to antibodies [1]. This discovery was preceded by the introduction of microscopy using optimized illumination and ultimately the discovery of the fluorescence phenomenon [2, 3]. Traditionally, fluorescent signals are measured using an epi-fluorescence setup, i.e. using an microscope objective for focusing the excitation beam coming from a light source onto the sample as well as collecting the emitted fluorescence photons for transfer to a charge-coupled device (CCD), a photo-multiplier tube (PMT), or the eye of the observer via ocular lenses. Pixels on a CCD-chip convert the incoming photons to electrons which are converted into digital units to create a digital image of the sample. As the resolution of the image is partially dependent on the size of the pixels (termed sampling in the image plane), a CCD chip that detects one band of color is preferred. Color pixels next to each other would lower the effective resolution of the camera. Hence, a spectral selection method has to be employed before the photons reach the CCD chip in order to distinguish between different fluorochromes.

A method is called multi-spectral if it can deliver a fixed number of wavelength ranges -termed spectral bands hereafter- to the camera. Current commercially available microscopes that are employing sets of interference filters and dichroic mirrors, usually mounted in so-called filter cubes, fit into this category [4]. Filter cubes are currently available for a wide variety of fluorescent probes; however, the number of filter cubes that can be installed in a microscopic system is limited. A spectral selection method is called hyper-spectral (although this classification is somewhat arbitrary) if it is able to transmit a large number of selected spectral bands to the camera. Examples of hyper-spectral imaging methods are Fourier based systems, dispersive or prism-based systems, and tunable filters by means of liquid crystals or acousto-optic modulators [5, 6, 7]. Each of these methods has its qualitative advantages and disadvantages in for instance ease of use, compatibility with fluorescent probes, and cost.

All these methods have been commercially available. However, it is currently unknown how these methods compare quantitatively in terms of transmission of photons per second versus spectral resolution in nanometer per bandwidth. Qualitatively, dedicated methods that are tuned, i.e. have high transmittance on specific wavelengths, generally can perform the task faster than methods that are very flexible and thus versatile. In this article, we present results from simulations of various spectral imaging methods in order to conclude which method provides the highest throughput and thus results in the fastest classification of simulated objects with low error using a commonly used set of fluorescent probes. We performed this simulation in order to optimize measurements on fluorescent in situ hybridized probes on centromeres of chromosomes, which are small and usually round signals

of $\sim 1 \mu\text{m}^2$. If all centromeres in a cell are to be measured in a fast and sensitive way, spectral overlap should be minimal, because the signals will be closely spaced in the nucleus. Finally, we provide measurements using a calibration source, to verify the simulations and to provide a quantitative comparison between these spectral imaging methods.

2.2 MATERIALS AND METHODS

2.2.1 INTRODUCTION

Four methods were simulated in Matlab 2009a (Mathworks, Natick, MA): i) an interferometer or Fourier spectral imager; ii) a dispersion or prism-based spectral imager; iii) a band selecting or liquid crystal tunable filter (LCTF) imaging system; iv) a dichroic mirror (DM) spectral imager. Each method will be explained in short below. However, before starting simulating it is important to provide a figure of merit: what do we want to measure to test the spectral imaging methods and why. Next, we recognize which fluorochrome combination is to be measured. Fluorescence emission theory and the spectral imaging methods are explained, and how these methods were optimized. Finally, it is shown how we set up a measure for the best spectral method and how the simulation was validated using actual microscopes.

2.2.2 TESTING THE SPECTRAL IMAGING METHODS: THE CHALLENGE

First, it is important to recognize how the methods are going to be compared. Spectral imaging methods may be rated in terms of transmission of photons per second, number of detectable colors, or spectral resolution: bandwidth per detector element. We chose a practical approach: ultimately, each method is to be used for measuring fluorescent objects and usually cells. To extend our DNA research in circulating tumor cells [8], we need to measure interphase fluorescence in situ hybridization (FISH) probes -termed dots from here on- on centromeres with low classification error. Therefore we stated the challenge for the spectral imaging methods as follows: which method can measure 24 different combinations of fluorochromes on FISH dots in the fastest way with low classification error. This challenge involves a trade-off between high throughput of photons, spectral resolution, and spectral overlap within detector elements.

2.2.3 FLUOROCHROME COMBINATIONS

In our view, a fluorochrome combination consisting of a nuclear dye and five FISH dot dyes should be sufficient for the task of measuring 24 different centromeres: the nuclear dye is used for locating the total area of the nucleus of the cell and the five dot dyes give us a total of $2^5 - 1 = 31$ combinations.

24 combinations are necessary to measure all centromeres of 22 pairs of autosomes and one pair of sex chromosomes. If all are to be measured at once, the emission spectra should be separated enough spectrally to be classified with low error, i.e. crosstalk or spectral overlap of signal from one fluorochrome in the detection channel of another should be minimal. Ideally, we would like to excite all these fluorochromes at the same wavelength. Only one laser would be required and all emission photons could be collected in parallel instead of in series of measurements.

In literature, different examples can be found using 6 or more combinations of fluorochromes for measuring either metaphase or interphase chromosomes [9, 10, 11]. However, all these combinations require multiple excitation sources and thus have to be measured either in a series of recording sessions or with very strictly defined band pass excitation optics. We chose a system consisting of the nuclear dye 4',6-diamidino-2-phenylindole (DAPI) and five quantum dots (Qdots), with in parentheses their full width half maximum: 525 (34) nm, 565 (31) nm, 605 (22) nm, 655 (29) nm, and 705 (66) nm. The emission peaks of the Qdots are almost equally spaced over the 500–750 nm range, DAPI emission peaks at 465 nm. All the fluorochromes can be excited in the blue region with for instance a 375 or 405 nm blue laser, molar absorption coefficients at 375 nm ranges from 540,000 $\text{M}^{-1}\text{cm}^{-1}$ for Qdot 525 to 10,500,000 $\text{M}^{-1}\text{cm}^{-1}$ for Qdot 705. Furthermore, Qdots are known to be less affected by bleaching and have a high photostability [12, 13, 14, 15, 16]. It is known that Qdots have a property termed blinking, i.e. they switch between dark and bright states [17]. Blinking of single quantum dots can be on a timescale in the order of seconds, and can thus have a serious influence on the measured fluorescence in single-molecule applications. However, in our application, the signal of many Qdots (~ 1000) is averaged, and we therefore assume a negligible influence on intensity. Table 2.1 shows the combinations necessary to reach 24 combinations of chromosomes: at most three Qdots together with DAPI per chromosome are required.

2.2.4 FLUORESCENCE EXCITATION AND EMISSION THEORY

The aforementioned spectral imaging methods were simulated by means of the number of photons excited by a our current system which was described elsewhere [18]. A 8 mW 375 nm laser was used in combination with a $40\times/0.6\text{NA}$ air immersion objective to calculate the number of photons arriving at the CCD camera from the sample as described below [19]. The fluorescence lifetime of DAPI is very short, however we recognize that the lifetime of Qdots ranges from 5 ns up to 100 ns. Nevertheless, we assume that each fluorochrome is excited in a linear and isotropic way, and is not saturated. The excitation energy absorbed by every fluorescent molecule is given by a probability defined by the absorption cross section σ_A [cm^2 molecule $^{-1}$].

TABLE 2.1 : Combinations of dyes used in the simulation. Different combinations of five Qdots were used in combination with the nuclear staining DAPI to create 24 combinations for labeling of all chromosomes.

combi	DAPI	Qdot ^a 525	Qdot565	Qdot605	Qdot655	Qdot705	# of probes
1	×	×					2
2	×		×				
3	×			×			
4	×				×		
5	×					×	
6	×	×	×				
7	×	×		×			
8	×	×			×		
9	×	×				×	
10	×		×	×			3
11	×		×		×		
12	×		×			×	
13	×			×	×		
14	×			×		×	
15	×				×	×	
16	×	×	×	×			4
17	×	×	×		×		
18	×	×	×			×	
19	×	×		×	×		
20	×	×		×		×	
21	×	×			×	×	
22	×		×	×	×		
23	×		×	×		×	
24	×			×	×	×	

^a Quantum dot.

$$\sigma_A = 10^3 \ln(10) \epsilon_\lambda N_{Avo} \quad (2.1)$$

where ϵ_λ represents the molar absorption coefficient [$M^{-1} \text{ cm}^{-1}$] and N_{Avo} is Avogadro's number. Every fluorochrome converts the absorbed light with a quantum efficiency Q_{em} thereby creating an isotropic light source with spectral density of radiant intensity I_λ^{em} [$\text{W sr}^{-1} \text{ nm}^{-1} \text{ molecule}^{-1}$]:

$$I_\lambda^{em} = E_{exc} \lambda_{exc} \sigma_A q(\lambda) Q_{em} / 4\pi \quad (2.2)$$

in which E_{exc} is the irradiance [W cm^{-2}], λ_{exc} [nm] is the excitation wavelength and $q(\lambda)$ is the spectral relative quantum yield distribution function according to the normalized emission spectrum $g(\lambda)$ by

$$q(\lambda) = \frac{g(\lambda)}{\int g(\lambda) \lambda d\lambda} \quad (2.3)$$

Because the source of emission could be viewed as an isotropic point source, the solid angle of the collection of the objective Ω is given by

$$\Omega = 2\pi \left(1 - \sqrt{1 - \left(\frac{NA}{n} \right)^2} \right) \quad (2.4)$$

in which NA is the numerical aperture of the objective and n the maximum refractive index in the medium between object and objective. Capturing the total transmittance of the optical components between sample and camera in τ_{em} , and the sensitivity of the spectral detector in Q_{CCD} , the total photoelectron flux [photoelectrons s^{-1} molecule $^{-1}$] in the CCD camera per pixel is then calculated by

$$\varphi_{CCD} = \frac{1}{4} 10^{16} (E_{ex} \lambda_{ex} \sigma_A) Q_{em} [\Omega \tau_{em}] \int q(\lambda) Q_{CCD}(\lambda) d\lambda \quad (2.5)$$

A rather lengthy expression in which the first constant term represents one over the energy of a photon at wavelength λ [J nm $^{-1}$]. The last part of the expression is integrated over the wavelength band as imposed by the spectral imaging method. We assume that the transmission of the optics and objective and the quantum yield Q_{em} are independent of the emission wavelength. Thus, the integral part of this equation is different in each spectral imaging method. Every method records a spectral cube consisting of two spatial axes and one spectral axis (x, y, λ) , although in different ways:

1. The Fourier method: the emission enters an interferometer which records an interferogram of the signal at the detector. Hence, information of all the wavelengths is recorded at the same time. The optical path length of the signal at the detector is varied by rotating a Sagnac interferometer, which consists of two mirrors and a beam splitter. Afterwards, the frequency signal is Fourier transformed to retrieve the wavelength information. Example: SpectraCube[®] by Applied Spectral Imaging (Carlsbad, CA).
2. The prism method: this method uses the CCD in a different way: one axis is used for spatial information, the other axis is used for spectral information. The emission signal enters a prism and is dispersed on the CCD. Excitation should be performed by slit illumination; the sample should be scanned to retrieve the second spatial axis. Example: P.A.R.I.S.S.[®] by Lightform Inc (NJ, U.S.).
3. the LCTF method: the emission light enters a liquid crystal tunable filter, which is tuned on a central wavelength, usually retrieving wavelength information ± 5 or ± 10 nm around this central wavelength. In a sequence of recordings, the tuned wavelength shifts across the desired wavelength spectrum. Example: Nuance[®] by Caliper Sciences (Hopkinton, MA).

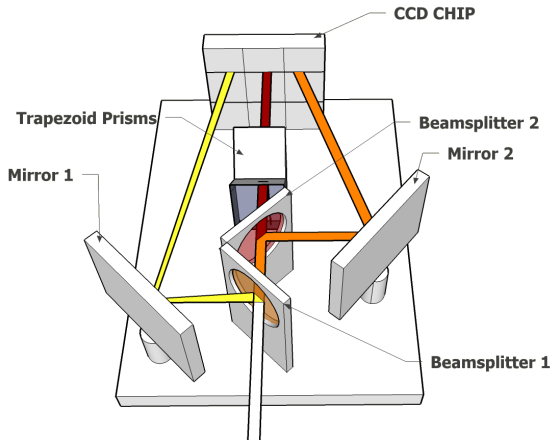


FIGURE 2.1 : Schematic example of a DM imaging setup. The white beam, coming from the objective, is split in three parts via dichroic mirrors and imaged on the camera using mirrors. Two trapezoid prisms are inserted in the path of the central beam to create equal path lengths for all three beams.

4. the DM method: Two CCD cameras which operate in time delay integration (TDI, see [18]) mode are illuminated in parallel. The emission light is split into six parts using dichroic mirrors, of which an example is given in figure 2.1. Two of the setups in figure 2.1 are needed if six fluorochromes are to be measured. This method could be compared to measuring the signal using filter cubes in a series of measurements, with the difference that the DM method works in parallel. Example: ImageStream[®] by Amnis Corporation (Seattle, WA).

2.2.5 TEST IMAGE FOR SIMULATION

Of all these methods the properties of current systems as built by current industrial companies were looked up and applied as well as possible. A 400×1024 pixel test image consisting of all 25 combinations of fluorochromes (24 Qdot combinations and DAPI alone) was created. In the test image, the objects that were created were 29 pixels ($\sim 1 \times 1 \mu\text{m}$ at $40\times$ magnification) in size and round in shape, which is typical for small objects such as FISH probes. In total, 720 objects were created that could conveniently fit in this space. In order to get maximum performance out of all of the methods, each method was optimized as described below.

2.2.6 INSERTING A ND FILTER FOR USING THE DM METHOD

The DM spectral imaging method is quite straightforward in its implementation based on formula 2.5. Two cameras are used in parallel to image the field of view in TDI mode; therefore their integration time should match. Qdots that are imaged on camera two have a substantial higher extinction

coefficient (factor ~ 20) and are therefore much brighter. In order to prevent saturation on camera two, while camera one still records noisy images, several neutral density (ND) filters were tested in order to decrease signal intensities at camera two.

2.2.7 OPTIMIZING THE LCTF METHOD

The LCTF method records images at bands that are 20 nm full width half maximum (FWHM) wide. The number of spectral bands that are recorded along the total range of the test images can be changed: the images could be recorded 1 nm apart or, for instance, 20 nm apart. In this way the total measuring time can be adjusted. By recording the test image at a different number of images -thus creating a different spectral resolution- the optimal spectral resolution was determined. A better resolution usually decreases the classification error, but requires longer recording times. Therefore, the total imaging time was set to 10 s, while varying the number of spectral bands. In this first step, the spectral bands were spread evenly over the total spectral range of 420–720 nm and the classification error was determined for each number of spectral bands. This classification error was plotted against the number of spectral bands to visualize a trend. In the next step, different combinations of spectral bands were tested in a brute force way in the optimal region to determine the definite optimal number of spectral imaging windows. As a last optimization step, it was recognized that it would be advantageous to distribute the integration time in a different manner over the chosen spectral windows: Qdots which emit at higher wavelengths have a higher extinction coefficient and are thus brighter. Furthermore, the transmission of the LCTF is also wavelength dependant. It was therefore determined which percentage of integration time should go to each spectral window to measure each channel with an equal signal to noise ratio (SNR).

2.2.8 OPTIMIZING THE FOURIER METHOD

The basis of the Fourier method is a Sagnac interferometer which splits the beam coming from the sample by means of a 50/50 beam splitter. The beams travels opposite path which are confined by two mirrors. A small optical path difference (OPD) is imposed by a slight tilt of one of the mirrors or the beam splitter. Two beams of a monochromatic source will interfere at the camera by means of the following relation [20]

$$I_{CCD}(\delta) = 4I_0 \cos^2 \left(\frac{\delta}{2} \right) \quad (2.6)$$

in which I_0 represents the emitted beam from the sample serving as input and δ is the phase difference imposed by the OPD. Rotating the interferometer varies this phase difference, and the number of steps and total angle result in a number of spectral images. The number of steps

that is required has a lower bound in the Nyquist criterion: the smallest spectral detail has to be sampled by at least two points to prevent aliasing. Higher sampling results in more images and a higher spectral resolution, but decreases the signal to noise ratio if the total integration time of the measurement is kept constant. After imaging, the so-called interferogram should be Fourier transformed in order to retrieve the original intensities present at the sample. A fast fourier transform (FFT) function was used, applying a Hann window for apodization. Using this window, the interferogram is smoothed by setting a passband in the shape of a bell-shaped function. Using the test image and a constant total integration time, the step size was varied to measure the error in classifying the Qdot combinations. In this way, the optimal step size was determined.

2.2.9 OPTIMIZING THE PRISM METHOD

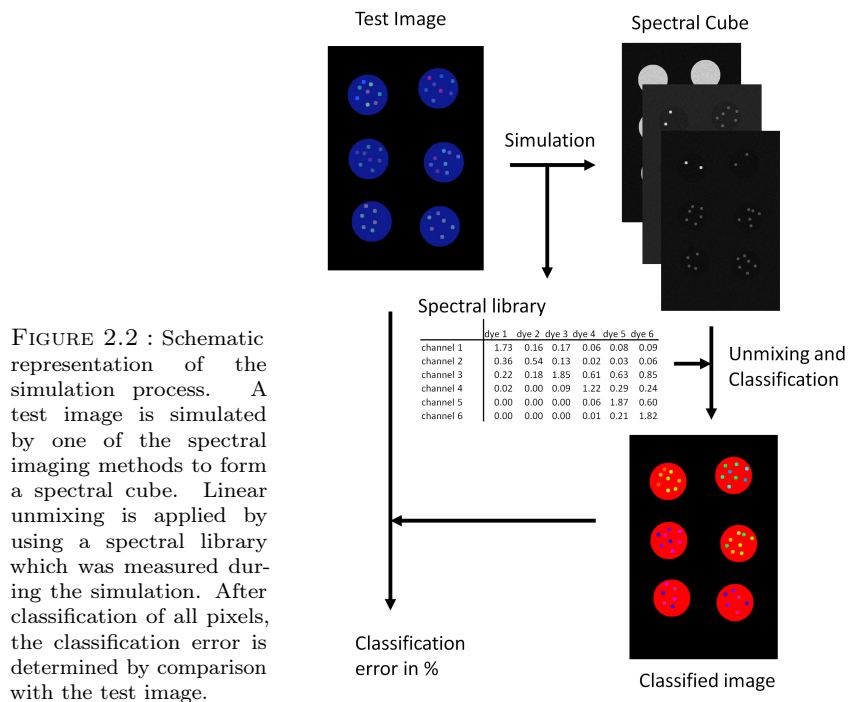
Prism based spectral imaging methods are based on a dispersive element in the emission path which spreads the spectral information within the signal from red to blue. Thus, one axis of the CCD camera is used for spectral information, the other axis is used for one of the spatial axes. In this setup, a slit scanning illumination profile should be used and the sample should be scanned to gather information about the second spatial axis. Using the empirical Sellmeier equation, the dispersion of a particular wavelength through a prism consisting of a certain type of glass can be derived.

$$n^2(\lambda) = 1 + \frac{B_1\lambda^2}{\lambda^2 - C_1} + \frac{B_2\lambda^2}{\lambda^2 - C_1} + \frac{B_3\lambda^2}{\lambda^2 - C_1} \quad (2.7)$$

Were n is the refractive index, and $B1$ to $B3$, and $C1$ to $C3$ are experimentally determined glass-specific Sellmeier coefficients. Using this equation, the wavelength dependent dispersion can be calculated. The deviation angle Θ_{dev} that results from the incident angle Θ_i on the prism after dispersion is then given by

$$\Theta_{dev} = \Theta_i + \sin^{-1} \left[\sin\alpha \sqrt{n^2 - \sin^2(\Theta_i)} - \sin(\Theta_i)\cos(\alpha) \right] - \alpha \quad (2.8)$$

in which α is the apex angle of the prism. This dispersed bundle of light is imaged onto a camera, blue light being more dispersed than red light. Spectral resolution at the lower wavelengths of the imaged spectrum is thus higher than that of the higher wavelengths, when using a common prism. If a 1280×1024 pixel camera is used, the dispersion results in a very high spectral resolution, which could be higher than needed for our application. Therefore, binning of CCD pixels in the spectral direction in order to create "super pixels" may be applied to trade some spectral resolution for SNR. Binned super pixels resulting from 2, 4, 8, 16, 32, and 64 regular pixels were created to optimize this trade-off.



2.2.10 CLASSIFICATION OF THE TEST IMAGE AFTER LINEAR UNMIXING

The simulation is illustrated in Figure 2.2. For each method, the 400×1024 test image was converted to an intensity stack by means of the simulation described above. Next to this intensity image, all fluorochromes were simulated separately to create a spectral library. Poisson noise was added to each intensity image, based on the photon count. Readout noise and a noise component from the laser were also added. The spectral cube that was recorded using each method was then sized $400 \times 1024 \times \lambda$, in which λ is given by the number of spectral bands. Next, the spectral cube was processed by a regular linear unmixing algorithm, based on a pseudo-inverse of the measured spectral library, to eliminate spectral crosstalk (see for example [21]). This unmixing procedure resulted in a $400 \times 1024 \times 6$ intensity image, a stack with an image containing information for each fluorochrome. Next, the intensities were normalized to 1 in the λ -direction.

For each pixel (a $1 \times 1 \times 6$ matrix), the Euclidean distances to predefined points in six dimensional space were derived, which represented the normalized fluorochrome combinations. The combination that had the minimum distance to the pixel was chosen. The classified pixels were grouped as objects, if the number of equally classified pixels in a region was between 20 and 40. In this way, we allowed for $\sim 33\%$ misclassified pixels per object.

We set the integration time in every method such that 5% of the 720 objects were allowed to be misclassified compared to the objects in the test image. These misclassification thresholds were set at arbitrary levels. However, in this way we could compare the methods in a transparent way, because a 0% classification error as a result of an experiment does not clearly tell us if improvement is still possible.

2.2.11 MEASURING CAMERAS AND SETUPS USING A CALIBRATION BOARD

To validate the simulation, measurements were performed on different spectral imaging systems. To exclude variance in sample quality, an electronic calibration board (which was kindly provided by the QI group, Delft University of Technology) was used. This board contains six LEDs that could be tuned at certain intensities very precisely. The input (measured in amperes through the LED) could therefore be controlled for every spectral imager in a controlled way. In this way, output (measured here in intensity counts per millisecond) from different setups can be compared in a quantitative way. Using a $20\times/0.4\text{NA}$ objective in a wide-field fluorescence setup in which a DAPI filter cube was fitted, images were recorded of the LEDs. This filter-cube has a 420 nm long-pass emission filter which passes the light of all LEDs. The LEDs used were emitting at a central wavelength of 460 nm (blue LED), 520 nm (green LED), and 640 nm (red LED) according to the specs. The calibration measurements on each of the setups was performed as follows:

1. One integration time for all measurements of the specific setup was set after determination of the LED that resulted in the highest intensity at a current of 300 mA. The integration time was set such that 3/4 of the maximum of the intensity range was reached. Saturation of the camera was thus avoided.
2. A black image was measured with the LED switched off at the designated integration time.
3. For all three LEDs, intensities were measured for different LED currents between 50 and 300 mA in steps of 50 mA.
4. Images were loaded in Matlab. For each measurement, the central and brightest part of the LED surface was selected in the image.
5. The mean intensity in this region was determined and normalized to 1 ms integration time after subtracting the black image.

Each selected region contained at least $3\cdot 10^5$ pixels. The cameras that were tested were the Hamamatsu c4742-95NR (Hamamatsu Photonics, Hamamatsu City, Japan), Meade DSI III (Meade Instruments, Irvine, CA), Smartek GC 1392M (Smartek d.o.o., Cakovec, Croatia), CRI Nuance

TABLE 2.2 : Camera specifications.

Camera	Hamamatsu C4742-95NR	Smartek GC 1392M CRI Nuance ASI SpectraCube Meade DSI III pro	SBIG ST 402ME
Sensor	Hamamatsu	Sony ICX285	KAF-0402/E/MF
Cooled(C)	yes (-20)	no / yes (0) / yes (-20) / no	yes (0)
Number of pixels	1280×1024	1344×1040	756×510
Pixel size (μm)	6.7×6.7	6.5×6.5	9×9
full well depth (e ⁻)	13300	16000	100000
Readout noise per pixel (e ⁻)	13	8	17
Dark current per pixel (e ⁻ /s)	0.05	0.01	1
QE ^a at 400 nm (%)	50	50	45
QE at 650 nm (%)	20	41	72

^a Quantum efficiency.

(Caliper Life Sciences, Hopkinton, MA), and the Spectracube (Carlsbad, CA). The latter two cameras employ a LCTF filter and an interferometer for spectral separation, respectively. Table 2.2 shows details of the CCD chips and performance as given by the manufacturers. The SpectraCube system was installed on a Leica DMXRA epi-fluorescence microscope, the other cameras were all installed on the same Nikon Eclipse E400 epi-fluorescence microscope. Unfortunately, a similar epi-fluorescence system using a prism was not available at the time of this study.

2.3 RESULTS

2.3.1 OPTIMIZING THE SPECTRAL IMAGING METHODS

Figure 2.3 shows the classification error of the FISH objects as function of the number of spectral LCTF bands. The bands were evenly spread over the range from 420–720 nm and the total integration was kept constant at 10 s. From this figure, it was clear that the optimum number of bands was in the range of 6 to 8 bands; it is thus best to record few bands with high SNR instead of more bands with lower SNR. The error is large at 7 bands, which was due to the chosen spread of bands over the wavelength range, which turned out to be a bad choice for seven bands.

To find the optimum setting within the range of 6–8 bands, it was therefore necessary to vary the spread of the bands over the total spectral range, instead of spreading them evenly over the whole range. Therefore, all combinations of putting the spectral bands 20 nm apart were tested for the case of 6 to 8 bands, resulting in 3003, 3432 and 3003 combinations respectively. To ensure every band has an optimal SNR, the total integration

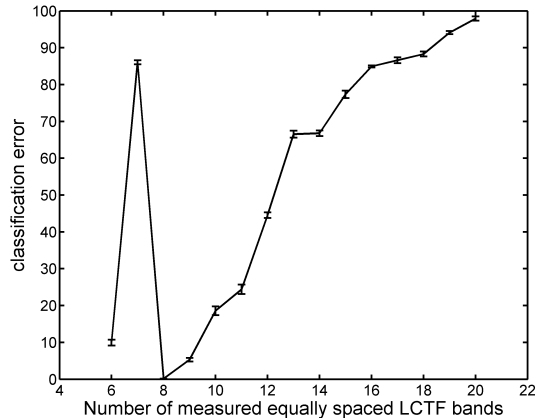


FIGURE 2.3 : The classification error as function of the number of spectral LCTF bands. The bands are spaced equally over the total bandwidth of 420–720 nm. The integration time for all bands together was set at 10 sec, which results in a lower integration time per band when more bands are measured. Error is the standard deviation of 10 experiments.

time was divided among the bands unevenly. In this way, bands with weaker fluorescence signals would contribute the same number of photons as stronger signals in other bands. The combination of 7 bands that produced the lowest classification error had center wavelengths at 460, 480, 520, 560, 600, 640, and 700 nm. The integration times used for each band were 24.1, 19.4, 8.9, 3.6, 1.3, 1.6, and 1.1 s, respectively.

The dispersion method records images with high spectral resolution per pixel, which varies from 0.16 nm at 400 nm to 1.18 nm at 850 nm for a CCD chip that is 1024 pixels in the spectral dimension and 1280 in the spatial dimension. The preferred method of binning was 8×1 on the CCD chip. This method lowers the total imaging time by a factor of 8, while still preserving enough spectral information to keep the classification error low.

Optimization of the Fourier method was performed by varying the number of rotation steps by the interferometer. Each step shifts the interference fringes over the CCD. In our test images, the lowest wavelength present is 400 nm, and the fringes produced by this wavelength should be sampled on the CCD by at least two points. The optimal sampling step was just above the Nyquist criterion: further increase of the sampling density did not contribute to a better classification within the same total integration time.

The DM method was equipped with a ND filter that would even out the intensity between the cameras: a ND 4 filter (75% of signal is absorbed at the camera recording the brightest Qdots) turned out to be necessary.

TABLE 2.3 : Results of simulation of spectral imaging methods. Total integration time needed for every method to arrive at the same classification error of 5% when classifying the test image containing 720 objects.

Spectral method	Integration time (s)
LCTF ^a (using 7 bands)	7.80
Prism (using 8×1 binning)	3.60
Interferometer (sampling just above Nyquist)	11.52
DM ^b method using 2 cameras (ND ^c 4 filter)	0.24

^a Liquid crystal tunable filter, ^b dichroic mirror, ^c neutral density

2.3.2 TOTAL INTEGRATION TIME FOR SIMULATION OF THE TEST IMAGE

Next, we compared the optimal configuration for every spectral imaging method for processing time by allowing a classification error of 5% using the 400×1024 pixel test image containing 720 objects. Results of the comparison are shown Table 2.3. Integration times were corrected for a field of view of 1280×1024 pixels: the integration time found using the DM method was multiplied by three, because this method only accepts field of views that are 400×1024 pixels in size.

2.3.3 MEASURING THROUGHPUT IN MICROSCOPES

All cameras were exposed to photons emitted from the same LEDs. Integration times were set such that the cameras were not saturated, and the 3/4 of the intensity range could be covered when the LEDs were set to 300 mA. However, the integration time of the SBIG camera could not be set low enough to avoid saturation, therefore the currents of the LEDs were varied from 30 to 55 mA in steps of 5 mA. Extra amplification of the signal by means of a gain were turned off where possible. Figure 2.4, panel A, B, and C show the results of the single shot cameras. The normalized intensity per ms is plotted against the current through the LEDs, on a log-scale. All cameras, but the MEADE camera showed linear behaviour in this region. The SpectraCube camera could not be included in this comparison, as there exists no single shot option on this system: for an interferometric measurement to be successful, a range of optical path differences has to be measured.

Next, measurements of a spectral range from 420–720 nm were performed using the CRI Nuance and the SpectraCube. Currents through the three LEDs were again varied from 50 to 300 mA in steps of 50 mA, results are shown in figure 2.4, panel D, E, and F. The measurements included 31 steps at 5 ms for the Nuance camera, and 128 steps at 2 ms for the SpectraCube camera. Total acquisition time for a measurement was 16 s for the Nuance camera, 17 s for the SpectraCube (which included the Fourier transformation). The intensity values per image were normalized on the total integration time that was used for the whole measurement. This

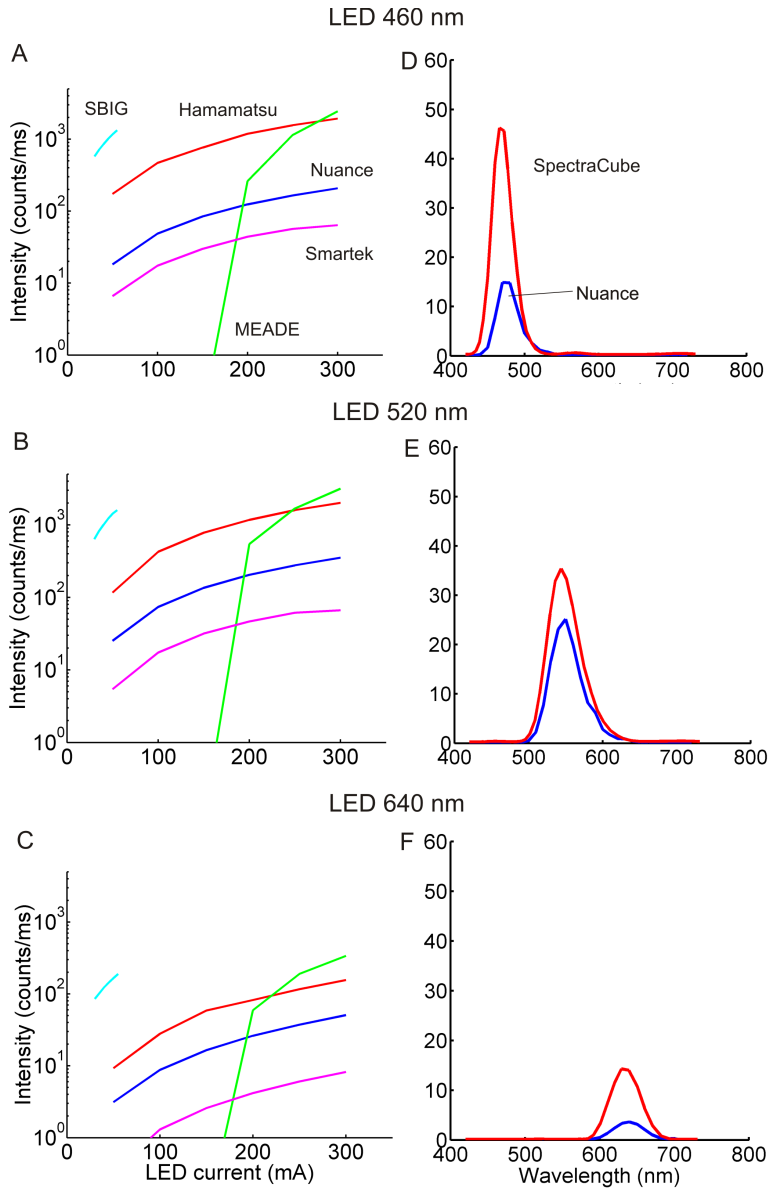


FIGURE 2.4 : Panel A, B, C: Intensity counts for a series of cameras mounted on a Nikon Eclipse E400 microscope as function of the current through a 460, 520, and 640 nm LED. Panel D, E, F: Intensity versus wavelength for the Nuance and SpectraCube methods at a current of 300 mA through LEDs of 460, 520, and 640 nm.

was necessary, as the measurements from the SpectraCube resulted in 41 images (not 128) after Fourier transformation. A per image comparison would therefore not yield a fair result. To measure if there existed a time dependant factor in the LED intensity, the first intensity measurement from the red LED was repeated after all other measurements were done, the variability between those two measurements was 0.25 counts, measured over $3.2 \cdot 10^6$ pixels.

2.4 DISCUSSION

Fluorescence spectral imaging has been around for several decades. Still, reported measurements and images are usually measured using a different systems, and are therefore difficult to compare. Calibration tools are necessary to avoid false claims of increased sensitivity and to compare systems in a quantitative way. In this article, spectral imaging systems were characterized by simulation and the results were compared to measurements on real setups for a potential powerful application: the application of several quantum dots in one assay in order to view a high number of different FISH probes. Quantum dots have narrow emission spectra (i.e. spectral overlap is low), and can be excited at the same wavelength, for instance using 375 or 405 nm lasers. However, we noticed in early experiments that the actual usage of quantum dots on cells proved to be difficult: large variations in staining efficiency were seen. Therefore we compared the spectral imaging methods using a LED calibration board. The emission power of LEDs can be tuned accurately by varying the current through the LEDs, thereby creating the ideal source for calibration of the emission path of a microscope system.

Simulation results in table 2.3 show that the DM spectral imaging method is about one order of magnitude more sensitive than the other three methods. This dedicated, tailored method for measuring a collection of quantum dots has the highest transmission. By using two parallel CCD cameras, photons from the whole spectral range can reach the cameras at the same time. This high measuring speed comes at a price: the system is very inflexible for other combinations of fluorescent probes. Looking at the other three systems, the prism method has the best throughput, although a slit-scan system has to be employed to use this system. Measuring a whole field of view therefore always requires a translational stage that moves, or scanning mirrors, and therefore may cause some motion artifacts. The LCTF system is very flexible: before a scan is performed all spectral windows can be determined exactly, thereby minimizing the recording of redundant wavelengths. A lot of photons are thrown away by selecting the LCTF window, therefore the speed of measuring is a factor two lower than the dispersion method.

The Fourier method performed a little worse than the LCTF method.

This method is suited for usage in combination with a large range of fluorochromes in one sample. No single shot option is available, because for accurate measuring of the optical path difference, a minimum number of images is always required to reach the Nyquist criterion for sampling. When looking at the simulated images, the Fourier system had difficulty with classifying the DAPI signal. It performs well on small emission peaks compared to other methods, because photons are measured from those peaks during the whole scan. A more broader wavelength spectrum, like DAPI, seems to suffer more from noise/background signal after classification of the pixels. Of note, we used a linear unmixing algorithm to classify pixels in this simulation. This approach might be slightly unfavorable for methods which record at high spectral resolution. More sophisticated methods of unmixing, such as support vector machines or neural networks, may be applied if there is more spectral information available. An algorithm could then be trained on a setup with possibly an improved accuracy in classification of pixels.

When we measured different cameras and spectral imaging methods using the LED calibration tool, large differences were seen (see figure 2.4), even between cameras employing the same CCD chip. Some cameras apply a standard gain that cannot be turned off, as is revealed by the difference in slope of the graphs. For instance, the MEADE camera applies a high gain that delivers high signals per ms, at the cost of dynamic range. The Hamamatsu camera performed the best: it has a high dynamic range, while returning high intensities per ms. Two setups were tested on a wide spectral range as can be seen in figure 2.4, panel D, E, and F: the LCTF Nuance system and the SpectraCube Fourier system. Measurements on both systems took ~ 17 s. The results suggest that the Fourier system is to be preferred if a large spectral range is to be measured at high spectral resolution: it delivers a higher signal by a factor of two.

However, if only three different signals are to be measured, the LCTF can be set such to only measure at the designated wavelengths. The LCTF values of figure 2.4A can for example be compared with the values of figure 2.4D. In our simulation it was derived that the optimal number of spectral windows for the LCTF was one more than the number of fluorochromes. If, as in the simulation, 6 different signals should have been measured, 7 LCTF bands are needed. This will increase the throughput of the LCTF by a factor of ~ 4 . Looking at figure 2.4 it can then be concluded that the only gain to use the LCTF would be in the order of 1.5 to 2. The Hamamatsu camera has a factor ten better throughput using the blue LED; this factor is six and three when using the green and red LED, respectively. If the Hamamatsu camera would be used for a DM setup, for instance using six cameras in parallel, it was derived that its throughput would be a factor ~ 5 higher than the LCTF setup.

Comparison of simulation and measurements on the basis of integration times show that the differences are smaller in measurements than was predicted in simulation. The DM method would still be the fastest, but

only by a factor of five instead of ~ 20 which came out of the simulation. The difference between LCTF and Fourier was similar in simulation and measurement: about a factor 1.5. However, when taking overhead time into account, the LCTF and especially the Fourier method perform worse than the DM method. The time it takes to switch from wavelength bands in the LCTF and to perform a Fourier transform are substantial. In conclusion, it is clear that if dedicated setup for fast scanning of a high number of fluorochromes is needed, the DM setup is to be preferred. It is tailored for six spectral windows and therefore does not record any redundant spectral information. If a setup is to be used for flexible essays, a proof of principle may be acquired in an easy way using the LCTF setup. This setup is user friendly and flexible, although some tuning of the ratios of integration time has to be done for maximum performance. Using the LED calibration board gives us the tool for compare setups in quantitative manner. It can be used for comparison of setups, but also as a quality control system for the throughput of one system in time.

ACKNOWLEDGEMENTS

The authors would like to thank Ronald Ligteringen and Ted Young from the Quantitative Imaging Group at Delft University of Technology for providing the LED calibration board and their kind assistance in measuring on the SpectraCube instrument.

2.5 REFERENCES

- [1] A. H. Coons, H. J. Creech, and R. N. Jones, "Immunological properties of an antibody containing a fluorescent group," *Proceedings of the Society for Experimental Biology and Medicine. Society for Experimental Biology and Medicine (New York, N.Y.)*, vol. 47, no. 2, pp. 200–202, 1941.
- [2] A. Koehler, "Ein neues beleuchtungsverfahren fuer mikrographische zwecke," *Zeitschrift fuer wissenschaftliche Mikroskopie und fuer Mikroskopische Technik*, vol. 10, no. 4, pp. 433–440, 1893.
- [3] G. G. Stokes, "On the change of refrangibility of light," *Philosophical Transactions of the Royal Society of London*, vol. 142, pp. 463–562, 1852.
- [4] P. H. Lissberger, "Properties of all-dielectric interference filters .1. a new method of calculation," *Journal of the Optical Society of America*, vol. 49, no. 2, pp. 121–125, 1959. ISI Document Delivery No.: WF915 Times Cited: 28 Cited Reference Count: 6 Lissberger, ph Amer inst physics Woodbury.
- [5] Z. Malik, D. Cabib, R. A. Buckwald, A. Talmi, Y. Garini, and S. G. Lipson, "Fourier transform multipixel spectroscopy for quantitative cytology," *Journal Of Microscopy-Oxford*, vol. 182, pp. 133–140, 1996.
- [6] J. M. Lerner, "Imaging spectrometer fundamentals for researchers in the biosciences - a tutorial," *Cytometry Part A*, vol. 69A, no. 8, pp. 712–734, 2006. ISI Document Delivery No.: 089WE Times Cited: 12 Cited Reference Count: 30 Lerner, Jeremy M. Wiley-liss Hoboken.

- [7] H. R. Morris, C. C. Hoyt, and P. J. Treado, "Imaging spectrometers for fluorescence and raman microscopy - acoustooptic and liquid-crystal tunable filters," *Applied Spectroscopy*, vol. 48, no. 7, pp. 857–866, 1994.
- [8] J. F. Swennenhuis, A. G. Tibbe, R. Levink, R. C. Sipkema, and L. W. Terstappen, "Characterization of circulating tumor cells by fluorescence in situ hybridization," *Cytometry Part A*, vol. 75, no. 6, pp. 520–7, 2009.
- [9] E. Schrock, S. duManoir, T. Veldman, B. Schoell, J. Wienberg, M. A. Ferguson-Smith, Y. Ning, D. H. Ledbetter, I. BarAm, D. Soenksen, Y. Garini, and T. Ried, "Multicolor spectral karyotyping of human chromosomes," *Science*, vol. 273, no. 5274, pp. 494–497, 1996.
- [10] M. R. Speicher, S. G. Ballard, and D. C. Ward, "Karyotyping human chromosomes by combinatorial multi-fluor fish," *Nature Genetics*, vol. 12, no. 4, pp. 368–375, 1996.
- [11] J. Fung, W. Hyun, P. Dandekar, R. A. Pedersen, and H. U. G. Weier, "Spectral imaging in preconception/preimplantation genetic diagnosis of aneuploidy: Multicolor, multichromosome screening of single cells," *Journal Of Assisted Reproduction And Genetics*, vol. 15, no. 5, pp. 323–330, 1998.
- [12] X. Y. Wu, H. J. Liu, J. Q. Liu, K. N. Haley, J. A. Treadway, J. P. Larson, N. F. Ge, F. Peale, and M. P. Bruchez, "Immunofluorescent labeling of cancer marker her2 and other cellular targets with semiconductor quantum dots (vol 21, pg 41, 2003)," *Nature Biotechnology*, vol. 21, no. 4, pp. 452–452, 2003.
- [13] D. R. Larson, W. R. Zipfel, R. M. Williams, S. W. Clark, M. P. Bruchez, F. W. Wise, and W. W. Webb, "Water-soluble quantum dots for multiphoton fluorescence imaging in vivo," *Science*, vol. 300, no. 5624, pp. 1434–1436, 2003.
- [14] L. A. Bentolila and S. Weiss, "Single-step multicolor fluorescence in situ hybridization using semiconductor quantum dot-dna conjugates," *Cell Biochemistry and Biophysics*, vol. 45, no. 1, pp. 59–70, 2006.
- [15] L. Ma, S. M. Wu, J. Huang, Y. Ding, D. W. Pang, and L. J. Li, "Fluorescence in situ hybridization (fish) on maize metaphase chromosomes with quantum dot-labeled dna conjugates," *Chromosoma*, vol. 117, no. 2, pp. 181–187, 2008.
- [16] S. M. Wu, X. Zha, Z. L. Zhang, H. Y. Xie, Z. Q. Tian, J. Peng, Z. X. Lu, D. W. Pang, and Z. X. Xie, "Quantum-dot-labeled dna probes for fluorescence in situ hybridization (fish) in the microorganism escherichia coli," *Chemphyschem*, vol. 7, no. 5, pp. 1062–1067, 2006.
- [17] M. Nirmal, B. O. Dabbousi, M. G. Bawendi, J. J. Macklin, J. K. Trautman, T. D. Harris, and L. E. Brus, "Fluorescence intermittency in single cadmium selenide nanocrystals," *Nature*, vol. 383, no. 6603, pp. 802–804, 1996. ISI Document Delivery No.: VQ144 Times Cited: 934 Cited Reference Count: 20 Nirmal, M Dabbousi, BO Bawendi, MG Macklin, JJ Trautman, JK Harris, TD Brus, LE Macmillan magazines ltd London.
- [18] T. M. Scholtens, F. Schreuder, S. T. Ligthart, J. F. Swennenhuis, A. G. J. Tibbe, J. Greve, and L. Terstappen, "Celltracks tdi: An image cytometer for cell characterization," *Cytometry Part A*, vol. 79A, no. 3, pp. 203–213, 2011.
- [19] T. M. Jovin and D. J. Arndtjovin, "Luminescence digital imaging microscopy," *Annual Review of Biophysics and Biophysical Chemistry*, vol. 18, pp. 271–308, 1989.
- [20] E. Hecht, *Optics*. Reading, MA: Addison Wesley Longman, Inc., 3 ed., 1998.
- [21] H. Tsurui, H. Nishimura, S. Hattori, S. Hirose, K. Okumura, and T. Shirai, "Seven-color fluorescence imaging of tissue samples based on fourier spectroscopy and singular value decomposition," *Journal Of Histochemistry & Cytochemistry*, vol. 48, no. 5, pp. 653–662, 2000.

FISH PROBE COUNTING IN
CIRCULATING TUMOR CELLS¹

Sjoerd T. Ligthart, Joost F. Swennenhuis, Jan Greve, and Leon W.M.M. Terstappen

3.1 INTRODUCTION

Presence of tumor cells in blood of patients with metastatic carcinomas has been associated with poor progression free and overall survival [1, 2, 3]. Assessment of treatment targets on circulating tumor cells (CTC) before initiation of therapy may provide means to guide therapy [4, 5, 6, 7, 8, 9, 10]. Characterization of CTC can be performed by fluorescence in situ hybridization (FISH), which has been used to prove that CTC are indeed malignant [11, 10], and that gene amplifications, deletions and translocations related to certain therapies can be detected [4, 7, 8].

CTC are extremely rare in most patients: 1–10/7.5 ml of blood [12], among about 50 million leukocytes and 50 billion erythrocytes within that volume. For accurate characterization of CTC it is thus of utmost importance that no cell loss is incurred and that the error in the interpretation of the results is kept to a minimum. The CellSearch[®] system is the only clinically validated system for counting CTC [12, 13]. It is based firstly on immunomagnetic enrichment of the blood sample using the epithelial cell adhesion molecule (EpCAM) labeled with ferrofluid. A second step consists of labeling the enriched sample with fluorescent dyes for the nucleus and cytokeratins 8, 18, and 19; CD45 labeling is added to recognize leukocytes.

¹published as a chapter in: Cytokeratins - Tools in Oncology, Rijeka: Intech - open access publisher, editor: Dr. Gerhard Hamilton, ISBN 978-953-308-10-7

In a third step, fluorescence images are recorded of the enriched and labeled sample. CTC candidates are presented to a trained reviewer who distinguishes CTC from debris and leukocytes. Limitation of the CellSearch system is that only CTC expressing both EpCAM and CK 8, 18 and or 19 will be detected.

Recently, we have developed a semi-automated method for FISH analysis of CTC after they have been identified by the CellSearch system. Interpretation of FISH results is however encumbered by apoptosis of CTC, which is observed frequently. In addition, counting of FISH dots can be tiring and subjective, and thus likely results in differences in intra-reviewer and inter-reviewer interpretation. Automation of counting of these FISH signals -termed FISH dots hereafter- could resolve these challenges. Other work has been done in the field of automated counting of FISH dots: on cell lines [14], blood from healthy individuals [15], cells from amniotic fluid [16], and on tissue [17]. An excellent overview of methods is available [18]. However, to our knowledge, automated dot counting has never been applied in samples containing CTC. The nuclei and dots of these cells are extremely heterogeneous in shape and intensity, and thus difficult to score, even by reviewers. Therefore, we investigated the error in counting FISH dots, and evaluated different methods to count FISH dots by a computer algorithm.

3.2 METHOD

3.2.1 PATIENT SAMPLES

A prospective multicenter clinical trial that evaluated the utility of counting CTC for predicting response to therapy, progression-free survival, and overall survival in metastatic castration-resistant prostate cancer patients was conducted [3]. A total of 65 clinical centers throughout the United States and Europe participated in this study after formal institutional review board approval. All patients were required to provide written informed consent. Blood was collected before starting a new treatment and at monthly intervals prior to the next cycle of therapy.

3.2.2 SAMPLE PREPARATION FOR CTC ENUMERATION

Blood samples were drawn into 10 ml evacuated blood draw tubes, maintained at room temperature, and processed within 96 hours of collection. The CellSearch System (Veridex LLC, Raritan, NJ) consists of the CellTracks Autoprep[®], CellTracks Magnest[®], CellSearch Epithelial Cell Kit[®] and the CellTracks Analyzer II[®]. The CellSearch Epithelial Cell Kit contains: -EpCAM labeled ferrofluids; -staining reagents 4',6-diamidino-2-phenylindole (DAPI), CD45-allophycocyan (CD45-APC), cytokeratin 8, 18 phycoerythrin, cytokeratin 19 phycoerythrin (CK-PE); -buffers to enhance cell capture [19], cell permeabilization and cell fixation. The CellTracks Au-

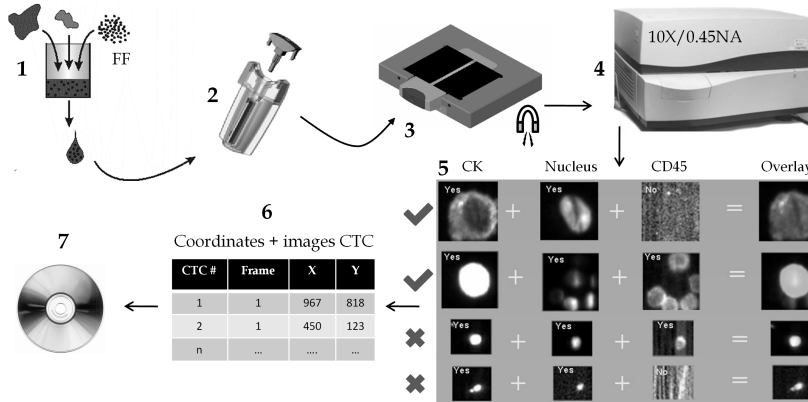


FIGURE 3.1 : Image acquisition using the CellSearch system. 1) Ferrofluid and staining reagents are added to the blood sample. 2) After immuno-magnetic enrichment and fluorescent labeling, the sample is inserted into a cartridge. 3) The cartridge is inserted into a Magnest to distribute the labeled cells over the analysis surface. 4) The cartridge is scanned at 10× in the CellTracks Analyzer II. 5) CK + DAPI+ objects are presented to the reviewer for CTC selection. 6) Coordinates of selected CTC are collected. 7) The coordinates and images from the scan are saved on a CD or DVD.

toprep immuno-magnetically enriches cells expressing EpCAM from 7.5 ml of blood, fluorescently labels the enriched cells with DAPI, CD45-APC and CK-PE, and re-suspends the cells in the cartridge placed in the CellTracks Magnest. The design of the magnets guides the magnetically labeled cells to the analysis surface [20].

3.2.3 DATA ACQUISITION FOR CTC ENUMERATION

The CellTracks Magnest containing the cartridge was placed inside the CellTracks Analyzer II, a semi-automated fluorescence-based microscopy system that acquires images using a 10×/0.45NA objective with filters for DAPI, FITC, PE, and APC to cover the complete surface area of the cartridge. The CellSearch software identifies objects staining with DAPI and PE in the same location and generates images for the DAPI, FITC, PE, and APC filters. A reviewer selects the CTC defined as nucleated DAPI+ cells larger than 4 μm , lacking CD45-APC and expressing CK-PE from the gallery of objects, which are tabulated by the computer. Figure 3.1 shows an overview of the image acquisition and identification of CTC. After a scan, the cartridges were stored at room temperature until the reviewer was finished reviewing the images. Accuracy, precision, linearity, and reproducibility of the CellSearch system have been described elsewhere [12].

3.2.4 SAMPLES USED FOR FISH COUNTING ALGORITHM DEVELOPMENT

For algorithm development, images from cells of a patient in which no CTC were detected were used. Leukocytes which were non-specifically carried over during the CTC enrichment procedure were used as targets for this purpose, as in these cells almost no chromosomal aberrations are present and each cell should therefore contain two copies of each chromosome and gene region. The sample was labeled with a centromere specific FISH probe for chromosome 17 and a probe identifying the Her-2 gene region. The centromere probe is larger than the Her-2 probe, a difference in size for which the automated algorithm should be trained.

3.2.5 SAMPLE PREPARATION FOR FISH PROBES ON CTC

For the validation of the algorithm, CTC from 47 patients with hormone refractory metastatic prostate cancer labeled with probes identifying the centromeres of chromosome 1, 7, 8 and 17 were used [10]. To preserve the location of the CTC for future interrogation the buffer inside the cartridge was carefully aspirated to avoid cell movement and replaced with methanol acetic acid. After fixation of the cells, the cartridges were dried using a forced air flow and processed for FISH or stored at -20°C for later use. FISH probes specific for the centromeric regions of chromosome 1, 7, 8, and 17 labeled with PlatinumBright-647, -550, -505, and -415, respectively, were used in this study (Kreatech, Amsterdam, The Netherlands). The probe mixture consisted of 50 μl of hybridization buffer (50% Formamide/1 x SSC/10% Dextran Sulfate) containing FISH probes against 1, 7, 8, and 17 at 2 ng/ μl each. The cartridges were placed on a 80°C hotplate for 2 min, with the glass facing towards the hotplate, and hybridized at 42°C for 16 h. After hybridization the cartridges were washed with PBS containing DAPI as a nuclear counter stain.

3.2.6 DATA ACQUISITION FOR FISH PROBE DETECTION IN CTC

After hybridizing the FISH probes, the samples were scanned on a modified CellTracks Analyzer II. This setup is equipped with a $40\times/0.63\text{NA}$ objective, to improve the resolution and light collection of the fluorescent FISH dots, and filter cubes to detect DAPI, PlatinumBright-647, -550, -505, and 415. The locations and images of the CTC identified in the initial $10\times$ scan were loaded from a CD. A software program was written to move to the locations of interest and record Z-stacks to capture signals at a range of depths of the objects of interest [10]. The DAPI signals are used to correlate the $40\times$ with the $10\times$ scan, thereby verifying if the CTC location is correct. This was necessary, as the cells could shift up to $\sim 200\ \mu\text{m}$ due to the FISH protocol. The image acquisition procedure for the FISH probe detection is shown in figure 3.2.

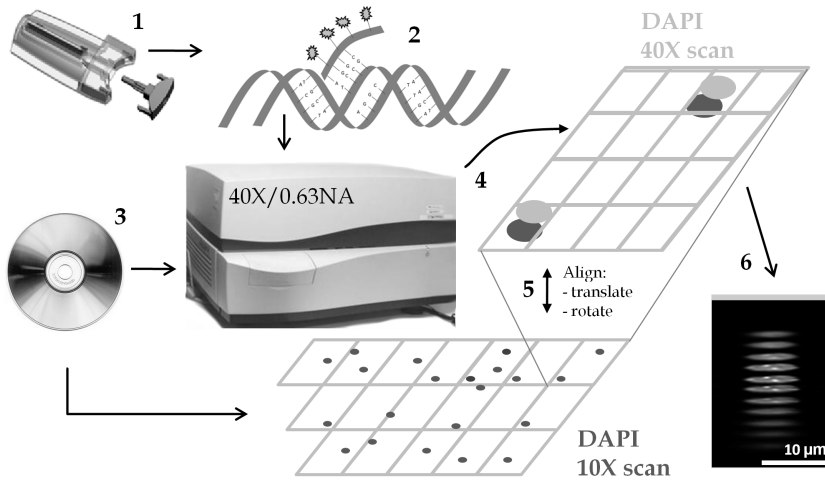


FIGURE 3.2 : Procedure for FISH probe imaging. 1) The cartridge is re-opened to carefully aspirate buffer and fixate the cells. 2) FISH reagents are added and the sample is hybridized for 16 h. 3) Coordinates and images from previously assigned CTC are loaded from a CD. 4) The modified CellTracks records DAPI images at $40\times$ at the designated coordinates and surroundings. 5) Cross correlation is performed between the DAPI images from the $40\times$ and $10\times$ scan to verify the correct location of the CTC. 6) After the right location is found, the FISH z-stacks are recorded.

3.2.7 ALGORITHM FOR COUNTING FISH SIGNALS

Maximum intensity profiles were created of all the Z-stacks to speed up the counting process for human reviewing. The algorithm to identify the nucleus and count the FISH probes within this nucleus consists of five general steps:

1. enhancement and segmentation of the outline of the nucleus;
2. enhancement and segmentation of potential FISH objects;
3. exclusion of objects that are too noisy;
4. measurement of intensity and morphological features of the potential objects;
5. exclusion of objects that do not meet inclusion criteria.

Nuclei were located in the DAPI image, which was enhanced using a zero-crossing filter in combination with a gradient magnitude filter [21]. In short, edges were enhanced using a morphological gradient magnitude filter. The image of this filter was multiplied by an image filtered by a laplace-plus-dgg filter. This filter created a combination of second order derivatives in x and y , and combined it with an image of the second derivative in

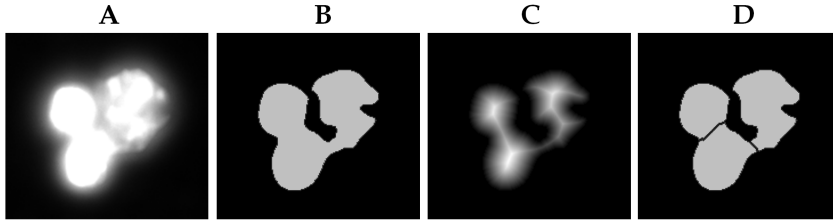


FIGURE 3.3 : Example of a distance transform aiding the watershed transform. Panel A shows the original DAPI image containing three nuclei which is thresholded to give the outline at panel B. This outline is difficult to watershed due to saturation in the original image. The distance transform shown in panel C creates a new input for the watershed transform, which successfully separates the nuclei as is shown in panel D.

the direction of the maximum gradient. The combined image of the two filters is thresholded using a fixed threshold. On the outline of objects larger than 2000 pixels a distance transform was applied: every outline pixel value is replaced by its closest distance to the edge of the outline. This procedure was then followed by a watershed transform, to verify if the outline consisted of multiple maxima and were thus two or more closely spaced objects. Figure 3.3 shows how a distance transform improves the watershed transform in case of saturated DAPI signals. The nucleus that was at least 250 pixels in size and located closest to the middle of the image was selected as the final outline: only objects inside this outline are considered for FISH dot counting.

In a next step, dot-like structures were enhanced. Usually this is performed by employing a tophat filter [14]. We used a method termed multiscale product [22] because it appeared better suited for dealing with the heterogeneity of the sizes of the FISH probes. This filter increases the intensity of objects in a range of radii, using a multiplication of Gaussian kernels with different σ (ranging from 1.5 to 3.0 pixels):

$$G_{product} = (G - G_0) \cdot (G_0 - G_1) \cdot (G_1 - G_2) \quad (3.1)$$

with the 2D Gaussian kernel defined as

$$G(\sigma_i) = \frac{1}{2\pi\sigma_i^2} e^{-\frac{x^2+y^2}{2\sigma_i^2}}. \quad (3.2)$$

Using a range of Gaussian kernel sizes improves robustness to variations in size of the objects of interest, compared to using a single kernel size as is done in the tophat transform method. After applying the multiscale product, objects were thresholded using the triangle threshold method [23] of which the value was multiplied by a factor of 0.1 to include all relevant dot-like structures, bright or dim. This thresholding method uses the intensity histogram of the image and is especially suited for images with

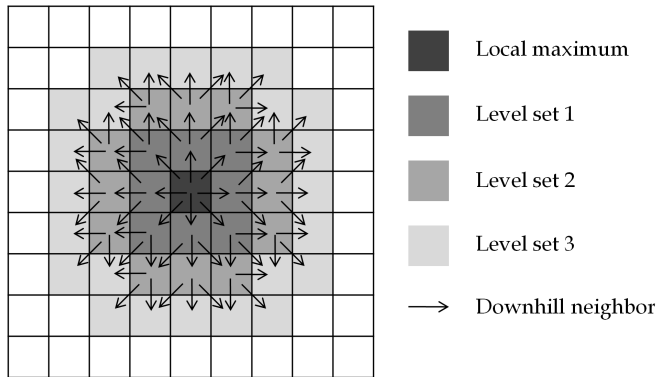


FIGURE 3.4 : Scheme for verifying downhill neighbors. By verifying objects in this way, only dome shape objects are included. (adapted from [18])

few object pixels. After thresholding, a final verification was performed to exclude objects that were too noisy. The dome finding method of Restif was applied to the coordinates of the maximum of each object [18]. This method checks nearest downhill neighbors in three level sets -up to three pixels around the maximum- and excludes objects that have more than one extra local maximum in this region. Figure 3.4 shows the method by which nearest downhill neighbors were verified.

Restif et al. determined that if one extra local maximum is allowed, at least 75% of the downhill neighbors should have a lower intensity. We allow one extra local maximum to include very closely spaced FISH dots. In this way, noisy objects are excluded, whether or not this noise is originating from a high or low intensity background. Finally, measurements were performed on these objects: size, maximum intensity, mean intensity, total intensity, relative intensity, roundness, and perimeter were saved for every object. Relative intensity was defined as the total intensity of the object related to the total intensity of the brightest object within the same nucleus. Using these measurements, different exclusion criteria were tested. Combinations of measurements were tested on the leukocyte training set to exclude debris and keep the true FISH dots. Figure 3.5 shows a schematic overview of the procedure.

3.2.8 EXPERT REVIEWING OF SAMPLES

Next to the algorithm, five expert reviewers counted FISH probe signals in the leukocyte images, using a macro written in the program ImageJ [24]. They reviewed the set two times: the first time they were asked to review all images, the second time they could skip images that were unclear in their view. In this way it could be measured how sure reviewers were. All

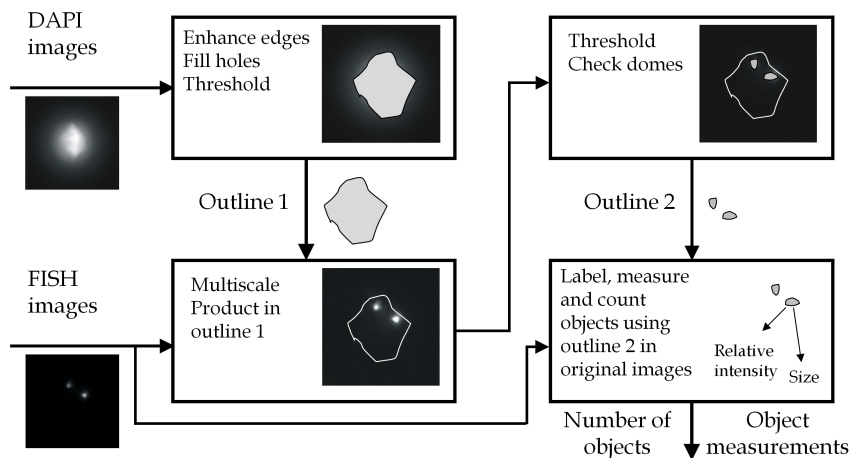
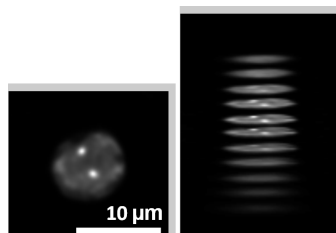


FIGURE 3.5 : Schematic representation of the counting process. DAPI images are loaded and their outline is determined. Within this outline, a multiscale product is applied. Objects within outline 1 are thresholded and checked for dome-like structures. Finally, the objects from outline 2 are counted and measured.

FIGURE 3.6 : Example of a recorded Z-stack of FISH probes in a leukocyte in top-view, panel A, and in slice-view, panel B. Distance between the slices is $1 \mu\text{m}$.



CTC samples were also reviewed by five expert reviewers.

3.3 RESULTS

492 leukocytes and 500 CTC were imaged by the modified CellTracks Analyzer II. Figure 3.6 shows an example of a FISH Z-stack from a leukocyte in top and slice view, from which a maximum profile was created. The profiles were processed by the algorithm, requiring 2 minutes for each sample and counted by human reviewers, requiring 2 hours for each sample. Figure 3.7 shows the different steps of the algorithm: segmentation of the nucleus, enhancement of dot-like structures and the final outline of nucleus and dots.

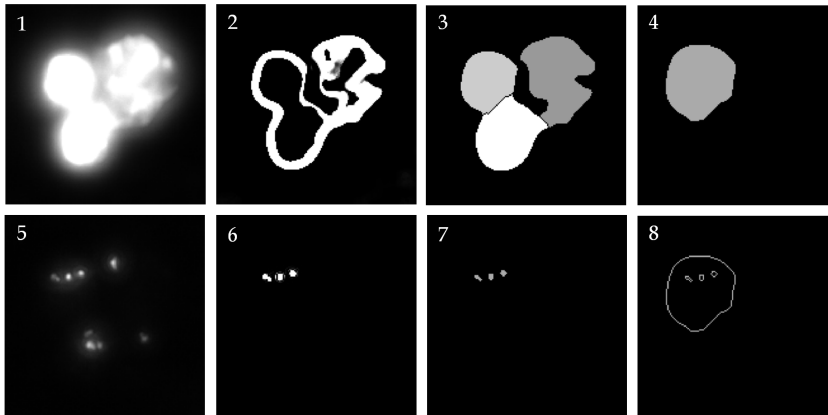


FIGURE 3.7 : Example of the image processing steps on a recorded DAPI and FISH image. 1) original DAPI image; 2) edge enhanced DAPI image; 3) outline after thresholding and watershed; 4) final DAPI outline; 5) original FISH image; 6) FISH image after multiscale product; 7) outline FISH image; 8) final outlines of DAPI and FISH.

3.3.1 COUNTING OF THE LEUKOCYTE TRAINING SAMPLES

After comparison of the manual and automated counts in the training sample, it became apparent that only the measurements "size" and "relative intensity" had a positive impact on the counting agreement of the algorithm. After objects were measured and counted in the Her-2 channel, the objects with a relative intensity lower than 30% of the brightest dot within that nucleus were excluded. For the centromere 17 channel this threshold was optimal at 25%. Objects smaller than 5 pixels were also excluded. Automatic counting of chromosomes in leukocytes resulted in an accuracy of 97.8% of the Her-2 dots and 97.5% of the centromere 17 dots. Accurate here means "equal to the manual count of the subset of images were all reviewers agreed upon" (N=409 for Her-2 and N=347 for centromere 17). The mean inter-reviewer agreement over the whole data set was $92.6\% \pm 2.3\%$ and $91.7\% \pm 1.7\%$ and the mean intra-reviewer agreement was $96.5\% \pm 2.7\%$ and $97.0\% \pm 1.8\%$ for the Her-2 and centromere 17 probes, respectively. Table 3.1 gives an overview of the counting agreement after review of Her-2 and centromere 17 of the whole data set and the data set containing only the images with objects that could be easily identified by the reviewer, compared with the count generated by the algorithm. In figure 3.8 the distribution of the count of the PC and five reviewers is shown. The count of the reviewers is represented by the mean and the standard deviation for each chromosome count.

TABLE 3.1 : Agreement between expert reviewers and the PC algorithm when reviewing leukocytes. In the upper right part of the table (not emphasized), the agreement of the first review is shown in which the reviewer had to review the full dataset (N=492). Second, in the lower left part of the table (emphasized bold) the agreement of the second review is shown, in which the reviewers only reviewed the cells they were certain of (the "obvious" dataset). Last, the intra-reviewer variation is given on the diagonal of the table (emphasized italic).

Her-2	Reviewer 1	Reviewer 2	Reviewer 3	Reviewer 4	Reviewer 5	PC
Reviewer 1	<i>96.2%</i> ±1.7%	91.3%±2.5%	93.9%±2.1%	93.5%±2.2%	95.5%±1.9%	90.3%±2.6%
Reviewer 2	91.3% ±2.5%	<i>92.5%</i> ±2.3%	89.9%±2.7%	89.3%±2.7%	90.9%±2.6%	85.0%±3.2%
Reviewer 3	93.9% ±2.2%	91.0% ±2.6%	<i>95.7%</i> ±1.8%	92.7%±2.3%	96.0%±1.8%	90.5%±2.6%
Reviewer 4	98.7% ±1.2%	97.7% ±1.6%	97.9% ±1.5%	<i>99.2%</i> ±1.0%	92.9%±2.3%	87.6%±2.9%
Reviewer 5	97.3% ±1.5%	93.0% ±2.3%	96.2% ±1.8%	99.2% ±1.0%	<i>98.9%</i> ±1.0%	92.1%±2.4%
PC	90.5% ±2.6%	87.4% ±2.9%	89.8% ±2.7%	96.9% ±1.8%	92.6% ±2.4%	<i>100.0%</i> -0.4%
Centromere 17						
Reviewer 1	<i>96.4%</i> ±1.7%	92.9%±2.3%	92.7%±2.3%	93.7%±2.2%	90.5%±2.6%	87.6%±2.9%
Reviewer 2	92.3% ±2.4%	<i>94.9%</i> ±2.0%	93.9%±2.1%	90.7%±2.6%	92.3%±2.4%	88.1%±2.9%
Reviewer 3	95.0% ±2.0%	94.5% ±2.1%	<i>96.0%</i> ±1.8%	90.3%±2.6%	91.7%±2.5%	85.4%±3.1%
Reviewer 4	96.1% ±1.9%	98.3% ±1.3%	97.0% ±1.7%	<i>99.0%</i> ±1.0%	88.5%±2.8%	89.1%±2.8%
Reviewer 5	96.0% ±1.9%	96.0% ±1.9%	98.1% ±1.4%	96.4% ±1.9%	<i>98.6%</i> ±1.2%	88.5%±2.8%
PC	87.4% ±2.9%	91.5% ±2.5%	88.4% ±2.9%	95.4% ±2.1%	91.7% ±2.7%	<i>100.0%</i> -0.4%

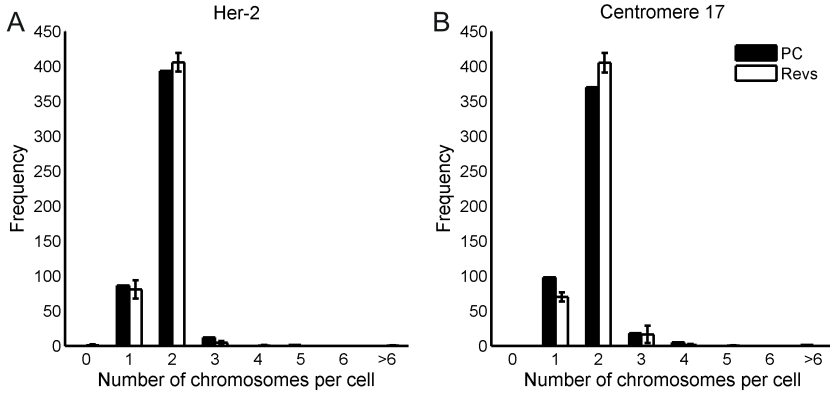


FIGURE 3.8 : Number of leukocytes for the Her-2 gene probe (panel A) and the centromere of chromosome 17 probe (panel B) as counted by the PC and the five reviewers (N=492).

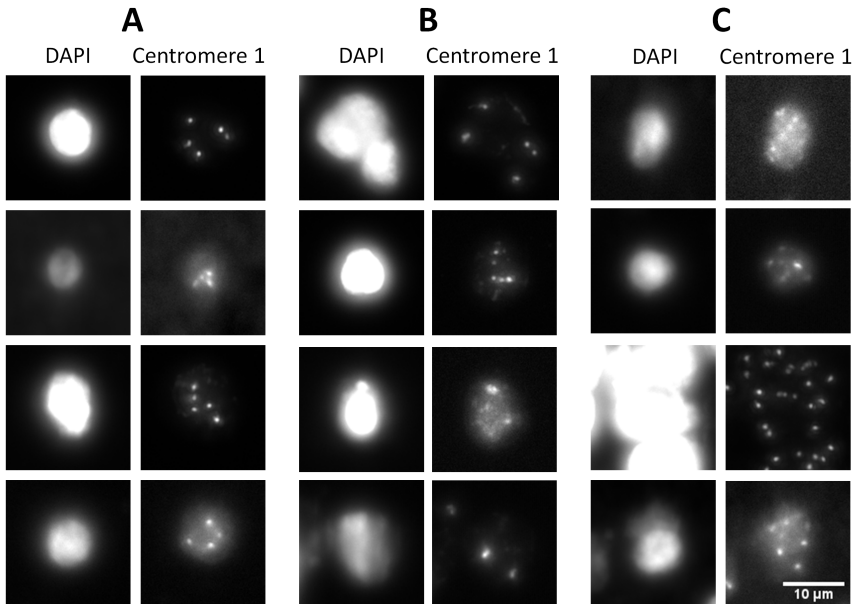


FIGURE 3.9 : Example of FISH images. Panel A shows cases where all reviewers and the PC agreed. Panel B shows examples where all the human reviewers agreed, but the PC disagreed on the number of FISH dots. Panel C shows examples with the largest discrepancy among reviewers.

3.3.2 COUNTING OF SAMPLES CONTAINING CTC

After processing a sample containing CTC it became clear that the threshold for relative intensity was strongly related to the quality of the FISH probe used, and thus probe dependent. Slight adjustment of the relative intensity criteria to a range from 14%-20% was necessary to ensure reasonable counting by the algorithm. This value was correlated with the average of the maximum intensity of all the objects in a channel: if this average was high, then the relative intensity should be set lower. Figure 3.9 shows three categories of examples from the data set: panel A shows examples where all the reviewers and the PC agreed; panel B shows examples where all reviewers agreed, but the PC did not; finally, panel C shows examples where there was a large discrepancy between all reviewers and the PC.

Agreement of the PC with the subset of cells on which all reviewers agreed was 76.1% (n=238), 83.9% (n=280), 86.6% (n=209), and 85.3% (n=251) for probes from centromere 1, 7, 8, and 17 respectively. Mean inter-reviewer agreement was 70.9%, 75.3%, 66.8%, and 72.3% for these four channels. Figure 3.10 shows the agreement between all reviewers in detail and the histogram of the count.

3.4 DISCUSSION

3.4.1 AUTOMATED COUNTING IS NECESSARY AND FEASIBLE

We have shown that reliable automated counting of FISH probes on EpCAM+DAPI+CK+CD45- cells is both necessary and feasible. Comparing expert reviews revealed that intra-reviewer variation -the same expert reviewing a data set twice- could be as high as 3.5% of the cells. Inter-reviewer variation was higher: 7.5%; these numbers were both acquired for the "easy" leukocyte samples with low copy numbers. Variation between reviewers while reviewing CTC samples could be as high as 33.2% (centromere chromosome 8), showing that the number of signals in a nucleus is of great influence on counting accuracy, as is the knowledge of the reviewer that he or she is dealing with CTC or leukocytes. Furthermore, reviewing 500 FISH nuclei in four channels takes several hours, while the computer only needs a few minutes.

When comparing these results to earlier work [14], not much improvement is seen in the percentages of correctly assigned dots. However, FISH dot counters are usually compared using cells from cell lines. These cells have a high homogeneity and are thus a relative easy target. Our CTC samples contain a very heterogeneous population of cells, of which some may be apoptotic (for more examples see [12]). Furthermore, the procedure for labeling FISH dots in our system, although optimized for dot quality, still induces large variations in staining intensity. The leukocytes that were used as a training set were labeled using the same procedure. It is a two-step

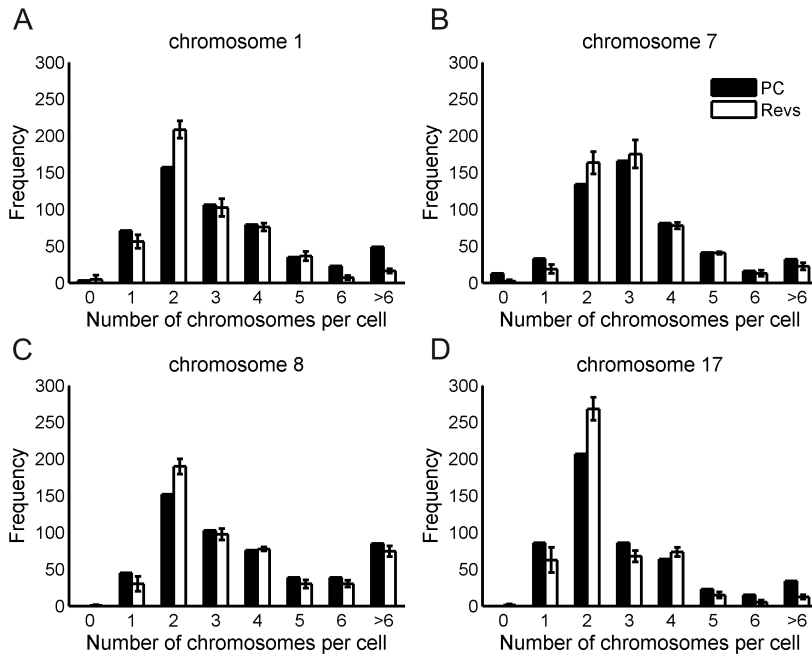


FIGURE 3.10 : Number of CTC containing 0-6 or >6 dots for the centromeres of chromosomes 1, 7, 8, and 17 probe (panel A-D) as counted by the PC and the five reviewers (N=500). The reviewer count is given as the mean of the frequency of each number and its standard deviation.

rocket: first localize the CTC by an immuno assay, then follow-up using the FISH procedure. Inter- and intra-reviewer variations in these samples are high. Although no progress was made in the percentage of correctly identified dots, acquiring these percentage with CTC samples is a greater challenge.

From the results it becomes clear that review of chromosome 1 and 8 was the most difficult, for both PC and reviewer. These probes had on average a factor two lower intensities than the probes from chromosomes 7 and 17. Thus, the inter-reviewer agreement was lower as well as the agreement with the PC. The dome finding part of the algorithm revealed the same: it removed objects that were too noisy in 17% and 13% of the nuclei in the channel from chromosome 1 and 8 respectively, and only in 8% of the nuclei from channels of chromosome 7 and 17. Signal to noise ratios were clearly lower in channels where the agreement was lower.

Agreement between PC and reviewer was good when control samples were reviewed, and reasonable when CTC were reviewed. The difference between the two data sets could be attributed to a few sources of error:

1) Nuclei were well separated in the leukocyte sample, the CTC samples contained more clusters. While these clusters are usually easily resolved by eye, the algorithm had more difficulty in this task. Figure 3.9, panel B shows an example in row 1: two closely spaced nuclei with almost saturated intensity. In this case the signals of the nuclei are close to saturation and although a distance transform and watershed transform was applied, they were still segmented as one. The PC thus over-counted in this example.

2) Because the DAPI signal from the nuclei can vary greatly between samples, some signals fall just outside the segmented outline of the nucleus as determined by the algorithm. This is the case when the signal from the nucleus is relatively dim, as is shown in figure 3.9, panel B, row 4, where the reviewers counted two probes and the PC counted only one. This challenge could be resolved by dilating the outline nuclei more than is done now. However, closely spaced nuclei will be resolved worse in this case. The heterogeneity of the shape and size of the nucleus is largely due to presence of ferrofluid in combination with the fixation step in the FISH procedure. The ferrofluid particles were added to keep the cells tightly at the imaging surface. However, due to the influence of these magnetic particles and the tendency of some cells to adhere to surfaces, the DNA spreads over the surface. Ferrofluid particles that line up under influence of the magnetic field force these cells to spread even further. Thus in the DAPI images even small islands of DNA were visible, that clearly were part of a bigger nucleus, making it more difficult for the algorithm to measure a perfect outline of the nucleus and include all the DNA in the dot counting. Figure 3.11 shows an example of this effect.

3) The CTC sample had a larger variety in signal quality. Although the segmentation algorithm is dynamic on the histogram, it is still difficult for the PC to distinguish between what a reviewer calls a "true signal" and debris. For example, when a reviewer sees two signals -a bright and a relatively dim one-, he or she will usually count two. However, when five bright signals and one dim signal are seen, the dim object is more often discarded as being debris. Figure 3.9 panel B rows 2 and 3 show examples of difference in counting because of relative intensity. In row 2, the reviewers counted five and the pc four, while in row three, the reviewers counted two and the PC three probes. The PC counts 100% reproducible, but does not take into account these human considerations. For this analysis, it is thus very difficult to get an absolute "golden truth".

4) It still is difficult for the algorithm to distinguish between a split probe (one chromosome that had two signals) and two closely spaced chromosomes. It is however not known how often a reviewer misclassifies such an object.

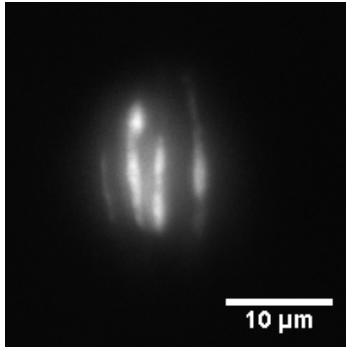


FIGURE 3.11 : Example of a nucleus spread due to fixation of the cell. Note the vertical lines that were created by ferrofluid aggregates that follow the local magnetic field. This type of nucleus is especially difficult to segment correctly.

A reviewer can structurally ignore or assign the split spots. The PC cannot and counts these items according to the algorithm. The number of cells that have these split spots may vary between samples and also within samples (e.g. between lymphocyte and CTC). The PC might not be able to distinguish between these, but if this factor appears to be of influence to the result, the PC could use the measurements of the probes -i.e. relative intensity coupled to size of closely spaced probes- to estimate the probability of these splits in the cells. Leukocytes could be used as an internal control for measuring the frequency of these splits and for estimating a relevant "size/relative intensity" threshold.

The above error sources may seem a big challenge, but are not of importance for the clinical relevant observations which is the presence or absence of aneuploidy ascertain cancerous origin and the presence of amplification or deletions of specific genes. CTC are very heterogeneous: within one patient a wide variety of chromosomal aberrations could be spotted. So whether or not a certain cell has five or six copies is of lesser importance than the fact that this number is greater than two. When comparing counts that are greater than two or not, the reviewer and PC concur in 87%, 93%, 94%, and 94% of the cells for centromere 1, 7, 8, and 17 respectively for the data set in which all reviewers agree. This demonstrates that in about 90% of the cases, the PC and reviewer will draw the same conclusion about the ploidy status of the cells identified as tumor cells.

Figure 3.9 panel C shows examples in which the reviewers greatly disagreed. Two examples of varying signal intensities (rows 1 and 2) and two examples of difficulty of locating the true outline of the nucleus (row 3 en 4) are given. It could be argued that the example of row 1 is not suitable for reviewing at all because the background staining is too high. For reviewers, there is no real quantitative criterion whether or not to reject a certain object based on its intensity distribution. However, the PC has such a criterion: it can easily verify if a maximum of an object is surrounded by more than two other local maxima. If this is the case, then an object should be excluded. We perform this verification by means of the dome

finding function. In this way, the PC performs more reliable than the human reviewers.

3.4.3 FUTURE RESEARCH

In the future, the algorithm may be optimized further by using clinical data. When coupling for instance response to a therapy of a patient to the aberration of the genes in the CTC, a better golden truth may be found. Furthermore, quality of FISH could still be improved. Split probes are still a big challenge for the PC, but also for establishing a good count by reviewers. Consequently a quality score could be set by the algorithm by measuring intensity variations, for instance in carried-over leukocytes. This score could be used as an internal control in each patient sample to adjust exclusion criteria and to reject cells that are not suitable for interpretation. Finally, removal of ferrofluid could greatly improve the segmentation of the nucleus. Aggregation of ferrofluid particles disturbs the natural shape of the nucleus and blocks a fraction of the fluorescence light. Implementation of physical filters to enrich CTC by size would not require any ferrofluid and could be an improvement in the next generation CTC capturing devices.

ACKNOWLEDGEMENTS

The authors would like to acknowledge Ronald Sipkema for his contribution to the software for the improved CellTracks Analyzer and Bernd Talsma for his work on segmentation of the nucleus.

3.5 REFERENCES

- [1] S. J. Cohen, C. J. A. Punt, N. Iannotti, B. H. Saidman, K. D. Sabbath, N. Y. Gabrail, J. Picus, M. Morse, E. Mitchell, M. C. Miller, G. V. Doyle, H. Tissing, L. Terstappen, and N. J. Meropol, "Relationship of circulating tumor cells to tumor response, progression-free survival, and overall survival in patients with metastatic colorectal cancer," *Journal of Clinical Oncology*, vol. 26, no. 19, pp. 3213–3221, 2008.
- [2] M. Cristofanilli, G. T. Budd, M. J. Ellis, A. Stopeck, J. Matera, M. C. Miller, J. M. Reuben, G. V. Doyle, W. J. Allard, L. Terstappen, and D. F. Hayes, "Circulating tumor cells, disease progression, and survival in metastatic breast cancer," *New England Journal of Medicine*, vol. 351, no. 8, pp. 781–791, 2004.
- [3] J. S. de Bono, H. I. Scher, R. B. Montgomery, C. Parker, M. C. Miller, H. Tissing, G. V. Doyle, L. Terstappen, K. J. Pienta, and D. Raghavan, "Circulating tumor cells predict survival benefit from treatment in metastatic castration-resistant prostate cancer," *Clinical Cancer Research*, vol. 14, no. 19, pp. 6302–6309, 2008.
- [4] G. Attard, J. F. Swermeuhuis, D. Olmos, A. H. M. Reid, E. Vickers, R. A'Hern, R. Levink, F. Coumans, J. Moreira, R. Riisnaes, N. B. Oommen, G. Hawche, C. Jameson, E. Thompson, R. Sipkema, C. P. Carden, C. Parker, D. Dearnaley, S. B. Kaye, C. S. Cooper, A. Molina, M. E. Cox, L. Terstappen, and J. S. de Bono, "Characterization of *erg*, *ar* and *pten* gene status in circulating tumor cells from

- patients with castration-resistant prostate cancer,” *Cancer Research*, vol. 69, no. 7, pp. 2912–2918, 2009.
- [5] J. S. de Bono, G. Attard, A. Adjei, M. N. Pollak, P. C. Fong, P. Haluska, L. Roberts, C. Melvin, M. Repollet, D. Chianese, M. Connely, L. Terstappen, and A. Gualberto, “Potential applications for circulating tumor cells expressing the insulin-like growth factor-i receptor,” *Clinical Cancer Research*, vol. 13, no. 12, pp. 3611–3616, 2007.
- [6] D. F. Hayes, T. M. Walker, B. Singh, E. S. Vitetta, J. W. Uhr, S. Gross, C. Rao, G. V. Doyle, and L. Terstappen, “Monitoring expression of her-2 on circulating epithelial cells in patients with advanced breast cancer,” *International Journal of Oncology*, vol. 21, no. 5, pp. 1111–1117, 2002.
- [7] S. Meng, D. Tripathy, S. Shete, R. Ashfaq, H. Saboorian, B. Haley, E. Frenkel, D. Euhus, M. Leitch, C. Osborne, E. Clifford, S. Perkins, P. Beitsch, A. Khan, L. Morrison, D. Herlyn, L. Terstappen, N. Lane, J. Wang, and J. Uhr, “upar and her-2 gene status in individual breast cancer cells from blood and tissues,” *Proceedings Of The National Academy Of Sciences Of The United States Of America*, vol. 103, no. 46, pp. 17361–17365, 2006.
- [8] S. D. Meng, D. Tripathy, S. Shete, R. Ashfaq, B. Haley, S. Perkins, P. Beitsch, A. Khan, D. Euhus, C. Osborne, E. Frenkel, S. Hoover, M. Leitch, E. Clifford, E. Vitetta, L. Morrison, D. Herlyn, L. Terstappen, T. Fleming, T. Fehm, T. Tucker, N. Lane, J. Q. Wang, and J. Uhr, “Her-2 gene amplification can be acquired as breast cancer progresses,” *Proceedings Of The National Academy Of Sciences Of The United States Of America*, vol. 101, no. 25, pp. 9393–9398, 2004.
- [9] D. A. Smirnov, D. R. Zweitzig, B. W. Foulk, M. C. Miller, G. V. Doyle, K. J. Pienta, N. J. Meropol, L. M. Weiner, S. J. Cohen, J. G. Moreno, M. C. Connelly, L. Terstappen, and S. M. O’Hara, “Global gene expression profiling of circulating tumor cells,” *Cancer Research*, vol. 65, no. 12, pp. 4993–4997, 2005.
- [10] J. F. Swennenhuis, A. G. Tibbe, R. Levink, R. C. Sipkema, and L. W. Terstappen, “Characterization of circulating tumor cells by fluorescence in situ hybridization,” *Cytometry Part A*, vol. 75, no. 6, pp. 520–7, 2009.
- [11] T. Fehm, A. Sagalowsky, E. Clifford, P. Beitsch, H. Saboorian, D. Euhus, S. D. Meng, L. Morrison, T. Tucker, N. Lane, B. M. Ghadimi, K. Heselmeyer-Haddad, T. Ried, C. Rao, and J. Uhr, “Cytogenetic evidence that circulating epithelial cells in patients with carcinoma are malignant,” *Clinical Cancer Research*, vol. 8, no. 7, pp. 2073–2084, 2002.
- [12] W. J. Allard, J. Matera, M. C. Miller, M. Repollet, M. C. Connelly, C. Rao, A. G. J. Tibbe, J. W. Uhr, and L. Terstappen, “Tumor cells circulate in the peripheral blood of all major carcinomas but not in healthy subjects or patients with nonmalignant diseases,” *Clinical Cancer Research*, vol. 10, no. 20, pp. 6897–6904, 2004.
- [13] M. Kagan, D. Howard, T. Bendele, J. Mayes, J. Silvia, M. Repollet, J. Doyle, J. Allard, N. Tu, T. Bui, T. Russell, C. Rao, M. Hermann, H. Rutner, and L. Terstappen, “A sample preparation and analysis system for identification of circulating tumor cells,” *Journal of Clinical Ligand Assay*, vol. 25, no. 1, pp. 104–110, 2002.
- [14] H. Netten, I. T. Young, L. J. vanVliet, H. J. Tanke, H. Vrolijk, and W. C. R. Sloos, “Fish and chips: Automation of fluorescent dot counting in interphase cell nuclei,” *Cytometry*, vol. 28, no. 1, pp. 1–10, 1997.
- [15] C. O. de Solorzano, A. Santos, I. Vallcorba, J. M. Garcia-Sagredo, and F. del Pozo, “Automated fish spot counting in interphase nuclei: Statistical validation and data correction,” *Cytometry*, vol. 31, no. 2, pp. 93–99, 1998.

- [16] B. Lerner, W. F. Clocksin, S. Dhanjal, M. A. Hulten, and C. M. Bishop, "Automatic signal classification in fluorescence in situ hybridization images," *Cytometry*, vol. 43, no. 2, pp. 87–93, 2001.
- [17] F. Raimondo, M. A. Gavrielides, G. Karayannopoulou, K. Lyroudia, I. Pitas, and I. Kostopoulos, "Automated evaluation of her-2/neu status in breast tissue from fluorescent in situ hybridization images," *Ieee Transactions On Image Processing*, vol. 14, no. 9, pp. 1288–1299, 2005.
- [18] C. Restif, *Segmentation and Evaluation of Fluorescence Microscopy Images*. PhD thesis, 2006.
- [19] C. G. Rao, D. Chianese, G. V. Doyle, M. C. Miller, T. Russell, R. A. Sanders, and L. Terstappen, "Expression of epithelial cell adhesion molecule in carcinoma cells present in blood and primary and metastatic tumors," *International Journal of Oncology*, vol. 27, no. 1, pp. 49–57, 2005.
- [20] A. G. J. Tibbe, B. G. de Grooth, J. Greve, G. J. Dolan, C. Rao, and L. Terstappen, "Magnetic field design for selecting and aligning immunomagnetic labeled cells," *Cytometry*, vol. 47, no. 3, pp. 163–172, 2002.
- [21] P. W. Verbeek and L. J. Vanvliet, "On the location error of curved edges in low-pass filtered 2-d and 3-d images," *Ieee Transactions on Pattern Analysis and Machine Intelligence*, vol. 16, no. 7, pp. 726–733, 1994.
- [22] B. J. Vermolen, Y. Garini, I. T. Young, R. W. Dirks, and V. Raz, "Segmentation and analysis of the three-dimensional redistribution of nuclear components in human mesenchymal stem cells," *Cytometry Part A*, vol. 73A, no. 9, pp. 816–824, 2008.
- [23] G. W. Zack, W. E. Rogers, and S. A. Latt, "Automatic-measurement of sister chromatid exchange frequency," *Journal Of Histochemistry & Cytochemistry*, vol. 25, no. 7, pp. 741–753, 1977.
- [24] T. J. Collins, "Imagej for microscopy," *Biotechniques*, vol. 43, no. 1, pp. 25–+, 2007.

IMAGE ANALYSIS ALGORITHM FOR THE RECOGNITION OF CIRCULATING TUMOR CELLS¹

Sjoerd T. Ligthart

ABSTRACT

Circulating tumor cells (CTC) in patients with metastatic carcinomas are associated with poor survival and can be used to guide therapy. However, classification of CTC remains subjective, as they are morphologically heterogeneous. We acquired digital images, using the CellSearch[®] system, of EpCAM positive circulating objects captured from the blood of 185 castration resistant prostate cancer (CRPC) patients and 68 healthy subjects to define CTC by computerized algorithms. Patient survival data was used as the training parameter for the computer defined CTC. The optimal CTC definition was selected based on: 1. High Cox Hazard ratio (HR) for both baseline and follow-up samples; 2. Higher HR for follow-up than baseline samples since these patients received drugs with antitumor activity; 3. Low relative and absolute count in control samples. Our computer-generated CTC definition resulted in HRs of 3.1 for baseline and 4.8 for follow-up samples after dichotomization on the median number of objects found at baseline, which is comparable to the manual CellSearch CTC definition (Baseline HR 2.9, follow-up HR 4.5). Processing of a sample using the automated classifier

¹This chapter contains the technical details of the image analysis approach for automated CTC enumeration as described in: S.T. Ligthart, F.A.W. Coumans, G. Attard, A. Mulick Cassidy, J. S. De Bono & L.W.M.M. Terstappen "Unbiased and Automated Identification of a Circulating Tumour Cell Definition that Associates with Overall Survival" *Plos One* 2011, DOI:10.1371/journal.pone.0027419

required no operator intervention, compared to up to 39 minutes of review per sample by highly trained operators in the CellSearch method. Summarizing, we have created a definition for CTC using automated counting that compares favorably to manual counting by the CellSearch definition, is rapid and reproducible.

4.1 INTRODUCTION

In recent years several studies have reported that counting circulating tumor cells (CTC) can indicate the effectiveness of a therapy for advanced cancer already after the first cycle of therapy [1, 2, 3, 4, 5, 6, 7, 8]. At present, the CellSearch method is the only FDA validated method for CTC enumeration [9]. In this system, CTC are enriched from 7.5 ml of blood and fluorescently labeled with cytokeratin-phycoerythrin (CK-PE), CD45-allophycocyan (CD45-APC) and the nuclear dye 4,6-diamidino-2-phenylindole (DAPI). The recorded fluorescence images of CK, DNA, CD45 and a FITC channel are segmented on the basis of being positive for CK and DNA and are then presented to a trained reviewer.

For CTC classification by a reviewer, objects need to be CK+, DAPI+, CD45-, $>4\mu\text{m}$ in size and have a cell-like morphology. This procedure is very laborious, time-consuming and can be highly subjective. Moreover, CTC are known to be morphologically heterogeneous [1] and different laboratories have slightly different definitions of what constitutes a CTC, especially in the case of objects that are dead or apoptotic [10]. Because CTC occur at very low frequencies, misjudging a few events may be very significant [11]. Furthermore, the definition of a CellSearch CTC that has been used to date may not be optimal. Our group recently reported that tumor micro particles (TMPs) -objects that are EpCAM+CK+CD45- and smaller than $4\mu\text{m}$ - have the same prognostic value as manually counted "CTC" [12], suggesting alternative definitions for CTC evaluation should be considered.

Here we present the results of a new approach to identify CTC in images recorded by ten CellSearch systems from samples of castration-resistant prostate cancer (CRPC) patients [13]. We recorded images before treatment (baseline samples) and from the first follow-up sample. Our hypothesis was that using solely survival data as a training feature, an automated algorithm could be optimized to count CTC with the same fidelity as the manual CellSearch method. We stipulated that such an algorithm needed to identify candidate CTC, extract several relevant properties and compare the candidate to a range of known parameters. The algorithm should in principal at first select all objects that are present: CTC are so heterogeneous that a minimal number of assumptions should be made about shape, intensity distribution, and size, in order to avoid excluding any objects that could be of importance. CTC are very rare objects and the maximum number of objects with clinical impact should therefore be included.

Automation of cell identification and characterization on cells other than CTC has been researched extensively. Several free-ware image processing packages are available that can be used for batch processing of cytological images such as ImageJ ([14], <http://rsbweb.nih.gov/ij/>), CellProfiler ([15], <http://www.cellprofiler.org/>), and for time laps images CellCognition ([16], <http://www.cellcognition.org/>). Commercial software packages are also available, examples are Imaris (Bitplane AG, Zurich Switzerland), Autoquant (Media Cybernetics, Inc., Bethesda, MD), or Cyteseer (Vala Sciences Inc, San Diego, CA). In addition numerical computation programs such as Matlab (Mathworks, Natick, MA) and acquisition software such as Labview (National Instruments, Austin TX) are available that contain image processing toolboxes, which can be programmed easily for cytological images. Next to options for image pre-processing for manual review, these packages often offer possibilities of classification of cells using numerous parameters.

Requirements for our data set were: (i) short processing times as thousands of images are to be processed, (ii) limited number of parameters as for each patient only the time of survival after image acquisition is used which easily can result in severe over-training, (iii) versatility in options and ease of programming. The program that best suited these needs was Matlab and this program was used for this study.

4.2 MATERIALS AND METHODS

4.2.1 PATIENTS

Development of image analysis algorithms for automated CTC enumeration was performed on stored images from 10 CellSearch systems (Veridex LLC, Raritan, NJ) from patients and controls participating in prospective studies, all provided written informed consent. 276 patients were enrolled in the IMMC-38 study [13], 231 were eligible and for 185 of those patients images could be imported for baseline and first follow-up. All patients had histologically confirmed prostate adenocarcinoma, castrate levels of testosterone (<50 ng/ml) and progressive disease as defined by three consecutively rising PSA values. Baseline samples were taken up to 19 days prior to commencement of new cytotoxic chemotherapy, follow-up samples were taken 2–6 weeks after start of therapy. 121 patients started their first line of chemotherapy. Of 185 IMMC-38 patients 118 (64%) died during the study, median survival was 20.7 months and median duration of follow-up for censored patients was 29.8 months. A total of 65 clinical centers in the United States and Europe participated in this study. Images of 68 healthy individuals participating in the IMMC-06 study (20, 21) were available as a control group. This prospective trial was conducted at 55 clinical centers throughout the US, the Netherlands, and the United Kingdom. Institutional review boards at each center approved both study protocols.

4.2.2 ISOLATION OF EPCAM+ OBJECTS AND FLUORESCENCE IMAGING

The CellSearch system was used to image CTC. The system consists of a CellTracks Autoprep[®] for sample preparation and a CellTracks Analyzer II[®] for sample analysis [9]. The CellTracks Autoprep immuno-magnetically enriches epithelial cells from 7.5 ml of blood using ferrofluids conjugated to EpCAM. The enriched sample is stained with CK-PE antibodies directed against cytokeratins 8, 18 and 19, a CD45-APC antibody for leukocyte detection, and the nuclear dye 4',6-diamidino-2-phenylindole (DAPI). This enriched sample is then transferred to a magnetic cartridge where all ferrofluid labeled objects are pulled towards a cover slip. The entire cartridge cover slip is imaged by the CellTracks Analyzer II, a four-color semi-automated fluorescence microscope that captures digital images for four different fluorescent dyes using a $10\times/0.45$ NA objective and a CCD camera with 6.7×6.7 μm sized pixels. Thus, although the objective has a resolution of 556 nm using 500 nm light according to Abbe's law, the sampling by the CCD of $6.7/10 = 670$ nm in the sample does not capture this resolution fully, according to the Nyquist-Shannon sampling criterion. Hence, the images in the CellTracks system are systematically undersampled. The smallest distinguishable details in the images from the setup are therefore ~ 1.3 μm , as the criterion tells us that the cut-off frequency should be sampled by two points minimally.

Next to the DAPI, PE and APC images, a fourth fluorescence channel (emission 535 ± 25 nm) is imaged as a control channel for exclusion of auto-fluorescent debris. This channel will be termed "FITC" channel. Per cartridge, 144–180 four-layer tiff images are saved per patient, which are 1280×1024 or 1380×1035 pixels in size, depending on when the instrument was built. Images are compressed to save space from 12-bit to 8-bit by normalizing each pixel value between 0–255, in which 0 is a mapping of the dimmest pixel values from the image histogram and 255 is a mapping of the brightest pixel values. Pixel values in between are mapped linearly between 0 and 255.

4.2.3 COUNTING OF MANUAL CTC BY HUMAN REVIEWERS

After imaging by the Celltracks Analyzer II, the Celltracks software selects objects that are DNA and CK positive. It performs this task by analyzing 10×10 pixel segments of the DNA and CK images. A segment is called positive if, after re-scaling the 10×10 part from 0 to 255 by incorporating a 50×50 environment around the segment, at least 10% of the pixels have an intensity above 100. In the next step, adjacent positive segments are merged into larger squares. Segments that are positive for both CK and DNA are shown in a thumbnail gallery. A trained operator reviews this gallery and selects manual CTC (mCTC) as objects positive for DNA and CK, negative for CD45 that are larger than 4×4 μm and have cell-like morphology.

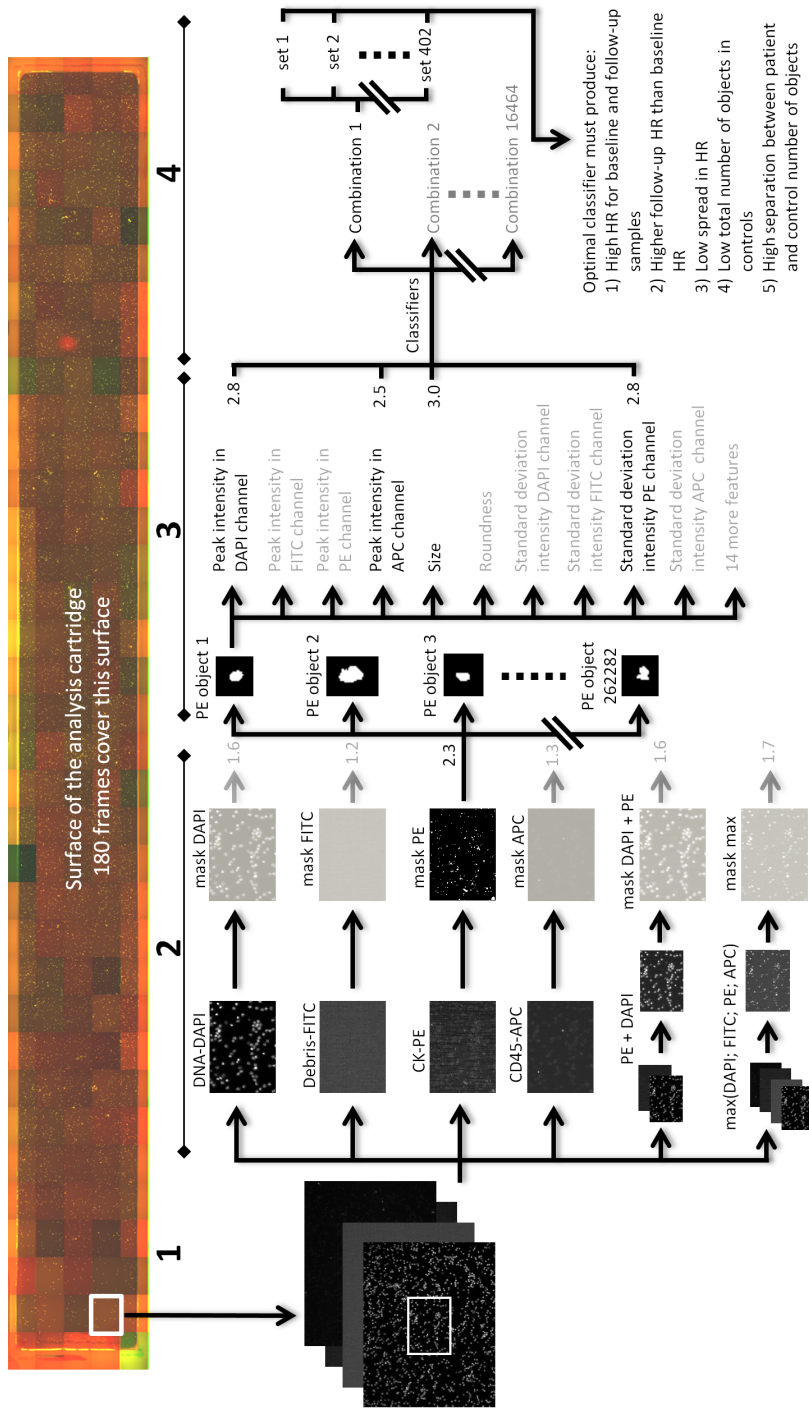


FIGURE 4.1 : Schematic overview of the aCTC classifier development process: 1) image acquisition, 2) object segmentation, 3) feature measurements, 4) creation and selection of classifiers.

4.2.4 ALGORITHM DEVELOPMENT FOR CLASSIFYING AUTOMATED CTC

Algorithm development for counting automated CTC (aCTC) was performed in Matlab 2009a (Mathworks, Natick, MA) using the DIPimage toolbox (Delft University of Technology, Delft, The Netherlands). This toolbox was chosen because it works very intuitively, while still having a lot of useful build-in image processing functions which work in any dimension. Batch processing of images and measuring of parameters is implemented relatively easy using the combination of Matlab, its parallel processing toolbox, and the DIPimage toolbox. This is usually more difficult when trying to extend dedicated software for image analysis. Figure 4.1 illustrates the method used to develop the CTC classifier, which can be divided into four major steps:

Step 1: Import images and select analysis area

CDs containing archived images of samples were collected for import to a central hard drive. All imported 8-bit multipage tiff images were scaled from 0–255 and had to be re-scaled to pseudo 12-bit using information stored in the tiff-header. This critical step had to be performed in order to quantitatively compare objects. Detection of the sample border was performed via thresholding in the debris-FITC channel and the area outside the border was excluded from further analysis. Cartridges were found to have very irregular edges, especially at the corners, and it was therefore necessary to compare the total selected area of the whole cartridge to a training set that was acquired manually. Figure 4.2 shows several of the following steps as an example of part of a cartridge. First, all FITC images from a scan were sub-sampled by a factor of eight to avoid excessive memory requirements and to neglect small details. Images expected to have a certain orientated border were convolved with a line-shaped kernel to amplify this orientation. This filtering was performed in order to close the border: differences in border intensities and gaps were common.

Next, the resulting images were connected to each other to construct an image of the surface of the total cartridge (see figure 4.2, panel 1). This image was filtered by a gradient magnitude filter, using a Gaussian derivative width of 8 pixels, to boost the edges (see figure 4.2, panel 2), and was then thresholded via the triangle threshold method [17] to identify the cartridge edges using the total image histogram (see figure 4.2, panel 3). The triangle method was used because the border only takes up a small part of the image, and thus does not show a large peak in the histogram. Selection of the area where cells were located was done after inversion of the thresholded mask and filling of the holes in the images. The area of interest was selected via a binary propagation algorithm and the result was verified by comparison to the possible area range between 72 and 92 mm² (see figure 4.2, panel 4). If the detected area failed this verification, boundaries were estimated using

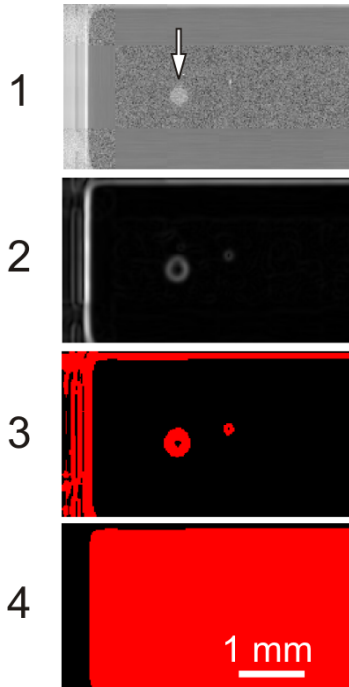


FIGURE 4.2 : Example of selection of cartridge scan area. 1: original FITC images of one side of a cartridge stitched together after application of linear convolution filter to border images (arrow indicates an air bubble), 2: border enhanced image by gradient magnitude filtering, 3: Binary image of thresholded borders (red color), 4: Selected scan area (red color) after inversion of image 3, binary propagation of center square, and size verification.

results from a fixed set of previously analyzed cartridges. This estimation had to be applied for 8 samples (of 370, 2%), whose cartridges had very irregular edges.

Step 2: Determine the optimal channel for segmentation

First step in automated image analysis is to determine where objects of interest are and what their outline is, so-called segmentation. The samples that were used for this step were all images from patient samples acquired before initiation of therapy. Segmentation techniques are generally contour based or region based: contour based techniques are aimed at finding the contour or edges of objects, usually by applying edge enhancement step(s). Region based methods can be subdivided in texture analysis, watershedding and intensity thresholding. The last, most common, method can be applied locally (i.e. in a part of an image) or globally (i.e. using all the images of one cartridge). Choosing the right segmentation method is critical for further processing and depends on the knowledge of the objects of interest. In our images, every object that was visible slightly above the background was of interest, therefore a very basic histogram-based thresholding algorithm was chosen which could segment objects that appeared in very low or high numbers: Zack's triangle threshold.

Figure 4.3 shows a comparison of segmentation methods using three

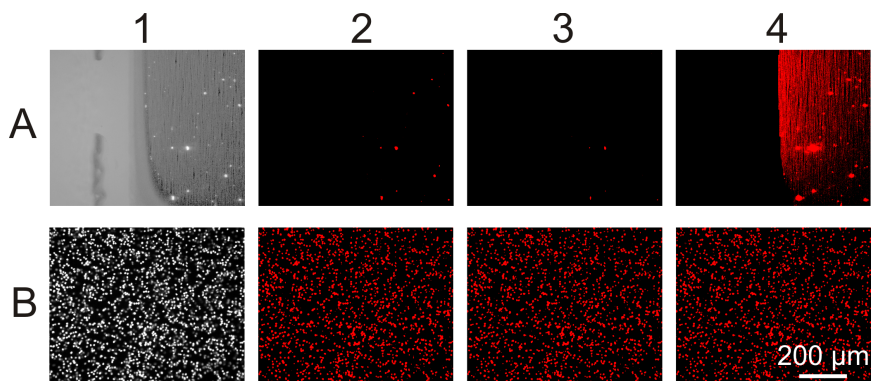


FIGURE 4.3 : Comparison of common thresholding procedures. Two original images containing a small (1A) and large number of objects (1B) were thresholded using three methods: triangle (2B and 2B), otsu (3A and 3B), and isodata (4A and 4B). The three methods give similar results on an image with a large number of objects, but triangle finds the correct number of objects in images which contain a small number of objects. Image A1 is shown using a logarithmic intensity scale to show the texture in the background; the left part of the image is part of the cartridge border.

methods on images containing low and high numbers of objects shown in panel 1A and 1B. Column 2 shows the triangle method applied on these images, column 3 the isodata method (iterative method to find average between foreground and background [18]), and column 4 the Otsu method (maximal inter-class variance [19]). Figure 4.3 panel 1A shows the image using a log scale to visualize the structure of the background. This background changed dramatically from cartridge to cartridge, not only in intensity (which could be compensated for by a common background subtraction method), but also in texture. Background subtraction was therefore not used, as this gave undesirable results (negative values, formation of extra contrast) when using a top hat based background subtraction procedure [20]. The proper method of background subtraction would be to record a black image with no objects present and subtracting this black image from the images with objects. However, such as black image was unfortunately unavailable. The task of the thresholding method was to capture everything just above this mixed background. The triangle threshold method is designed for this task. Figure 4.4 shows a detail of the image of figure 4.3, panel 1A, in which the three methods area again compared.

This algorithm was applied to the histogram of the total cartridge (minus the excluded border) to avoid local disturbances by for example air bubbles. Objects were segmented in six different channels: DAPI, FITC, PE, APC, sum of DAPI and PE, and a maximum intensity profile of all four channels. A constant offset of 20 was added to each threshold value for each channel in order to be as sensitive as possible, but not to segment large areas of noise. This offset was particularly necessary because of the shape

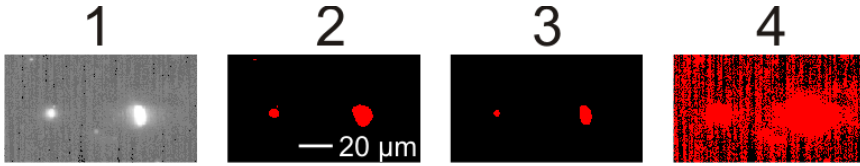


FIGURE 4.4 : Detail of a PE image (1), and masks as thresholded by the triangle (2), otsu (3), and isodata (4) methods.

of histogram after the conversion of intensity values from 8-bit to pseudo 12-bit. This conversion caused undesired valleys in the background part of the histogram, which could lead to threshold value which was too low. The sum of DAPI and PE channel was chosen because CellSearch selects objects positive in both PE and DAPI for review. The maximum intensity profile was expected to approach the actual outline of cells in the cartridge most accurately. Objects smaller than 9 pixels and larger than 2000 pixels were rejected. There were too many objects smaller than 9 pixels, most of which looked like noise on close examination. Objects larger than 2000 pixels ($\sim 900 \mu\text{m}^2$) cannot be single cellular objects. For each investigated segmentation channel the HR was determined after dichotomization of the patient group on the median number of segmented objects per patient. The channel, which yielded the highest HR was selected for further analysis.

Step 3: Select four independent properties with highest impact on HR

Objects can be characterized by measuring many features, see for rigorous examples on cell lines [21, 22]. For each object that was found in the selected segmentation method, 24 basic features were extracted; peak intensity, standard deviation of intensity, mean intensity and total intensity in the DAPI, FITC, PE, APC channels, as well as roundness, size, height and width, perimeter, first and second moments of inertia, and center of gravity in the segmentation channel. This set of basic features was chosen because the $10\times/0.45$ images do not show enough detail to measure higher level morphological features and the CTC are very heterogeneous. We required at least 15 samples per feature included in the classifier to prevent over fitting, four features were selected because there were only 68 samples in the healthy controls [23]. Selection of high impact features was achieved by univariate analysis. For each feature, objects were excluded if they were above or below a variable threshold. At each threshold tested, the HR was determined by dichotomizing patients on the median number of objects found in the patient population. The maximum obtained HR was determined for each feature and features were ranked according to HR. The covariance matrices were also derived for all features. Features with correlation coefficients less than 0.4 with all features ranked higher in HR were selected for further classifier development.

Step 4: Create classifiers and use bootstrap aggregating to test for robustness

A classifier required an object to fall within the inclusion criteria for each of the 4 features. 16464 different classifiers were created by varying the inclusion criteria independently. Performance of each classifier was tested against our training set of baseline and follow-up samples. To prevent over-training of the classifier due to possible high impact of a few dominating samples, the bootstrap aggregating method [24] was employed to test the robustness of each classifier. Each classifier was tested on 402 sets of 185 random pickings -with replacement- from the total of 185 samples. When picking n samples with replacement from a group with n samples using a large set, 63% of the samples are expected to be unique. Taking 402 sets of 185 random numbers results in each sample appearing approximately 402 times with a coefficient of variation of 5%. In this way, strong dependency on a few samples is averaged out. Furthermore, when deriving a HR after 402 picks, the spread of this HR provides a confidence interval (CI) and thus a comparative measure for stability. The optimal classifier needed to meet the following requirements:

1. High HR in both baseline and follow-up samples.
2. Large ratio between the number of classified objects found in patients and controls.
3. Very low absolute number of classified objects in the control samples.
4. HR from the follow-up samples higher than that of the baseline samples.
5. CI of the HR is not above average of the other classifiers.

4.3 RESULTS

4.3.1 PROCESSING OF IMAGES

Stored images from the 185 CRPC patients with samples taken at baseline and at first follow-up (2–5 weeks after baseline) and 68 healthy donors were imported. Figure 4.1 shows a schematic representation of the procedure that was used to process the images in such a way that an optimal classifier for selecting aCTC could be acquired. Automated counting and classification of objects took about 5 minutes per sample. Identification of candidate CTC in the current system takes approximately 5 minutes. However, classification of these objects by a human operator takes 0.5–39 minutes per sample (median 5.0 min, mean 7.9 min, SD 8.4 min, N=43) according to the main Veridex CellSearch processing lab. The reproducibility of counting aCTC by the algorithm is 100%. Detailed results of each step are described in the following sections.

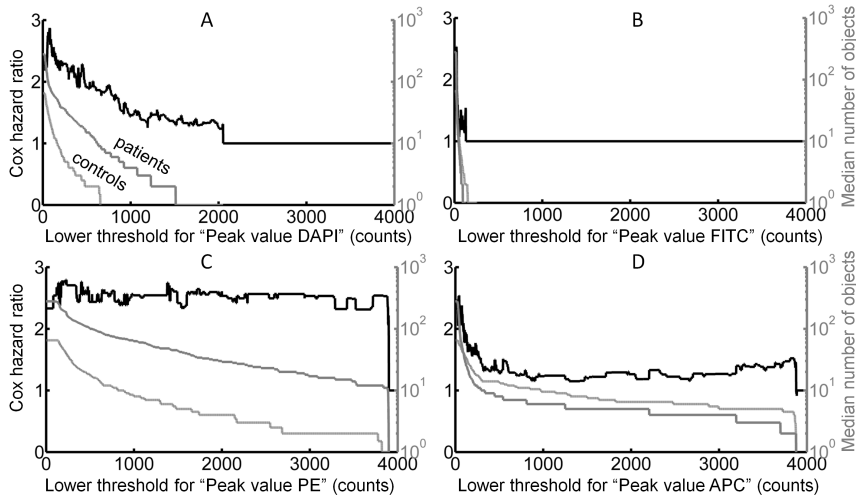


FIGURE 4.5 : Example of the influence of thresholding of a feature on the Cox hazard ratio. After segmentation of objects in the PE channel, the patient group is dichotomized on the median count and a base HR is derived. Varying a lower threshold from 0–4000 grey level counts on the feature "peak value" for all four channels DAPI, FITC, PE, and APC (panel A–D respectively) resulted in a varying HR (black line) and decreased the median number of objects of the patients and controls (grey lines) that were included.

4.3.2 OPTIMAL CHANNEL FOR SEGMENTATION

Segmentation in the four image channels, the PE/DAPI sum and the maximum intensity profile was performed on 185 baseline samples. The HR after dichotomization on the median number of objects found was 1.2 in the FITC channel, 1.3 in the APC channel, 1.6 in the DAPI and the DAPI/PE sum channels, 1.7 in the maximum intensity profile and 2.3 in the PE channel. In the 185 baseline samples a total of 262282 objects were found within the PE channel, these objects were used for the feature selection step in the algorithm development process.

4.3.3 FEATURES WITH HIGHEST IMPACT ON HR

The values of the 24 measured features were determined for each object. For each feature, all objects were tested against thresholds at various levels and for each level the HR was determined. The maximum HR was determined for each feature. Figure 4.5 shows the results for four of these analyses, including only events that exceeded a minimum "peak signal intensity" in the four channels. Panel A and C show that selecting on this feature - and thereby rejecting debris - results in a higher HR; these features are therefore of value. Panel B shows that objects have very low FITC peak values, HR is near 1 for any minimum value and already at low minimum value the count

drops to 0. Panels D shows that a minimum CD45-APC peak intensity results in lower HRs than the base HR after segmentation.

The object features with the best discriminatory power based on the above analysis were: Size (HR 3.0), total PE intensity (HR 3.0), total DAPI intensity (HR 3.0), standard deviation of the PE intensity (HR 2.8), mean PE intensity (HR 2.8), peak PE intensity (HR 2.8), peak DAPI intensity (HR 2.8), mean DAPI intensity (HR 2.8), total APC intensity (HR 2.8), standard deviation DAPI intensity (HR 2.6), and peak APC intensity (HR 2.5). After deriving the covariance matrix of these features, the ρ -correlation value between each feature was determined. The feature with the highest HR was chosen first. Next, features were chosen in descending order of HR provided they had a correlation of less than 0.4 with features already chosen. As a result the total intensity features were not selected as they had too much correlation with the size of the feature. Four features with both high influence on the HR and low inter-feature correlation were selected for further analysis:

1. Size of object
2. Standard deviation of signal in the PE-CK channel
3. Peak value of signal in the DAPI-DNA channel
4. Peak value of signal in the APC-CD45 channel

4.3.4 RULES TO ASSIGN aCTC

Acceptance ranges of features were varied simultaneously to construct 16464 different classifiers. In some classifiers the acceptance ranges for one or more features were set wide open, effectively reducing a classifier to 2 or 3 features. Each of these classifiers was tested by means of bootstrapping; the acquired variation in HR was used as a measure for the robustness of the classifier. Follow-up samples from 185 patients and 68 control samples were now included to determine the optimal classifier.

Figure 4.6A shows the average HR after bootstrapping of the baseline samples from the 16464 classifiers plotted against the HR on the follow-up samples, after patient groups were dichotomized on the median number of objects. The figure illustrates that many classifiers in this data set can be chosen that provide a higher HR for both baseline and follow-up samples as compared to the manual CellSearch CTC definition (indicated with the red cross in the figure). Figure 4.6B shows the performance of the classifiers on 4 of our 5 selection rules (the spread of the HR after bootstrapping is not shown in this figure). The optimal classifier is indicated by the arrow in figures 4.6A and 4.6B. In this data set, the selected classifier has only one background object classified as an aCTC in the 68 controls, a high mean bootstrapped HR of 4.2 and a HR for follow-up samples that is 1.5 times

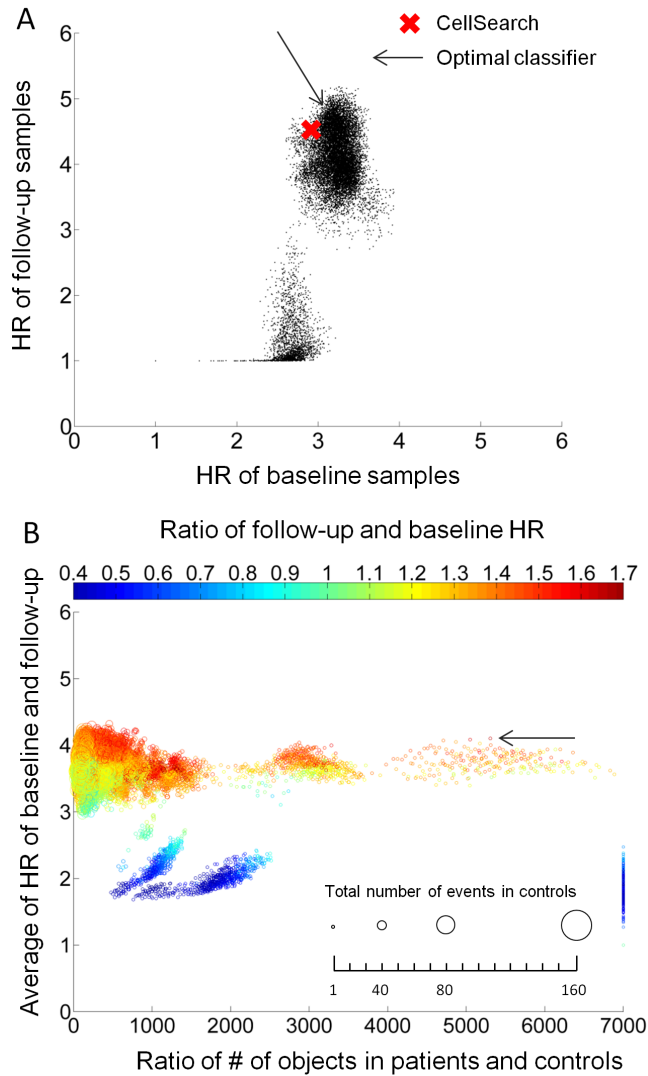


FIGURE 4.6 : Scatter plots showing the relation between the HR of 185 baseline and 185 follow-up samples and the 16464 classifiers. Panel A shows the HRs of baseline versus follow-up samples for all the 16464 classifiers. The arrow depicts our optimal classifier, the red cross depicts the HR derived using the CellSearch CTC definition. Panel B shows the mean HR of baseline and follow up samples versus the mean number of particles found in patient samples divided by the mean number of objects found in control samples of all the 16464 classifiers. This ratio was set to 7000 if there were no objects found in the control samples. The color indicates the HR of the follow-up samples divided by the HR of the baseline samples and the size of the circles indicates the number of objects classified in all 68 control samples. This size was set equal to 1 object if there were no objects found in the control samples.

higher than the baseline HR. The optimal classifier included objects with a CK-PE standard deviation higher than 50 counts, a size range of 75–500 pixels (34–224 μm^2), a DNA-DAPI peak value of at least 170 counts and a CD45-APC peak value less than 60 counts. Standard deviation (SD) of the spread of the HR after bootstrapping for this classifier was 0.7 and 1.3 for baseline and follow-up samples respectively.

4.3.5 AUTOMATIC CTC COUNT VERSUS MANUAL CTC COUNT IN PATIENTS AND CONTROLS

In the baseline samples the aCTC counts ranged from 0 to 3384 (total 14439, median 5, mean 78, SD 333) compared to mCTC counts of 0 to 5925 (total 18706, median 7, mean 101, SD 497). The R^2 -correlation between aCTC (x) and mCTC (y) was 0.80 (slope = 1.33, intercept = -3.03). In the follow-up samples the aCTC counts ranged from 0 to 870 (total 4992, median 2, mean 27, SD 86) compared to mCTC counts of 0 to 545 (total 5546, median 2, mean 30, SD 87). The R^2 -correlation was 0.67 (slope = 0.85, intercept = 7.18). In the 68 control samples 1 object was classified as aCTC and 0 objects as mCTC. Cox regression yielded a HR, by dichotomizing the patient group on the median baseline count, of 3.1 (95% CI 2.1–4.7) for aCTC and a HR of 2.9 (95% CI 2.0–4.4) for mCTC. For first follow-up samples we found a HR of 4.8 (95% CI 3.2–7.3) for aCTC and a HR of 4.5 (95% 3.0–6.8) for mCTC.

4.4 DISCUSSION

Multicenter prospective clinical studies have shown a significant relation between the presence of mCTC defined by the CellSearch CTC definition and poor progression-free and overall survival [2, 3, 4, 5, 6, 7, 8, 13]. It is envisioned that the clinical use of CTC studies will rapidly increase in the near future. Translational studies are, however, hindered by the laborious, manual CTC identification process, which increases study costs and is associated with variability of CTC classification by different operators. We have recently shown that all circulating EpCAM+CK+CD45- objects predict overall survival in castration-resistant prostate cancer and that different morphological definitions for CTC can be used, supporting the possibility that automated and clinically relevant CTC identification is feasible [12].

Utilizing overall survival from a cohort of patients with metastatic castration resistant prostate cancer, we have evaluated all the EpCAM positive circulating objects detected by ten CellSearch systems to determine the most clinically relevant automated definition for CTC identification. We have now shown that automated counting of EpCAM+ CK+ objects is fast, 100% reproducible and comparable to the manual count performed utilizing the established FDA-cleared method. It must be stressed that our

automatic count was only compared to the one manual count performed during the study [13]. In an earlier study it was found that in only 50% of cases two independent readers will read the same number of mCTC [11]. Although this does not always change the assignment of a patient as favorable or unfavorable, it shows how variable human perception can be.

Optimization of the aCTC classifier was a challenge during this research. There are many ways of testing and optimizing classifiers to improve robustness. However, most of these methods require a training set for which an absolute ground truth is known. Unfortunately in this research, this ground truth cannot be established. After enrichment of the blood sample, many objects still remain that have no influence on patient survival. The impact on survival of each object cannot be verified by inspecting it. The Cox hazard ratio and the number of objects in controls are our measure if a classifier is working properly. This indirect measure, which can only be evaluated by looking at a group of patients, cannot assure which objects are 100% responsible for patient survival. By using the bootstrap method we have chosen the classifier that was least influenced by certain dominating patients and gave consistent results for every subset of the total patient group. In this way we reduced the risk of over-training of the classifier. We found that the classifiers that were most promising had an equal spread in HR.

In the field of image processing, many pre-processing steps may be applied to the images that are recorded. Before segmentation, edge enhancement may be applied, as well as smoothing, background correction and many more. The impact of such procedures may increase or reduce the fidelity of a classifier in unpredictable ways. We did not apply any pre-processing to avoid making any presumptions about what kind of particles we were looking for. This led to an independent analysis of the aCTC definition.

The aCTC classifier that was chosen resulted in the identification of one object in the control samples, which is a very low false positive rate. Usually, such an auto-fluorescent object is discarded if it is positive in the FITC channel. The reviewers use this channel as a verification measure for debris. However, using the feature "maximum value in the FITC channel" in the classifier did not yield an improvement over the existing classifier. It was therefore concluded that some of the objects positive in the FITC channel were important for patient survival. This was surprising, as no fluorescent label was used in this channel. Bleed through of fluorescence signal from bright PE objects into this FITC channels was the likely cause for this observation. Not requiring a FITC feature is advantageous since this channel can now be used to not only detect the presence but also quantify the expression of biomarkers on the CTC such as Her-2 [25, 26, 27, 28].

For the purpose of the development of the aCTC classifier patients were divided into two groups based on the median aCTC count in the training set of 185 baseline samples. This division was chosen to minimize

statistical error in the HR. The median for the chosen aCTC classifier on the baseline samples was 5 aCTC and resulted in a HR of 3.1. The current standard CellSearch method presents to the reviewer CK+ DAPI+ objects for classification with a threshold of 5 or more mCTC being used to discriminate between those patients with a favorable and unfavorable prognosis. In an earlier study we argued that this threshold of 5 mCTC could be mainly attributed to error introduced by human interpretation [11]. As this error is eliminated by the automated method one could argue that the presence of any CTC with the automated method can be used to identify patients at risk.

Although the patient groups were dichotomized, it is clear that the presence of larger number of CTC is related to a poorer prognosis; determination of a meaningful increase or decrease in CTC upon treatment is however difficult to assess reliably, as the number of CTC that are detected are so low that a decrease or increase is often statistically insignificant. Exploration of alternative definitions that for example include tumor micro particles that occur in larger numbers may resolve this issue.

4.5 REFERENCES

- [1] W. J. Allard, J. Matera, M. C. Miller, M. Repollet, M. C. Connelly, C. Rao, A. G. J. Tibbe, J. W. Uhr, and L. Terstappen, "Tumor cells circulate in the peripheral blood of all major carcinomas but not in healthy subjects or patients with nonmalignant diseases," *Clinical Cancer Research*, vol. 10, no. 20, pp. 6897–6904, 2004.
- [2] G. T. Budd, M. Cristofanilli, M. J. Ellis, A. Stopeck, E. Borden, M. C. Miller, J. Matera, M. Repollet, G. V. Doyle, L. Terstappen, and D. F. Hayes, "Circulating tumor cells versus imaging - predicting overall survival in metastatic breast cancer," *Clinical Cancer Research*, vol. 12, no. 21, pp. 6403–6409, 2006.
- [3] S. J. Cohen, C. J. A. Punt, N. Iannotti, B. H. Saidman, K. D. Sabbath, N. Y. Gabrail, J. Picus, M. A. Morse, E. Mitchell, M. C. Miller, G. V. Doyle, H. Tissing, L. Terstappen, and N. J. Meropol, "Prognostic significance of circulating tumor cells in patients with metastatic colorectal cancer," *Annals of Oncology*, vol. 20, no. 7, pp. 1223–1229, 2009.
- [4] M. Cristofanilli, G. T. Budd, M. J. Ellis, A. Stopeck, J. Matera, M. C. Miller, J. M. Reuben, G. V. Doyle, W. J. Allard, L. Terstappen, and D. F. Hayes, "Circulating tumor cells, disease progression, and survival in metastatic breast cancer," *New England Journal of Medicine*, vol. 351, no. 8, pp. 781–791, 2004.
- [5] M. Cristofanilli, D. F. Hayes, G. T. Budd, M. J. Ellis, A. Stopeck, J. M. Reuben, G. V. Doyle, J. Matera, W. J. Allard, M. C. Miller, H. A. Fritsche, G. N. Hortobagyi, and L. Terstappen, "Circulating tumor cells: A novel prognostic factor for newly diagnosed metastatic breast cancer," *Journal of Clinical Oncology*, vol. 23, no. 7, pp. 1420–1430, 2005.
- [6] D. F. Hayes, M. Cristofanilli, G. T. Budd, M. J. Ellis, A. Stopeck, M. C. Miller, J. Matera, W. J. Allard, G. V. Doyle, and L. Terstappen, "Circulating tumor cells at each follow-up time point during therapy of metastatic breast cancer patients predict progression-free and overall survival," *Clinical Cancer Research*, vol. 12, no. 14, pp. 4218–4224, 2006.

- [7] D. Olmos, H. T. Arkenau, J. E. Ang, I. Ledaki, G. Attard, C. P. Carden, A. H. M. Reid, R. A'Hern, P. C. Fong, N. B. Oomen, R. Molife, D. Dearnaley, C. Parker, L. Terstappen, and J. S. de Bono, "Circulating tumour cell (ctc) counts as intermediate end points in castration-resistant prostate cancer (crpc): a single-centre experience," *Annals of Oncology*, vol. 20, no. 1, pp. 27–33, 2009.
- [8] J. Tol, M. Koopman, M. C. Miller, A. Tibbe, A. Cats, G. J. M. Creemers, A. H. Vos, I. D. Nagtegaal, L. Terstappen, and C. J. A. Punt, "Circulating tumour cells early predict progression-free and overall survival in advanced colorectal cancer patients treated with chemotherapy and targeted agents," *Annals of Oncology*, vol. 21, no. 5, pp. 1006–1012, 2010.
- [9] M. Kagan, D. Howard, T. Bendele, J. Mayes, J. Silvia, M. Repollet, J. Doyle, J. Allard, N. Tu, T. Bui, T. Russell, C. Rao, M. Hermann, H. Rutner, and L. Terstappen, "A sample preparation and analysis system for identification of circulating tumor cells," *Journal of Clinical Ligand Assay*, vol. 25, no. 1, pp. 104–110, 2002.
- [10] J. Kraan, S. Sleijfer, M. H. Strijbos, M. Ignatiadis, D. Peeters, J. Y. Pierga, F. Farace, S. Riethdorf, T. Fehm, L. Zorzino, A. G. Tibbe, M. Maestro, R. Gisbert-Criado, G. Denton, J. S. de Bono, C. Dive, J. A. Foekens, and J. W. Gratama, "External quality assurance of circulating tumor cell enumeration using the cellsearch((r)) system: A feasibility study," *Cytometry B Clin Cytom*, vol. 80B, no. 2, pp. 112–118, 2010.
- [11] A. G. J. Tibbe, M. C. Miller, and L. Terstappen, "Statistical considerations for enumeration of circulating tumor cells," *Cytometry Part A*, vol. 71A, no. 3, pp. 154–162, 2007.
- [12] F. A. W. Coumans, C. J. M. Doggen, G. Attard, J. S. de Bono, and L. W. M. M. Terstappen, "All circulating epcam+ck+cd45- objects predict overall survival in castration-resistant prostate cancer," *Annals of Oncology*, vol. 21, no. 9, pp. 1851–7, 2010.
- [13] J. S. de Bono, H. I. Scher, R. B. Montgomery, C. Parker, M. C. Miller, H. Tissing, G. V. Doyle, L. Terstappen, K. J. Pienta, and D. Raghavan, "Circulating tumor cells predict survival benefit from treatment in metastatic castration-resistant prostate cancer," *Clinical Cancer Research*, vol. 14, no. 19, pp. 6302–6309, 2008.
- [14] M. Abramoff, P. Magelhaes, and S. Ram, "Image processing with imagej," *Biophotonics International*, vol. 11, no. 7, pp. 36–42, 2004.
- [15] A. E. Carpenter, T. R. Jones, M. R. Lamprecht, C. Clarke, I. H. Kang, O. Friman, D. A. Guertin, J. H. Chang, R. A. Lindquist, J. Moffat, P. Golland, and D. M. Sabatini, "Cellprofiler: image analysis software for identifying and quantifying cell phenotypes," *Genome Biology*, vol. 7, no. 10, 2006.
- [16] M. Held, M. H. A. Schmitz, B. Fischer, T. Walter, B. Neumann, M. H. Olma, M. Peter, J. Ellenberg, and D. W. Gerlich, "Cellcognition: time-resolved phenotype annotation in high-throughput live cell imaging," *Nature Methods*, vol. 7, no. 9, pp. 747–U118, 2010.
- [17] G. W. Zack, W. E. Rogers, and S. A. Latt, "Automatic-measurement of sister chromatid exchange frequency," *Journal Of Histochemistry & Cytochemistry*, vol. 25, no. 7, pp. 741–753, 1977.
- [18] T. W. Ridler and S. Calvard, "Picture thresholding using an iterative selection method," *Ieee Transactions on Systems Man and Cybernetics*, vol. 8, no. 8, pp. 630–632, 1978.
- [19] N. Otsu, "Threshold selection method from gray-level histograms," *Ieee Transactions on Systems Man and Cybernetics*, vol. 9, no. 1, pp. 62–66, 1979.

- [20] F. Meyer, "Iterative image transformations for an automatic screening of cervical smears," *Journal Of Histochemistry & Cytochemistry*, vol. 27, no. 1, pp. 128–135, 1979.
- [21] C. Bakal, J. Aach, G. Church, and N. Perrimon, "Quantitative morphological signatures define local signaling networks regulating cell morphology," *Science*, vol. 316, no. 5832, pp. 1753–1756, 2007.
- [22] Z. E. Perlman, M. D. Slack, Y. Feng, T. J. Mitchison, L. F. Wu, and S. J. Altschuler, "Multidimensional drug profiling by automated microscopy," *Science*, vol. 306, no. 5699, pp. 1194–1198, 2004.
- [23] P. Peduzzi, J. Concato, A. R. Feinstein, and T. R. Holford, "Importance of events per independent variable in proportional hazards regression analysis .2. accuracy and precision of regression estimates," *Journal of Clinical Epidemiology*, vol. 48, no. 12, pp. 1503–1510, 1995.
- [24] L. Breiman, "Bagging predictors," *Machine Learning*, vol. 24, no. 2, pp. 123–140, 1996.
- [25] S. D. Meng, D. Tripathy, S. Shete, R. Ashfaq, B. Haley, S. Perkins, P. Beitsch, A. Khan, D. Euhus, C. Osborne, E. Frenkel, S. Hoover, M. Leitch, E. Clifford, E. Vitetta, L. Morrison, D. Herlyn, L. Terstappen, T. Fleming, T. Fehm, T. Tucker, N. Lane, J. Q. Wang, and J. Uhr, "Her-2 gene amplification can be acquired as breast cancer progresses," *Proceedings Of The National Academy Of Sciences Of The United States Of America*, vol. 101, no. 25, pp. 9393–9398, 2004.
- [26] D. F. Hayes, T. M. Walker, B. Singh, E. S. Vitetta, J. W. Uhr, S. Gross, C. Rao, G. V. Doyle, and L. Terstappen, "Monitoring expression of her-2 on circulating epithelial cells in patients with advanced breast cancer," *International Journal of Oncology*, vol. 21, no. 5, pp. 1111–1117, 2002.
- [27] S. Meng, D. Tripathy, S. Shete, R. Ashfaq, H. Saboorian, B. Haley, E. Frenkel, D. Euhus, M. Leitch, C. Osborne, E. Clifford, S. Perkins, P. Beitsch, A. Khan, L. Morrison, D. Herlyn, L. Terstappen, N. Lane, J. Wang, and J. Uhr, "upar and her-2 gene status in individual breast cancer cells from blood and tissues," *Proceedings Of The National Academy Of Sciences Of The United States Of America*, vol. 103, no. 46, pp. 17361–17365, 2006.
- [28] S. T. Lighthart, F. A. W. Coumans, G. Attard, A. Mulick Cassidy, J. S. De Bono, and L. W. M. M. Terstappen, "Unbiased and automated identification of a circulating tumour cell definition that associates with overall survival," *Plos One*, 2011.

UNBIASED AND AUTOMATED IDENTIFICATION OF A CIRCULATING TUMOUR CELL DEFINITION THAT ASSOCIATES WITH OVERALL SURVIVAL¹

Sjoerd T. Ligthart, Frank A.W. Coumans, Gerhardt Attard, Amy Mulick Cassidy, Johann S. de Bono, and Leon W.M.M. Terstappen

ABSTRACT

Circulating tumour cells (CTC) in patients with metastatic carcinomas are associated with poor survival and can be used to guide therapy. Classification of CTC however remains subjective, as they are morphologically heterogeneous. We acquired digital images, using the CellSearch[®] system, from blood of 185 castration resistant prostate cancer (CRPC) patients and 68 healthy subjects to define CTC by computer algorithms. Patient survival data was used as the training parameter for the computer to define CTC. The computer-generated CTC definition was validated on a separate CRPC dataset comprising 100 patients. The optimal definition of the computer defined CTC (aCTC) was stricter as compared to the manual CellSearch CTC (mCTC) definition and as a consequence aCTC were less frequent.

¹published as S.T. Ligthart, F.A.W. Coumans, G. Attard, A. Mulick Cassidy, J. S. De Bono, and L.W.M.M. Terstappen "Unbiased and Automated Identification of a Circulating Tumour Cell Definition that Associates with Overall Survival" *Plos One* 2011, DOI:10.1371/journal.pone.0027419

The computer-generated CTC definition resulted in hazard ratios (HRs) of 2.8 for baseline and 3.9 for follow-up samples, which is comparable to the mCTC definition (baseline HR 2.9, follow-up HR 4.5). Validation resulted in HRs at baseline/follow-up of 3.9/5.4 for computer and 4.8/5.8 for manual definitions. In conclusion, we have defined and validated CTC by clinical outcome using a perfectly reproducing automated algorithm.

5.1 INTRODUCTION

In recent years, several studies have reported that a change in circulating tumour cell (CTC) count could indicate whether a therapy for advanced cancer is effective [1, 2, 3, 4, 5, 6, 7]. It is envisioned that the clinical use of CTC as a pharmacodynamic and predictive biomarker will rapidly increase in the near future, especially in advanced prostate and breast cancers [8]. Currently, the CellSearch[®] method is the only clinically validated and FDA-cleared method for CTC enumeration [9]. In this system, objects that are positive for epithelial cell adhesion molecule (EpCAM) antigen are enriched from 7.5 ml of blood and then stained with cytokeratin-phycoerythrin (CK-PE), CD45-allophycocyanin (CD45-APC) and the nuclear dye 4',6-diamidino-2-phenylindole (DAPI). The recorded fluorescence images of CK-PE, DNA-DAPI, CD45-APC and a debris-fluorescein (FITC) channel are segmented on the basis of being positive for CK-PE and DAPI and are then presented to a trained reviewer for identification of CTC that are CK-PE positive, CD45-APC negative, 4 μm in diameter, DAPI-positive, and have a cell-like morphology.

This manual procedure is laborious, time-consuming and can be highly subjective. For example, others have described an inter-reviewer variation in manual CTC enumeration of 4% to 31% (median 14%) [10]. Moreover, CTC are known to be morphologically heterogeneous and in fact, different laboratories have used different definitions for what constitutes a CTC, especially for objects that are dead or apoptotic [2, 10]. CTC can occur at very low frequencies and therefore misjudging a few events could be very significant [11]. Also, the definition of what to call a CTC that is currently used may not be optimal. A recent report showed that tumour micro particles (TMPs, objects that are smaller than 4 μm and EpCAM+CK+CD45-) are also associated with unfavorable prognosis, suggesting alternative definitions for CTC evaluation should be considered [12].

Here we present the results of a new approach to identify CTC in images captured by the system in samples from castration-resistant prostate cancer (CRPC) patients. We recorded images before treatment (baseline samples) and from a follow-up sample taken 2–6 weeks after start of therapy. Our hypothesis was that using survival data as the only training parameter, an automated algorithm could be optimized to define and count CTC with the same fidelity as the manual CellSearch method (mCTC). This algorithm needs to identify automated CTC (aCTC) candidates, characterize them

and compare the candidates to a range of known features. Replacement of manual CTC counting with an automated method would significantly reduce cost and more importantly, eliminate inter- and intra-laboratory variation that could be clinically important in cases with low CTC counts. Moreover, a consensus definition for what constitutes a CTC is urgently required [13]. By using an unbiased approach to identify clinically important events, our analyses informs on the validity of different criteria currently being used, which were validated on an independent data set.

5.2 METHODS

5.2.1 ETHICS STATEMENT

Development of image analysis algorithms for automated CTC enumeration was performed on stored images from ten CellSearch systems (Veridex LLC, Raritan, NJ) from patients participating in the prospective IMMC-38 study (NCT00133900) and healthy individuals participating in the IMMC-06 study (NCT00133913) were available [7, 14]. For validation of the algorithm, images were used from samples from patients participating in Phase I and II clinical studies of abiraterone acetate (NCT00473512) conducted at the Royal Marsden NHS Foundation Trust and reported previously [15, 16, 17]. Samples were processed at The Institute of Cancer Research (ICR) (Sutton, UK) and archived images were sent for automated analysis at the University of Twente (Netherlands). The University of Twente was blinded to survival data for the validation samples. These studies were approved by the Ethics Review Committees of the participating centres: the United States Institutional Review Board for IMMC-38; the United States Food and Drug Administration and the United Kingdom Medicines and Healthcare Products Regulatory Agency for abiraterone acetate. All patients and healthy individuals provided written informed consent.

5.2.2 PARTICIPANTS

All patients had histologically confirmed prostate adenocarcinoma, castrate levels of testosterone (<50 ng/ml) and progressive disease as defined by three consecutively rising PSA values [18]. Patients included in IMMC-38 were commencing a new cytotoxic therapy. Patients with brain metastases or a history of other malignancies within the last 5 years were excluded. 276 patients were enrolled in IMMC-38, 231 met eligibility criteria and for 185 of those patients images could be imported for baseline and first follow-up [14]. Baseline samples were taken up to 19 days prior to commencement of a new cytotoxic chemotherapy, follow-up samples were taken 2-6 weeks after the start of therapy. 121 patients started their first line of chemotherapy. A total of 65 clinical centres in the United States and Europe participated in this study.

In the abiraterone acetate studies, samples were collected from a total of 100 patients. 89 patients contributed both a baseline and a follow-up sample, 7 contributed only follow-up, 4 only baseline. Fifty-one patients were chemotherapy naïve, and 44 patients were docetaxel-pretreated. Samples collected up to 14 days before initiation of abiraterone acetate (93 samples) and after one cycle (28 days) of therapy (96 samples) were used for this analysis. Of 185 IMMC-38 patients 118 (64%) died during the study, in the abiraterone acetate studies 73 of 100 (73%) died. Median survival was 20.7 months for IMMC-38 and 31.5 months for abiraterone acetate. Median duration of follow-up for censored patients was 29.8 months for IMMC-38 and 41.8 months for abiraterone acetate. In addition, samples of 68 healthy individuals participating in the IMMC-06 study were available [7]. Healthy individuals donated blood at three clinical centres in the US, the Netherlands, and the United Kingdom.

5.2.3 MANUAL COUNTING OF CIRCULATING TUMOUR CELLS (mCTC)

The CellSearch system was used to isolate and image EpCAM+ objects. The CellSearch system consists of a CellTracks Autoprep for sample preparation [2, 9] and a CellTracks Analyzer II for sample analysis. The CellTracks Autoprep immuno-magnetically enriches epithelial cells from 7.5 ml of blood using ferrofluids conjugated to epithelial cell adhesion molecule antibodies (EpCAM). The enriched sample is stained with phycoerythrin-conjugated (PE) antibodies directed against cytokeratins 8, 18, and 19 (CK), an allophycocyanin-conjugated (APC) antibody to CD45 and the nuclear dye 4',6-diamidino-2-phenylindole (DAPI). This enriched sample is transferred to a magnetic cartridge where all ferrofluid labeled objects are pulled towards an analysis surface. The entire analysis surface is imaged by the CellTracks Analyzer II, a four-color semi-automated fluorescence microscope that captures digital images for four different fluorescent dyes using a 10×/0.45NA objective. In addition to the DAPI, PE and APC images, a fourth fluorescence channel (emission 53 ± 25 nm) is imaged as a control channel for exclusion of auto-fluorescent debris. This channel will be termed "FITC" channel. Per cartridge, a whole scan consists of 144–180 4-layer tiff images that are saved for each patient. After imaging by the Celltracks analyzer, the software selects objects that are DNA and CK positive and presents them to an operator in a thumbnail gallery. The operators are trained to reviews these galleries to select the mCTC among the objects. An mCTC is positive for DNA and CK, is negative for CD45, is larger than $4 \times 4 \mu\text{m}$ and has morphological features that are consistent with those of a cell.

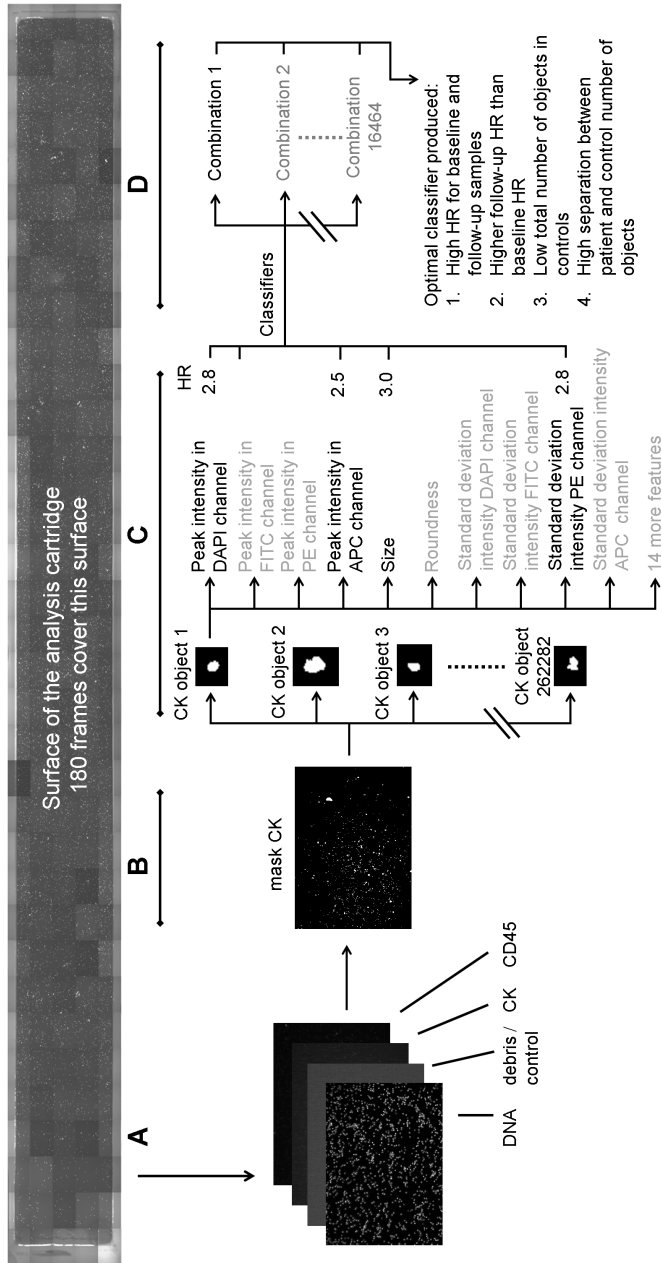


FIGURE 5.1 : Schematic overview of the aCTC classifier development process. A: importing of images; B: object segmentation; C: feature measurements; D: classification of aCTC.

5.2.4 AUTOMATED COUNTING OF EPCAM+ OBJECTS USING A COMPUTER ALGORITHM (aCTC)

CDs containing up to 180 archived four channel tiff images for each sample belonging to the respective studies were collected for import to a central hard drive. Objects were detected and classified using an automated algorithm developed in Matlab 2009a (Mathworks, Natick, MA) using the DIPimage toolbox (www.diplib.org). An outline of the method is given below and shown in figure 5.1. The method was applied for each patient sample separately. First, the true imaging area where all the objects were located was determined via sample cartridge edge detection in the debris-FITC channel. Candidate CTC objects were selected via object segmentation in the CK-PE channel. Segmentation was performed using a threshold which was determined for each sample via the channel image histogram [19]. Applying this threshold to the CK-PE images returned the outline, size and location of the objects.

In the next step, measurements, termed features from here on, were performed on these objects and the features providing the largest Cox hazard ratio (HR, shown next to the features in figure 5.1) and low correlation with other selected features were chosen for classification of these objects: the standard deviation of the signal in the CK-PE channel, the peak signal value in both the DNA-DAPI and CD45-APC channels and the size of the objects. Finally, selection of aCTC was performed by comparing every object to numerical inclusion criteria for these four features. The combined inclusion criteria -termed classifier from here on- were varied to find the aCTC definition that most strongly associated with high HR for baseline and follow-up samples, a higher HR for follow-up than baseline samples, and a low relative and absolute count in control samples. Bootstrap aggregation was used to test the stability of the optimal classifier [20].

5.2.5 STATISTICAL ANALYSIS

The primary objective was to identify a CTC definition with the largest HR between favourable and unfavourable patient groups and a low background in the control group. During algorithm development, the median number of events found was used to dichotomize patients into two groups. This approach allowed quick selection of a threshold, while ensuring that sufficient patients are present in both the 'at risk' and the 'not at risk' groups. It also allowed comparison of HR determined for different features and minimized error in HR. For this approach to work there needs to be a continuous relationship between survival and CTC count, which was previously demonstrated for the IMMC-38 data [21]. After algorithm optimization, all patient samples were processed by the algorithm and the training and validation patient groups were dichotomized on cut-off values ranging 1–10 of aCTC and mCTC to derive HR and median overall survival (OS) for baseline and

first follow-up samples. Furthermore, a linear regression was performed for comparison between the aCTC and mCTC count. Classifiers with reduced features were tested to determine the impact of each feature. Pearson coefficient of determination R^2 was determined between these populations using Matlab. Statistical HR and Kaplan Meier analysis for the training dataset was performed by S.T.L. using Matlab and GraphPad Prism v5. Statistical HR and Kaplan Meier analysis for the validation dataset was performed by A.M.C. at the ICR using Stata v10.1 (StataCorp) and GraphPad Prism v5.

5.3 RESULTS

5.3.1 CHOOSING THE OPTIMAL CLASSIFIER AND PROCESSING OF SAMPLES

The classifier resulting in the optimal aCTC definition that most strongly associated with high HR for baseline and follow-up samples was chosen. The features that most strongly associated with OS were: a CK-PE standard deviation >50 counts, a size range of 75–500 pixels (34–224 μm^2), a DAPI-DNA peak value >170 counts and a CD45-APC peak value <60 counts. For every patient sample, the objects meeting these inclusion criteria were added up to arrive at a final aCTC count per patient. The mCTC count was performed by trained reviewers. Time needed for preparation of images for mCTC assignment was similar to the time needed for complete aCTC enumeration: both took 5 minutes. However, enumeration of these mCTC by a human operator takes an additional 8 minutes per sample on average (median 5, range 1–39, SD 8 minutes, $N=43$).

5.3.2 AUTOMATED CTC COUNT COMPARED TO MANUAL CTC COUNT IN PATIENTS AND CONTROLS

After all the objects meeting the criteria of the optimal classifier were summed for each patient sample, the aCTC count was compared with the mCTC count. In the baseline samples the aCTC counts ranged from 0 to 3384 (median 5, mean 78, SD 333) compared to mCTC counts of 0 to 5925 (median 7, mean 101, SD 497). The R^2 between aCTC and mCTC was 0.80 (slope = 1.33, intercept = -3.03), indicating that were more mCTC counted than aCTC. In the follow-up samples aCTC counts ranged from 0 to 870 (median 2, mean 27, SD 86) compared to mCTC counts of 0 to 545 (median 2, mean 30, SD 87). The R^2 was 0.67 (slope = 0.85, intercept = 7.18). Figure 5.2 shows a scatter plot of the baseline and follow-up samples with the linear regression and corresponding statistics of the combined baseline and follow-up samples. In 68 control samples only one object was classified as aCTC and zero objects as mCTC.

Figure 5.3 shows the frequency distributions of mCTC (median 7) and aCTC (median 5) in baseline samples from patients and in controls for the

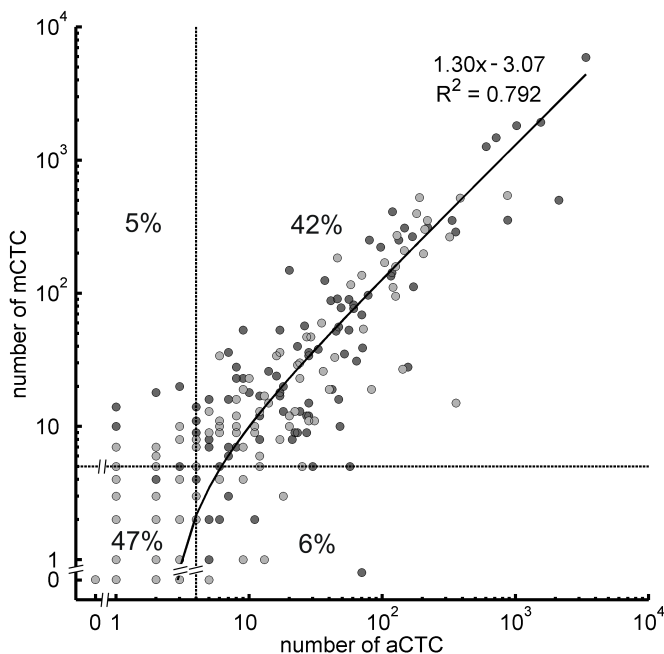


FIGURE 5.2 : Scatter plot of baseline (dark grey) and first follow-up (light grey) samples, counted by the aCTC and mCTC methods. The linear regression statistics apply to the total data set. Quadrants were defined by the clinically used cut-off value of 5 mCTC and the empirical determined value of 4 aCTC (dashed lines). In each quadrant, the percentage of patients is shown.

optimal aCTC definition and three other definitions that are less strict: without the CD45 feature (median 10), without the DAPI feature (median 40), and for TMP objects that are EpCAM+CK+CD45- (CK standard value >10 counts; CD45 peak value <60 counts) and $<4 \mu\text{m}$ in diameter (median 104). R^2 between aCTC and mCTC was 0.78. Between aCTC and the objects found with the classifiers without CD45 and DAPI the R^2 were 0.95 and 0.82, respectively. Between the aCTC and TMP definition, the R^2 was 0.56 ($p < 0.0001$ for all R^2).

5.3.3 DEFINING CUT-OFF VALUES FOR aCTC AND mCTC

To arrive at a clinically relevant cut-off value for aCTC comparable to the mCTC cut-off of 5 used in routine clinical practice, we used the linear regression slope of 1.33 between aCTC and mCTC baseline measurements. This resulted in a cut-off of 4 aCTC. In the scatter plot of figure 5.2 the cut-offs are indicated creating four quadrants: two with concordant and two with discordant results. The percentage of patients in each quadrant

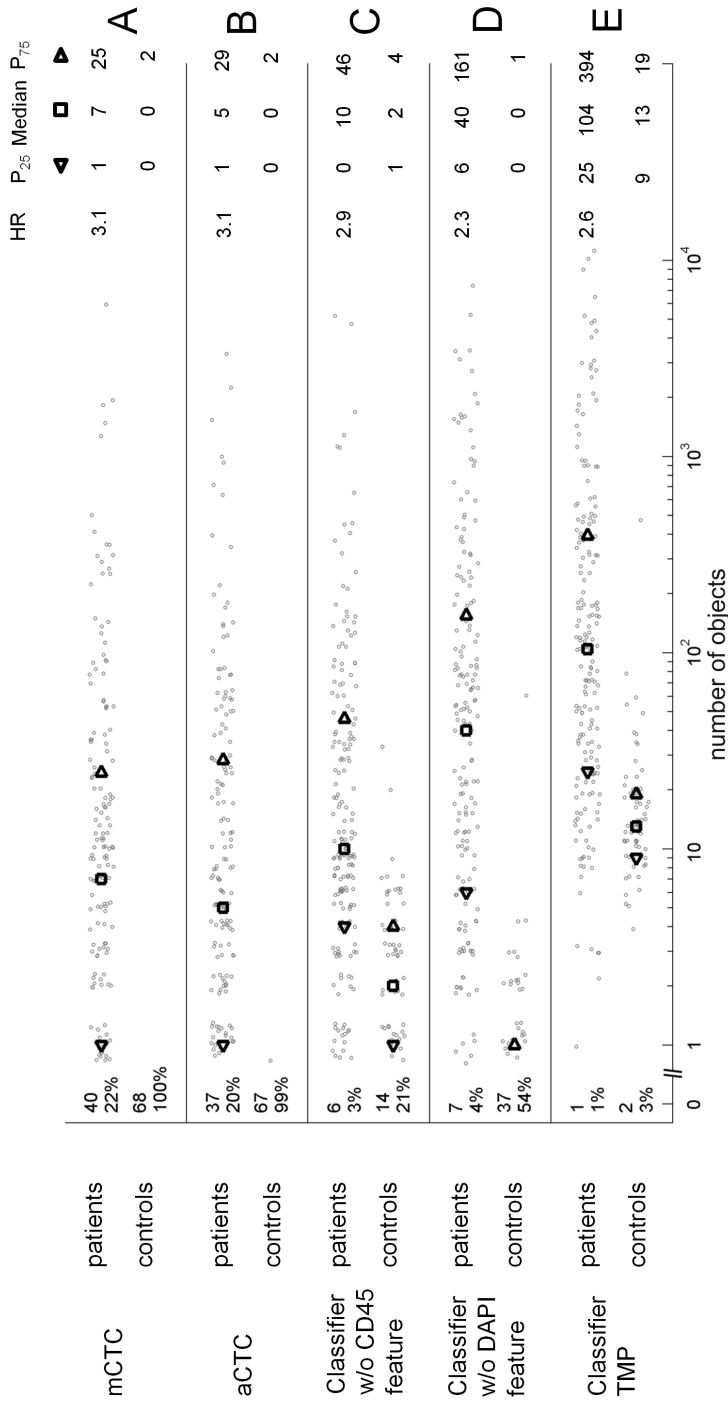


FIGURE 5.3 : Frequency distributions of mCTC and aCTC from patients (N=185) and control samples (N=68). The top row shows the mCTC frequency distribution (panel A). The next rows show the number of aCTC for the optimal aCTC definition (panel B), the classifier without the CD45 exclusion criterion (panel C), without the DAPI criterion (panel D), and of TMP objects that are EpCAM+CK+CD45- and <4 μ m in diameter (panel E). The percentage of patients with 0 objects is shown numerically on the left. On the right the HRs derived by dichotomizing on the median number of objects in patients are shown, together with the 25, 50, and 75 percentiles. These percentiles are also indicated in the plot.

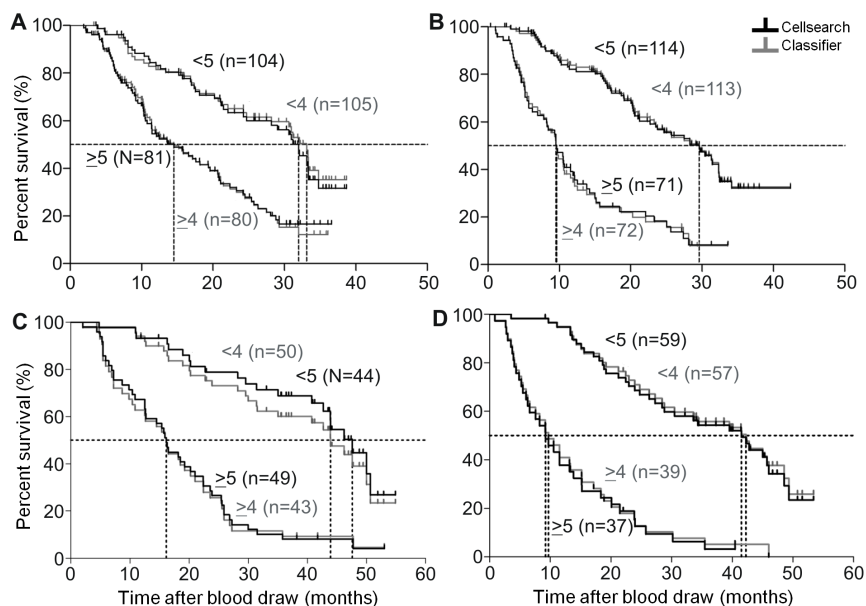


FIGURE 5.4 : Kaplan-Meier plots of the classifier (grey lines) and the manual CellSearch (black lines) definition. The training set is shown in panel A (baseline, $N=185$) and panel B (follow-up, $N=185$). Kaplan-Meier plots for the validation set are shown in panel C (baseline, $N=93$) and panel D (follow-up, $N=96$). Censoring is indicated by vertical marks on the Kaplan-Meier plot.

is provided. A total of 11% of patients had discordant results based on the CTC cut-offs of mCTC and aCTC. The influence of other CTC cut-off values from ≥ 1 to ≥ 10 CTC on the number of patients affected, the median OS, HR and its significance for both baseline and follow-up samples were determined and are shown in table 5.1 (p-values for all HRs < 0.0001 , except baseline cut-off=1: $p=0.0003$ for aCTC and $p=0.004$ for mCTC).

Kaplan-Meier plots were generated for 185 baseline and 185 follow-up samples using the standard cut-off value of 5 for mCTC and the cut-off value for aCTC of 4. Figure 5.4, panel A shows the Kaplan-Meier plot for the baseline samples. Cox regression yielded a HR of 2.8 (95% CI 1.9–4.1) for aCTC and a HR of 2.9 (95% CI 2.0–4.4) for mCTC. Figure 5.4, panel B shows the Kaplan-Meier plot for the follow-up samples. For the first follow-up samples we found a HR of 3.9 (95% CI 2.6–5.9) for aCTC and a HR of 4.5 (95% CI 3.0–6.8) for mCTC.

5.3.4 VALIDATION OF AUTOMATED CTC COUNT

To validate the aCTC count an independent data set was used from 100 metastatic prostate cancer patients treated with abiraterone acetate. The

TABLE 5.1 : aCTC and mCTC cut-off values with HRs, median OS, for baseline and first follow-up samples for the training set enrolled in the IMMC-38 study.

	Classifier			Cellsearch			HR	
	CO ^a n ≥ CO (%)	OS ^b 50% OS < CO	50% OS ≥ CO	HR ^c	n ≥ CO (%)	50% OS < CO		50% OS ≥ CO
Baseline (n=185)								
1	147 (79)	33.0	17.5	2.6 (1.6-4.5)	145 (78)	32.0	17.4	2.0 (1.2-3.2)
2	126 (68)	33.3	15.8	3.0 (1.9-4.8)	128 (69)	33.3	15.6	2.8 (1.8-4.4)
3	115 (62)	33.3	15.2	3.0 (1.9-4.6)	118 (64)	33.1	15.2	2.9 (1.9-4.4)
4	105 (57)	33.1	14.5	2.8 (1.9-4.1)	108 (58)	32.1	14.5	2.5 (1.7-3.7)
5	93 (50)	31.2	13.6	3.1 (2.1-4.7)	104 (56)	32.0	14.5	2.9 (2.0-4.4)
6	87 (47)	32.0	12.5	3.0 (2.0-4.5)	96 (52)	32.0	13.6	2.8 (1.9-4.2)
7	82 (44)	31.0	11.5	3.2 (2.2-4.8)	95 (51)	32.0	13.6	3.1 (2.1-4.5)
8	75 (41)	30.6	11.4	3.1 (2.1-4.6)	87 (47)	31.2	11.5	3.2 (2.1-4.7)
9	72 (39)	30.6	11.4	3.0 (2.0-4.4)	84 (45)	31.2	11.4	3.4 (2.3-5.0)
10	70 (38)	29.0	11.4	2.9 (1.9-4.3)	78 (42)	30.6	10.6	3.2 (2.1-4.7)
Follow-up (n=185)								
1	129 (70)	31.4	15.2	2.3 (1.6-3.5)	110 (59)	33.4	12.5	3.0 (2.1-4.5)
2	98 (53)	29.6	11.6	3.6 (2.4-5.3)	94 (51)	31.1	11.5	3.9 (2.6-5.7)
3	79 (43)	29.6	9.8	4.0 (2.7-5.9)	81 (44)	31.2	10.3	4.4 (2.9-6.5)
4	72 (39)	29.6	9.6	3.9 (2.6-5.9)	77 (42)	31.2	10.3	3.8 (2.5-5.6)
5	67 (36)	28.3	9.5	4.8 (3.2-7.3)	71 (38)	30.6	10.5	4.5 (3.0-6.8)
6	62 (34)	28.3	9.4	4.4 (2.9-6.6)	64 (35)	29.3	10.2	4.7 (3.1-7.1)
7	57 (31)	25.7	9.4	4.4 (2.9-6.6)	62 (34)	29.3	10.2	4.5 (3.0-6.8)
8	57 (31)	25.7	9.4	4.5 (2.9-6.8)	58 (31)	28.2	9.2	4.7 (3.1-7.1)
9	53 (29)	25.7	9.4	4.0 (2.6-6.1)	55 (30)	28.2	9.2	4.7 (3.1-7.2)
10	49 (26)	25.6	9.4	3.9 (2.6-6.0)	51 (28)	27.5	9.2	4.3 (2.8-6.6)

^a Cut-off. ^b Overall Survival. ^c Hazard Ratio.

set included 93 baseline samples 96 follow-up samples. In the baseline samples the aCTC counts range was 0–1258 (median 3, mean 46, SD 152) and a range of 0–1108 (median 6, mean 53 SD 151) was found for mCTC. R^2 between aCTC and mCTC was 0.28 (slope = 0.52, intercept = 28.76). Exclusion of a single outlier resulted in a R^2 of 0.90 (slope 1.72, intercept 4.07). In the follow-up samples the aCTC counts range was 0–2490 (median 2, mean 78, SD 326) and a range of 0–3573 (median 2, mean 74, SD 390) with mCTC. R^2 between aCTC and mCTC both was 0.83 (slope = 1.09, intercept = -11.43). Kaplan-Meier plots were generated for 93 baseline and 96 follow-up samples using the standard CTC cut-off of 5 for mCTC and 4 for aCTC. Figure 5.4 panel C shows the Kaplan-meier plot from the baseline samples. Cox regression resulted in a HR of 3.9 (95% CI 2.4–6.6) for aCTC and a HR of 4.8 (95% CI 2.8–8.3) for mCTC. The Kaplan-Meier plot from follow-up samples is presented in figure 5.4, panel D. A HR of 5.4 (95% CI 3.2–8.9) was found for aCTC and a HR of 5.8 (95% CI 3.4–9.8) for mCTC (p-values for all HRs <0.0001). Table 5.2 shows the influence of other cut-off values on the HR and OS.

5.4 DISCUSSION

This is the first report of an algorithm-based automated method for unbiased determination of a clinically significant definition for what constitutes a CTC. We used stored images recorded by the CellSearch system from 185 patients with metastatic CRPC. While the algorithm was developed using patients with metastatic CRPC receiving a cytotoxic agent (training cohort), validation on patients receiving the highly active hormonal agent abiraterone acetate (validation cohort) confirmed reproducibility of the enumeration algorithm.

For the purpose of the development of the CTC classifier, patients were divided into two groups based on the median aCTC count in the training set of 185 baseline samples. This division was chosen to minimize statistical error in the HR. The median count for the chosen aCTC classifier on the baseline samples was 5 aCTC and resulted in a HR of 3.1. The current standard CellSearch method presents the reviewer with CK+ DAPI+ objects for classification: a threshold of 5 or more mCTC is used to discriminate between patients with a favourable versus an unfavourable prognosis. In an earlier study we reported that this threshold of 5 mCTC could be mainly attributed to error introduced by human interpretation [11]. The variability of counting aCTC by the algorithm is 0% compared to inter-reviewer variability of 4% to 31% for mCTC (median 14%) [10]. As we have eliminated this variability by using an automated method, one could argue that the presence of any CTC could now be used to identify patients at risk.

To identify a threshold for aCTC we used the correlation statistics

TABLE 5.2 : aCTC and mCTC cut-off values with HRs, median OS, for baseline and first follow-up samples for the validation set of patients enrolled in Phase I/II trials of abiraterone acetate at Royal Marsden Hospital.

	Classifier			Cellsearch			HR
	CO ^a n ≥ CO (%)	50% OS ^b < CO	50% OS ≥ CO	n ≥ CO (%)	50% OS < CO	50% OS ≥ CO	
Baseline (n=93)							
1	69 (74)	50.0	20.3	60 (65)	50.0	18.5	3.8 (2.1–6.9)
2	54 (58)	46.2	18.1	56 (60)	47.6	18.1	4.0 (2.3–7.0)
3	49 (53)	46.2	16.3	50 (54)	47.6	16.1	4.8 (2.8–8.4)
4	43 (46)	44.0	16.1	49 (53)	47.6	16.1	4.8 (2.8–8.3)
5	43 (46)	44.0	16.1	49 (53)	47.6	16.1	4.8 (2.8–8.3)
6	41 (44)	43.9	15.9	47 (51)	46.2	16.1	4.4 (2.6–7.4)
7	37 (40)	43.9	15.4	45 (48)	44.0	16.3	3.9 (2.3–6.5)
8	36 (39)	42.7	14.5	43 (46)	44.0	16.1	3.8 (2.3–6.4)
9	35 (38)	42.7	15.4	41 (44)	43.9	15.9	3.6 (2.2–5.9)
10	33 (35)	40.7	15.4	40 (43)	43.9	15.7	3.5 (2.1–5.7)
Follow-up (n=96)							
1	58 (60)	48.5	13.8	55 (57)	48.5	13.8	5.4 (3.1–9.6)
2	51 (53)	45.0	13.3	50 (52)	45.0	11.5	5.0 (2.9–8.5)
3	40 (42)	42.2	9.7	44 (46)	42.2	11.1	4.8 (2.9–8.0)
4	39 (41)	42.2	9.7	40 (42)	42.2	9.7	6.6 (3.8–11.4)
5	37 (39)	41.5	9.7	37 (39)	41.5	9.2	5.8 (3.4–9.8)
6	34 (35)	41.5	9.2	36 (38)	41.5	9.2	6.2 (3.7–10.5)
7	33 (34)	41.5	9.2	36 (38)	41.5	9.2	6.2 (3.7–10.5)
8	32 (33)	41.5	9.2	35 (36)	41.5	9.2	5.9 (3.5–10.0)
9	30 (31)	39.5	9.2	32 (33)	39.5	7.6	5.3 (3.2–8.9)
10	28 (29)	34.4	9.2	32 (33)	39.5	7.6	5.3 (3.2–8.9)

^a Cut-off. ^b Overall Survival. ^c Hazard Ratio.

between mCTC and aCTC and proposed a cut-off for aCTC of 4. As shown in figure 5.4, the Kaplan-Meier plots before and after one cycle of treatment using aCTC and mCTC are equivalent. The aCTC classifier that was chosen detected up to one object in the control samples as can be seen in figure 5.3. From this figure it becomes clear that a CD45 exclusion criterion is necessary to suppress the number of background objects, although the impact on the HR is small. The influence of the DAPI exclusion criterion is large on HR: only cells with sufficient DNA should be included.

Although one would expect that counting tumour related events that occur at a higher frequency -such as TMPs [12]- is more sensitive and robust, the relationship with clinical outcome was less strong. The aCTC definition was stricter as compared to the mCTC definition as is exemplified by the frequency differences. R^2 between TMP and aCTC was 0.56 and TMPs were also present in the control group (see figure 5.3 bottom row). This may suggest that the current definition of TMPs is a proxy for the number of viable CTC, but in addition enumerates objects unrelated to tumor metastasis (e.g. originating from cell death in tumor or healthy tissue). Higher numbers of events are needed to improve robustness. TMPs may provide these higher numbers, but additional markers are needed to suppress the background signal in healthy volunteers.

The aCTC definition was validated using an independent data set. This validation set showed that the classifier performs well with equivalent HRs to those obtained with mCTC. Correlation with mCTC was quite low (R^2 0.28) due to one outlier. For this outlier, the algorithm counted 1258 aCTC, whereas the operator only counted 67 mCTC. Closer inspection of this sample revealed that this sample had a very high density of cells. This resulted in an overestimation of the number of CTC by the algorithm and an underestimation of the number of CTC by the human operator. Kaplan-Meier plots of baseline and follow-up using aCTC and mCTC from the validation set illustrated in figure 5.4 strongly support the use of the aCTC definition for routine clinical use. Whether or not the same definition for an aCTC can be used for other cancers remains to be determined and is currently being investigated in a large number of samples from breast and colorectal cancer patients. The definition of aCTC in this study was optimized towards the clinical outcome of patients and developed using stored images taken with a $10\times/0.45\text{NA}$ objective. The imaged objects were selected immuno-magnetically targeting the EpCAM antigen and stained with DAPI, CD45-APC, cytokeratin 8,18 and 19-PE. Alteration of the microscope or reagents used to identify CTC will obliterate the aCTC definition. To use this approach for other CTC capturing methods clinical studies will need to be conducted and images stored for relating the particular CTC definition to clinical outcome.

The CellSearch system is the first and currently the only clinically validated method for CTC enumeration. The system was introduced in 2004 and its initial users were well-trained clinical researchers. The need

for CTC counts in the clinic to manage patients with metastatic disease is however rising quickly and is accompanied by a need for simplification, higher reproducibility and a reduction of time needed to obtain a result, i.e. cost reduction. The introduction of aCTC addresses these issues as the need for extensively trained reviewers is eliminated, the algorithm is perfectly reproducible and no operator time is needed to review the images.

In conclusion, we have identified and validated a definition for CTC using an unbiased, automated algorithm that confirms that CK+DAPI+CD45-cells are the EpCAM positive events most strongly associated with survival. Moreover, automated counting of CTC using our classifier compares favourably to manual counting using the CellSearch system.

5.5 REFERENCES

- [1] E. Racila, D. Euhus, A. J. Weiss, C. Rao, J. McConnell, L. Terstappen, and J. W. Uhr, "Detection and characterization of carcinoma cells in the blood," *Proceedings Of The National Academy Of Sciences Of The United States Of America*, vol. 95, no. 8, pp. 4589–4594, 1998.
- [2] W. J. Allard, J. Matera, M. C. Miller, M. Repollet, M. C. Connelly, C. Rao, A. G. J. Tibbe, J. W. Uhr, and L. Terstappen, "Tumor cells circulate in the peripheral blood of all major carcinomas but not in healthy subjects or patients with nonmalignant diseases," *Clinical Cancer Research*, vol. 10, no. 20, pp. 6897–6904, 2004.
- [3] M. Cristofanilli, G. T. Budd, M. J. Ellis, A. Stopeck, J. Matera, M. C. Miller, J. M. Reuben, G. V. Doyle, W. J. Allard, L. Terstappen, and D. F. Hayes, "Circulating tumor cells, disease progression, and survival in metastatic breast cancer," *New England Journal of Medicine*, vol. 351, no. 8, pp. 781–791, 2004.
- [4] G. T. Budd, M. Cristofanilli, M. J. Ellis, A. Stopeck, E. Borden, M. C. Miller, J. Matera, M. Repollet, G. V. Doyle, L. Terstappen, and D. F. Hayes, "Circulating tumor cells versus imaging - predicting overall survival in metastatic breast cancer," *Clinical Cancer Research*, vol. 12, no. 21, pp. 6403–6409, 2006.
- [5] D. Olmos, H. T. Arkenau, J. E. Ang, I. Ledaki, G. Attard, C. P. Carden, A. H. M. Reid, R. A'Hern, P. C. Fong, N. B. Oomen, R. Molife, D. Dearnaley, C. Parker, L. Terstappen, and J. S. de Bono, "Circulating tumour cell (ctc) counts as intermediate end points in castration-resistant prostate cancer (crpc): a single-centre experience," *Annals of Oncology*, vol. 20, no. 1, pp. 27–33, 2009.
- [6] D. F. Hayes, M. Cristofanilli, G. T. Budd, M. J. Ellis, A. Stopeck, M. C. Miller, J. Matera, W. J. Allard, G. V. Doyle, and L. Terstappen, "Circulating tumor cells at each follow-up time point during therapy of metastatic breast cancer patients predict progression-free and overall survival," *Clinical Cancer Research*, vol. 12, no. 14, pp. 4218–4224, 2006.
- [7] S. J. Cohen, C. J. A. Punt, N. Iannotti, B. H. Saidman, K. D. Sabbath, N. Y. Gabrail, J. Picus, M. Morse, E. Mitchell, M. C. Miller, G. V. Doyle, H. Tissing, L. Terstappen, and N. J. Meropol, "Relationship of circulating tumor cells to tumor response, progression-free survival, and overall survival in patients with metastatic colorectal cancer," *Journal of Clinical Oncology*, vol. 26, no. 19, pp. 3213–3221, 2008.
- [8] J. S. de Bono and A. Ashworth, "Translating cancer research into targeted therapeutics," *Nature*, vol. 467, no. 7315, pp. 543–549, 2010.

- [9] M. Kagan, D. Howard, T. Bendele, J. Mayes, J. Silvia, M. Repollet, J. Doyle, J. Al-lard, N. Tu, T. Bui, T. Russell, C. Rao, M. Hermann, H. Rutner, and L. Terstappen, "A sample preparation and analysis system for identification of circulating tumor cells," *Journal of Clinical Ligand Assay*, vol. 25, no. 1, pp. 104–110, 2002.
- [10] J. Kraan, S. Sleijfer, M. H. Strijbos, M. Ignatiadis, D. Peeters, J. Y. Pierga, F. Farace, S. Riethdorf, T. Fehm, L. Zorzino, A. G. Tibbe, M. Maestro, R. Gisbert-Criado, G. Denton, J. S. de Bono, C. Dive, J. A. Foekens, and J. W. Gratama, "External quality assurance of circulating tumor cell enumeration using the cellsearch(r) system: A feasibility study," *Cytometry B Clin Cytom*, vol. 80B, no. 2, pp. 112–118, 2010.
- [11] A. G. J. Tibbe, M. C. Miller, and L. Terstappen, "Statistical considerations for enumeration of circulating tumor cells," *Cytometry Part A*, vol. 71A, no. 3, pp. 154–162, 2007.
- [12] F. A. W. Coumans, C. J. M. Doggen, G. Attard, J. S. de Bono, and L. W. M. M. Terstappen, "All circulating epcam+ck+cd45- objects predict overall survival in castration-resistant prostate cancer," *Annals of Oncology*, vol. 21, no. 9, pp. 1851–7, 2010.
- [13] G. Attard and J. S. de Bono, "Utilizing circulating tumor cells: challenges and pitfalls," *Current Opinion in Genetics & Development*, vol. 21, no. 1, pp. 50–58, 2011.
- [14] J. S. de Bono, H. I. Scher, R. B. Montgomery, C. Parker, M. C. Miller, H. Tissing, G. V. Doyle, L. Terstappen, K. J. Pienta, and D. Raghavan, "Circulating tumor cells predict survival benefit from treatment in metastatic castration-resistant prostate cancer," *Clinical Cancer Research*, vol. 14, no. 19, pp. 6302–6309, 2008.
- [15] G. Attard, A. H. M. Reid, T. A. Yap, F. Raynaud, M. Dowsett, S. Settatree, M. Barrett, C. Parker, V. Martins, E. Folkard, J. Clark, C. S. Cooper, S. B. Kaye, D. Dearnaley, G. Lee, and J. S. de Bono, "Phase i clinical trial of a selective inhibitor of cyp17, abiraterone acetate, confirms that castration-resistant prostate cancer commonly remains hormone driven," *Journal of Clinical Oncology*, vol. 26, no. 28, pp. 4563–4571, 2008.
- [16] G. Attard, A. H. M. Reid, R. A'Hern, C. Parker, N. B. Oommen, E. Folkard, C. Messiou, L. R. Molife, G. Maier, E. Thompson, D. Olmos, R. Sinha, G. Lee, M. Dowsett, S. B. Kaye, D. Dearnaley, T. Kheoh, A. Molina, and J. S. de Bono, "Selective inhibition of cyp17 with abiraterone acetate is highly active in the treatment of castration-resistant prostate cancer," *Journal of Clinical Oncology*, vol. 27, no. 23, pp. 3742–3748, 2009.
- [17] A. H. M. Reid, G. Attard, D. C. Danila, N. B. Oommen, D. Olmos, P. C. Fong, L. R. Molife, J. Hunt, C. Messiou, C. Parker, D. Dearnaley, J. F. Swennenhuis, L. Terstappen, G. Lee, T. Kheoh, A. Molina, C. J. Ryan, E. Small, H. I. Scher, and J. S. de Bono, "Significant and sustained antitumor activity in post-docetaxel, castration-resistant prostate cancer with the cyp17 inhibitor abiraterone acetate," *Journal of Clinical Oncology*, vol. 28, no. 9, pp. 1489–1495, 2010.
- [18] G. J. Bubley, M. Carducci, W. Dahut, N. Dawson, D. Daliani, M. Eisenberger, W. D. Figg, B. Freidlin, S. Halabi, G. Hudes, M. Hussain, R. Kaplan, C. Myers, W. Oh, D. P. Petrylak, E. Reed, B. Roth, O. Sartor, H. Scher, J. Simons, V. Sinibaldi, E. J. Small, M. R. Smith, D. L. Trump, R. Vollmer, and G. Wilding, "Eligibility and response guidelines for phase ii clinical trials in androgen-independent prostate cancer: Recommendations from the prostate-specific antigen working group," *Journal of Clinical Oncology*, vol. 17, no. 11, pp. 3461–3467, 1999.

- [19] G. W. Zack, W. E. Rogers, and S. A. Latt, "Automatic-measurement of sister chromatid exchange frequency," *Journal Of Histochemistry & Cytochemistry*, vol. 25, no. 7, pp. 741–753, 1977.
- [20] L. Breiman, "Bagging predictors," *Machine Learning*, vol. 24, no. 2, pp. 123–140, 1996.
- [21] H. I. Scher, X. Y. Jia, J. S. de Bono, M. Fleisher, K. J. Pienta, D. Raghavan, and G. Heller, "Circulating tumour cells as prognostic markers in progressive, castration-resistant prostate cancer: a reanalysis of immc38 trial data," *Lancet Oncology*, vol. 10, no. 3, pp. 233–239, 2009.

MORPHOLOGY OF CIRCULATING TUMOR CELLS AND SURVIVAL IN BREAST, COLORECTAL, AND PROSTATE CANCER¹

Sjoerd T. Ligthart, Frank A.W. Coumans, Francois-Clément Bidard, Lieke H.J. Simkens, Cees J.A. Punt, Marco de Groot, Gerhardt Attard, Johann S. de Bono, Jean-Yves Pierga, and Leon W.M.M. Terstappen

ABSTRACT

Background: Presence of circulating tumor cells (CTC) in patients with metastatic breast, colorectal, and prostate cancer is indicative for poor progression free and overall survival. An automated CTC (aCTC) algorithm was developed to eliminate the variability associated with manual counting of CTC (mCTC). This algorithm permits extraction of morphological data. Here we validated the aCTC algorithm, which was trained using survival data from prostate cancer patients, on CTC from breast and colorectal cancer patients and investigated the role of quantitative morphological parameters. **Methodology:** Stored images of samples from patients with benign and malignant breast, colorectal, and prostate cancer, and healthy controls were obtained using the CellSearch system. Images were re-analyzed, aCTC identified, and their morphological parameters measured and correlated with survival. **Results:** Correlation R^2 between aCTC and mCTC was 0.48 for 725 breast, 0.84 for 1729 colorectal, and 0.68 for 559 prostate cancer. However, no significant difference in the relation between survival and aCTC or mCTC

¹Submitted.

was found. Counts in 204 control samples were ≤ 1 aCTC. Size of aCTC and nuclear-to-cytokeratin ratio had significant ($p < 0.05$) impact on survival for breast cancer patients only. Conclusions: Morphological parameters of CTC differ between different cancer types. Although correlations between manual and automated count are low, an algorithm can be used to identify CTC of all cancer types with high prognostic value.

6.1 INTRODUCTION

Case studies showing the presence of tumor cells in blood of cancer patients have been reported for more than a century [1, 2, 3, 4]. These circulating tumor cells (CTC) are extremely rare and technology to reliably enumerate CTC has only become available in recent years [5]. Studies conducted with this validated system established the relation between the presence of CTC and poor outcome [6, 7, 8, 9, 10]. The ability to assess the presence of treatment targets on CTC demonstrates the potential for a real time liquid biopsy and has certainly spurred the interest in CTC [11, 12, 13, 14, 15, 16].

The recent emergence of different technologies to identify CTC is accompanied with a large range of CTC definitions urging the need for standardization [5, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30]. Although different CTC characteristics may indeed relate to clinical outcome, it is of utmost importance that counting and characterization of CTC can be done accurately, reproducibly, and is validated in controlled clinical studies [31]. Studies have shown that even in the FDA cleared CellSearch system intra-reviewer, inter-reviewer, and inter-laboratory variability is substantial [32, 33].

Recently, we developed a CTC detection algorithm that counts CTC in images recorded by the CellSearch system [34]. This algorithm was not developed to copy human reviewers that assign events as CTC, but it used survival data of metastatic prostate cancer patients to arrive at a definition that optimally stratified the patients into groups with favorable and unfavorable survival. This algorithm eliminates reviewer variability, is fast and decreases the cost of the CTC assay by elimination of the time a reviewer spends on reviewing the images. Furthermore, it delivers quantitative information about the objects it counts as CTC.

From case studies in breast and colorectal cancer it was found that breast cancer CTC are somewhat rounder than cells from colorectal samples [35, 36]. However, all studies show that CTC are morphologically very heterogeneous [5]. In this study, we validated the CTC algorithm on images of patients with metastatic colorectal and breast cancer and evaluated the correlation of quantitative morphological parameters with clinical outcome of patients with metastatic prostate, colorectal and breast cancer before and after initiation of a new line of therapy.

TABLE 6.1 : Patient characteristics.

Study	IMMC-01	IC2006-04	CAIRO-2	IMMC-06	IMMC-38	Abiraterone
Type	Breast	Breast	Colorectal	Colorectal	Prostate	Prostate
N Patients	179	248	417	34	185	100
N Samples	283	442	1691	40	370	189
N Censored	75	193	180	10	67	26
Survival (months) ^a	15.1 (0.4–48.6)	13.1 (0.2–33.3)	18.7 (0–60.8)	23.4 (0.8–39.1)	16.9 (1.9–38.7)	22.2 (0.9–53.2)
Follow-up Censored ^b (months)	20.7 (1.3–48.6)	14.2 (1.4–33.3)	36.7 (0–60.8)	31.8 (22.8–39.1)	25.9 (2.4–38.7)	40.8 (9.2–53.2)
Therapy line ^c	43/15/42%	100/0/0%	100/0/0%	62/29/9	67/16/17%	0/20/80%
Age ^a	58 (27–86)	57 (28–84)	63 (27–84)	65 (45–83)	69 (49–92)	70 (50–84)

^a Median values shown with ranges in parentheses, ^b median duration of follow-up for censored patients with ranges in parentheses, ^c % of patients with 1st, 2nd, 3+ line of chemotherapy (% 1st/ % 2nd/ % 3+).

6.2 MATERIALS AND METHODS

6.2.1 PATIENTS

Samples of patients before they received treatment (baseline) and at first follow-up after initiation of a new line of therapy from prospective multi-center CTC studies were used in this study. For metastatic breast cancer, fluorescence CellSearch images of 283 samples from 179 patients could be imported from the IMMC-01 study and 482 samples from 248 patients from the IC2006-04 study [7, 37]. For metastatic colorectal cancer, CellSearch images of 1691 samples from 507 patients could be imported from the CAIRO-2 and IMMC-06 studies [6, 38]. For castration resistant prostate cancer (CRPC), images of 370 samples from 185 patients could be imported from the IMMC-38 study and 189 samples from 100 patients could be imported from the Abiraterone study [8, 11, 39]. From the IMMC-26 study, images of 317 samples from 93 patients with benign colorectal disease and 200 samples from 61 patients with benign breast cancer disease were imported [40, 41, 42]. In addition, 204 samples from healthy controls from the IMMC-01 and IMMC-06 study were included in this study. Table 6.1 shows detailed characteristics of the patients enrolled in these studies. All patients and healthy volunteers provided written informed consent. The institutional review boards at each participating center approved the study protocol.

6.2.2 MANUAL ENUMERATION OF CIRCULATING TUMOR CELLS

The CellSearch system (Veridex LLC, Raritan, NJ) was used to enumerate CTC. The system consists of a CellTracks Autoprep[®] for sample preparation and a CellTracks Analyzer II[®] for sample analysis [5, 43]. The CellTracks Autoprep immuno-magnetically enriches epithelial cells from 7.5 ml of blood targeting the Epithelial Cell Adhesion Molecule (EpCAM). The enriched sample is labeled with phycoerythrin-conjugated (PE) antibodies directed against cytokeratins (CK) 8, 18 and 19, an allophycocyanin-conjugated (APC) antibody against CD45 and the nuclear dye 4',6-diamidino-2-phenylindole (DAPI). After preparation the enriched sample is transferred to a cartridge contained in a CellTracks MagNest. All ferrofluid labeled objects are pulled towards the imaging surface of the cartridge, which is placed on the CellTracks Analyzer II for image acquisition. The analyzer is a four-color semi-automated fluorescence microscope that captures digital images in four different fluorescent channels using a $10\times/0.45$ NA objective and a charge-coupled device camera with 6.7×6.7 μm sized pixels. For each cartridge, 144–180 4-layer tiff images of DAPI, FITC, PE, APC are saved per patient. After imaging, the system preselects objects that are positive for DNA and CK and shows the images of all fluorescence channels together with a DNA/CK overlay of all selected objects in a gallery. Finally, a trained operator selects objects from the gallery as manual CTC (mCTC) if they are DNA+CK+, CD45-, have a cell-like morphology, and are larger than 4×4 μm .

6.2.3 AUTOMATED ENUMERATION OF CIRCULATING TUMOR CELLS

The recorded images from the patient samples were copied to central hard drives and analyzed by an algorithm developed in Matlab 2009a, using the DIPimage toolbox (Delft University of Technology, Delft, The Netherlands). After loading the 144–180 images per patient, objects in the CK-PE channel are selected by dynamic thresholding and their outline is saved [44]. Next, the outline of each object is used to perform measurements in the DNA-DAPI, CK-PE, and CD45-APC channels. Finally, an algorithm, optimized using survival data from metastatic prostate cancer patients, selects automated CTC (aCTC). Details of the algorithm are described in more detail elsewhere [34]. The original inclusion criteria for an aCTC were a CK-PE standard deviation higher than 50 counts, a DNA-DAPI peak value of at least 170 counts, a CD45-APC peak value less than 60 counts and a size within 75–500 pixels (35 – 225 μm^2). However, the upper aCTC size limit was increased to 2000 pixels (~ 900 μm^2) after the observation that aCTC sizes in metastatic breast cancer are larger [45]. For every patient sample, the included objects were summed to arrive at a final aCTC count per patient. To compare the overall aCTC and mCTC enumeration, a frequency plot was created using the data from the three patient groups,

benign tumors and healthy controls.

6.2.4 ADDITIONAL MORPHOLOGICAL MEASUREMENTS ON aCTC

We performed a number of measurements on the aCTC, including cell roundness, nuclear to cytokeratin (NCK) ratio, the presence of cell clusters, and the number of speckles within the cytoplasm. Cell roundness was defined as the ratio of the signal perimeter squared to the signal area times 4π (termed P2A in the literature). Thus, roundness equals one for a perfect circle. The NCK-ratio was defined as the area of the DNA-DAPI signal, divided by the area of the CK-PE signal. For this parameter, an extra segmentation step for DNA-DAPI was performed to find DNA-DAPI+ objects larger than $10 \mu\text{m}^2$ (22 pixels, the smallest DNA-DAPI nucleus area). If multiple DNA-DAPI+ objects were found close to the CK signal, the center-of-mass distances between the CK and DNA signals were determined. The object with (partial) overlap between CK and DNA and the smallest distance (no more than $17 \mu\text{m}$, or half the size limit as set by the algorithm) was selected to determine the NCK-ratio. Clusters of cells were defined as the number of DNA-DAPI objects larger than $10 \mu\text{m}^2$ and located within each CK-PE signal. Finally, speckles within the CK-PE signals are indicative of damaged or apoptotic CTC [46]. The number of dot-like structures within the CK-PE signal was counted and used as a measure of apoptosis. Dot-like structures were counted by an algorithm developed for the detection of fluorescence in situ hybridization probes within a cell nucleus [47].

6.2.5 IDENTIFICATION AND MORPHOLOGICAL MEASUREMENTS OF LEUKOCYTES

Due to a-specific binding of EpCAM-ferrofluid particles during the CTC enrichment, leukocytes are carried over and are present at the analysis surface of the sample cartridge. Leukocytes are identified by their DAPI staining, expression of CD45 and lack of cytokeratin expression. The outline of the leukocytes was determined in the CD45-APC channel, using the triangle threshold method [44]. Leukocytes were defined as objects that had CK-PE peak value lower than 100 counts, DNA-DAPI peak value of at least 100 counts, CD45-APC peak value of at least 100 counts and a size range between 50 and 2000 pixels ($21\text{-}900 \mu\text{m}^2$). Morphological parameters were extracted as described for aCTC.

6.2.6 STATISTICAL ANALYSIS

SPSS 16 (IBM, New York, USA) was used for survival analysis. Least-squares linear regression analysis was performed to compare aCTC and mCTC, baseline and follow-up samples, and to compare different studies of the same type of cancer. Kaplan-Meier analysis was performed by

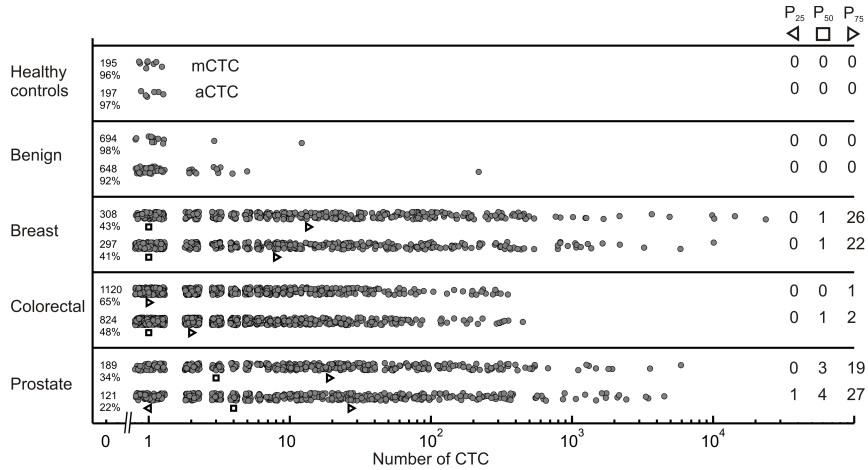


FIGURE 6.1 : Scatter plot of frequencies of mCTC and aCTC found in healthy controls, patients with benign disease and patients with metastatic breast, colorectal, and prostate cancer before initiation of a new line of therapy. The dots in the scatter plot are spread for viewing purposes. The 25th, 50th, and 75th percentiles are given by the black markers and in numerical values in the right part of the figure. At the start of each row, the percentage and number of patients that have 0 CTC are shown.

dichotomizing the patient population on the clinically used cut-off value of five CTC for breast and prostate cancer, and three CTC for colorectal cancer. Next to these cut-offs, other cut-off values of 1, 10, and 100 for aCTC or mCTC were tested to dichotomize the patient population per study for univariate Cox hazard regression. Log-rank tests were used to compare survival between groups. Cox hazard ratios (HR) with 95% confidence intervals (CI) were determined and collected in a table. Survival was defined as the time between baseline sample and death from any cause.

To determine whether the morphological parameters determined on aCTC provide additional information that is relevant for prognosis, multivariate Cox proportional hazards regression was performed (forward conditional, with p_{in} 0.10 and p_{out} 0.20), controlling for cancer type (colorectal, prostate, breast), with all morphological parameters and the aCTC count as continuous variables. The number of aCTC was $^{10}\log$ transformed before inclusion in the regression [48, 49]. Because the morphological parameters could not be determined on patients without aCTC, only patients with at least 1 aCTC were included. For parameters that are determined for each individual aCTC, the median value within a patient was used since this represents the predominant characteristic.

The nonparametric Mann-Whitney U-test was used to compare morphological parameters of different types of cancer. The t-test could not be used because the parameters are not normally distributed.

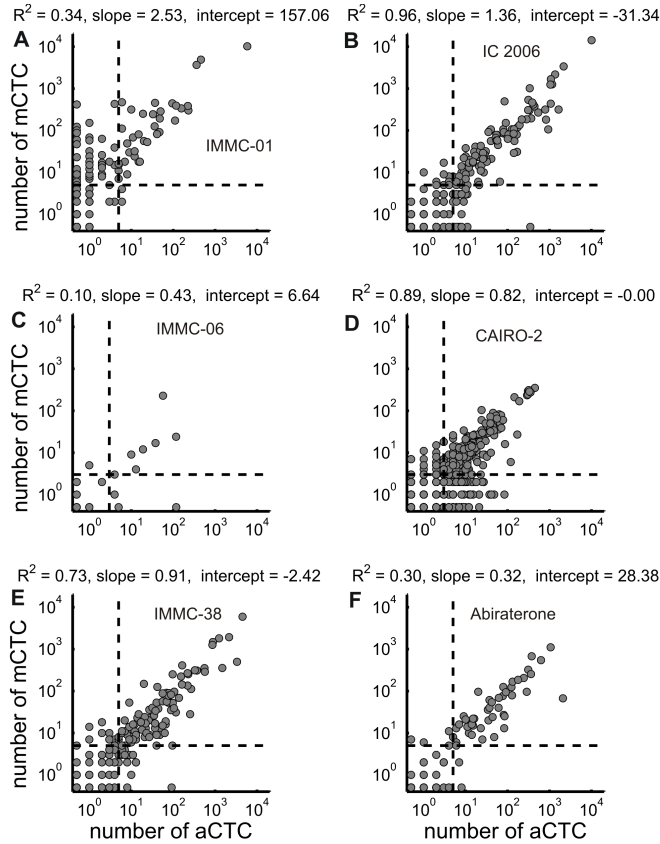


FIGURE 6.2 : Scatter plot of the number of aCTC versus mCTC for every study. Horizontal and vertical dashed lines were drawn to show clinically used cut-offs, which are ≥ 5 for breast and prostate, and ≥ 3 for colorectal cancer. At the top of each graph, the linear regression coefficients are shown.

6.3 RESULTS

6.3.1 FREQUENCY OF ACTC VERSUS MCTC

aCTC and mCTC were enumerated in healthy controls, patients with benign disease and patients with metastatic colorectal, prostate and breast cancer before initiation of therapy. The frequency distribution of the CTC is shown in figure 6.1 and the 25th, 50th, and 75th percentiles are provided at the right hand side of the figure. Figure 6.2 shows the scatter plot of aCTC versus mCTC per study, as well as the linear regression coefficients.

In summary, breast CTC counts ranged from 0–10111 aCTC (median 2/0 for baseline/follow-up) and 0–23618 mCTC (median 3/0 for baseline/follow-up); colorectal aCTC counts from 0–448 (median 1/0) and 0–351 mCTC

TABLE 6.2 : Measurements on aCTC and leukocytes for baseline and follow-up samples. Shown are mean and standard deviation.

	Time point		Breast	Colorectal	Prostate	Leukocytes
Size (μm^2)	BL ^a	μ^c	290	186	180	71
		σ^d	200	153	145	44
	FU1 ^b		220	209	174	66
			142	192	141	47
NCK-ratio ^e	BL		0.8	1.0	1.0	N/A ^g
			2.8	3.5	2.1	
	FU1		0.5	1.1	1.3	N/A
			2.0	4.9	4.3	
Roundness ^g	BL		1.5	1.5	1.5	1.8
			0.6	1.4	0.9	1.2
	FU1		1.1	1.3	1.3	1.4
			0.7	1.6	0.9	1.1

^a Baseline measurement, ^b first follow-up, ^c mean, ^d standard deviation, ^e nuclear-cytokeratin ratio, 1 indicates area of cytokeratin equal to area of nucleus, ^f not applicable, ^g 1 indicates perfectly round.

(median 0/0); prostate CTC counts ranged from 0–4970 aCTC (median 7/3) and 0–5925 mCTC (median 7/2). Benign patients ranged from 0–218 (median 0) aCTC and 0–12 mCTC (median 0). Finally, healthy controls had up to 1 aCTC and mCTC (both medians 0). Correlation R^2 between mCTC and aCTC was 0.48 for 725 breast samples, 0.84 for 1729 colorectal samples, and 0.68 for 559 prostate samples.

6.3.2 MEASUREMENTS OF MORPHOLOGICAL PARAMETERS OF ACTC

Recording of the outline of aCTC permits the extraction and quantification of morphological parameters from the identified aCTC. The distribution of the size, NCK-ratio and roundness of aCTC in blood of patients with metastatic breast, colorectal, and prostate cancer before initiation of therapy and at first follow up after initiation of therapy is shown in figure 6.3. The aCTC size distribution is given in μm^2 and is illustrated above panels A, B and C that show the results for colorectal, breast and prostate cancer respectively. The NCK-ratio of the aCTC ranged from 0.5 to 5, examples that show the range of these values are illustrated above panels E, F and G, for breast, colorectal, and prostate cancer respectively. aCTC appear in a large range of shapes. To illustrate this diversity in CTC shape, a measure of roundness ranging from 1 to 5 is shown in panels H, I, and J, and illustrated above these panels, for colorectal, breast and prostate cancer, respectively.

Below the aCTC measures, the size and roundness properties of leukocytes are shown in figure 6.3, panels D and K. The NCK-ratio for leukocytes is not shown as leukocytes do not express cytokeratin. R^2 correlation be-

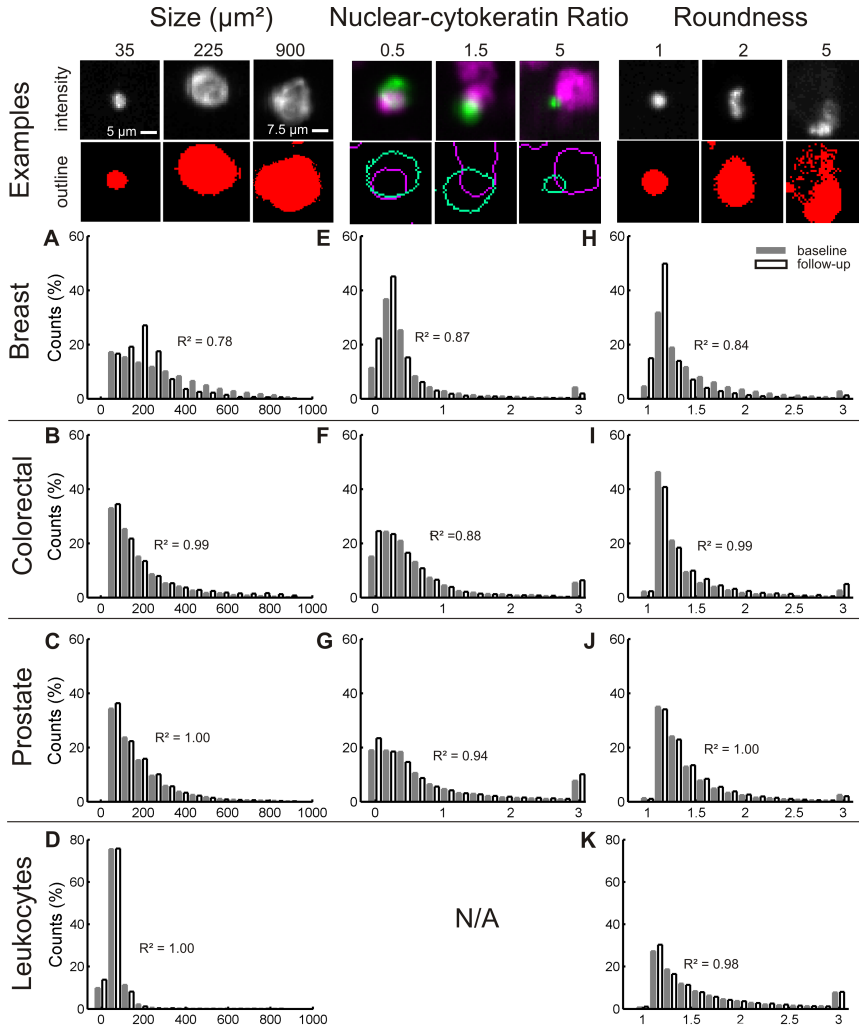


FIGURE 6.3 : Histograms of morphological parameters of aCTC for breast (panel A, E, H), colorectal (panel B, F, I) and prostate samples (panel C, G, J), next to morphological parameters of leukocytes (panel D, K). Baseline samples (black bars) are shown next to first follow-up samples (white bars). Coefficients of determination of a linear regression between the time points are shown next to the bars. Examples of morphological parameter values of aCTC are given at the top of the figure, as well as examples of the outlines after segmentation. The objects seem smaller in the images than in the outlines. This is due to scaling of the image for print, the outlines show the area of the object that is distinguishable from noise, and thus represents the true size of the object. The 5 μm scale bar applies to all images if not stated differently. N/A: not applicable.

TABLE 6.3 : Cluster and speckled aCTC measurements on baseline and follow-up samples that had ≥ 5 aCTC from breast, colorectal and prostate cancer patients. aCTC were considered clustered if they had more than one DNA-DAPI object within their CK-PE outline, and speckled if they had more than two dot-like structures in their CK-PE outline.

Patients ≥ 5 aCTC	Clusters - mean % (median %)		Speckled - mean % (median %)	
	BL ^a	FU1 ^b	BL	FU1
Breast (N=111)	3.8 (2.2)	3.2 (0.0)	3.5 (2.1)	5.5 (2.9)
Colorectal (N=100)	5.2 (0.0)	8.6 (0.0)	3.2 (0.0)	6.0 (0.0)
Prostate (N=49)	5.1 (2.9)	4.7 (3.5)	3.4 (2.1)	3.3 (0.9)
Leukocytes (N=20)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)	0.0 (0.0)

^a Baseline measurement, ^b first follow-up measurement.

tween measurements before and after initiation of therapy is provided in each panel. aCTC morphological parameters were quantified by mean and standard deviation and are provided in table 6.2. For inter-study comparison, morphological parameters from the respective studies were compared and plotted in figure 6.4 and their R^2 correlation is given in every graph.

To test whether the morphological parameters from table 6.2 were significantly different between cancer types, Mann-Whitney U-test p-values were derived using all samples of each type of cancer. Derived p-values for median size of aCTC between colorectal and prostate was 0.90 and 0.35 for median NCK-ratio between breast and colorectal. The percentage apoptotic cells between breast and prostate had $p=0.18$. For median roundness of aCTC p-values were 0.27, 0.008, and 0.15 for comparison between breast-colorectal, breast-prostate, and colorectal-prostate, respectively. For all other morphological parameters from table 6.2 p was <0.001 between all types of cancer.

The size of the objects can be influenced by the presence of multiple CTC representing CTC-clusters. These clusters are counted as one CTC in both the manual identification as well as the automated identification of CTC. The presence of aCTC clusters was estimated by the number of nuclei that could be detected within the CK-PE area. Percentages of clustered cells per patient are shown in table 6.3 for patients with ≥ 5 aCTC for both baseline and follow-up samples. None of the patients showed a significant rise or drop (outside 95% CI) in the number of clusters after initiation of therapy.

The presence of speckles of cytokeratin is associated with CTC undergoing apoptosis. A measure of the cytokeratin distribution was determined to capture this phenomenon and the percentages of cells having 3 or more dot-like structures in their cytokeratin signal is shown in table 6.3 for both baseline and follow-up samples. None of the patients had a significant rise or drop in the number of speckled objects between baseline and first follow-up measurement.

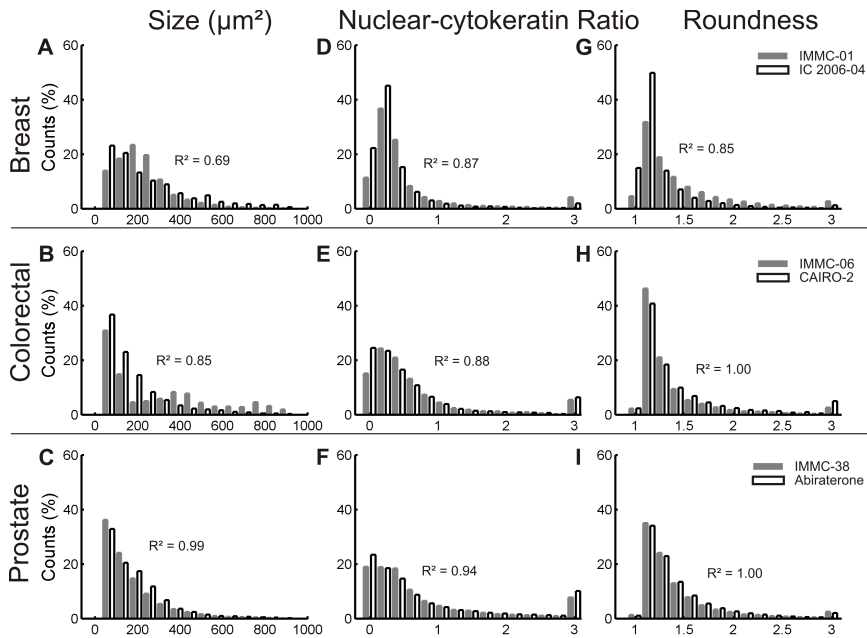


FIGURE 6.4 : Histograms of morphological parameters of aCTC for breast (panel A, D, G), colorectal (panel B, E, H) and prostate samples (panel C, F, I) per study. Coefficient of determination of a linear regression between studies is shown next to the bars.

6.3.3 MCTC, ACTC AND MORPHOLOGICAL PARAMETERS VERSUS SURVIVAL IN METASTATIC BREAST, COLORECTAL AND PROSTATE CANCER

To illustrate the relationship between survival and CTC, Kaplan-Meier plots were generated dividing the patients in favorable and unfavorable groups based on a threshold of 3 CTC for metastatic colorectal cancer and 5 CTC for metastatic prostate and breast cancer. The Kaplan-Meier plots shown in figure 6.5 are for aCTC and mCTC before initiation of therapy (panels A, C, E) and at first follow-up after initiation of therapy (panels B, C, F). The relationship between aCTC number and overall survival in metastatic breast, colorectal, and prostate cancer is shown in table 6.4. All p-values were <0.001 , except for breast cancer samples using a cut-off of ≥ 1 for baseline samples ($p=0.018$) and follow-up samples ($p=0.003$). To investigate the impact of aCTC morphological parameters on survival, univariate and multivariate analysis was performed using the measured morphological parameters as covariates next to the number of aCTC or mCTC. HRs were calculated for each covariate as a continuous variable. Table 6.5A shows the HRs, derived p-values, and Chi-squared.

TABLE 6.4 : Correlation of CTC numbers with overall survival in metastatic colorectal, breast, and prostate cancer

cut-off	aCTC			mCTC		
	n at risk	% at risk	HR ^a (95% CI ^b)	n at risk	% at risk	HR (95% CI)
Breast baseline measurement						
≥1	270	66	1.5 (1.1–2.2)	279	68	2.8 (1.8–4.2)
≥5	160	39	2.3 (1.7–3.3)	194	47	3.0 (2.1–4.2)
≥10	120	29	2.5 (1.8–3.5)	151	37	3.1 (2.2–4.3)
≥100	43	10	2.3 (1.4–3.6)	54	13	2.7 (1.8–4.1)
Breast first follow-up measurement						
≥1	151	48	1.8 (1.2–2.6)	139	44	3.3 (2.2–5.0)
≥5	56	18	3.7 (2.5–5.7)	74	24	4.2 (2.9–6.3)
≥10	40	13	3.7 (2.3–5.9)	55	18	3.9 (2.6–5.9)
≥100	6	2	8.3 (3.0–23.5)	14	4	4.1 (2.1–7.7)
Colorectal baseline measurement						
≥1	565	64	1.7 (1.4–2.0)	437	49	1.8 (1.5–2.1)
≥3	322	36	1.7 (1.5–2.0)	249	28	1.9 (1.6–2.2)
≥10	139	16	1.9 (1.6–2.3)	114	13	2.3 (1.8–2.8)
≥100	13	1	3.9 (2.2–6.9)	11	1	6.1 (3.3–11.3)
Colorectal first follow-up measurement						
≥1	340	40	1.5 (1.2–1.7)	172	20	2.1 (1.8–2.6)
≥3	108	13	2.0 (1.6–2.5)	70	8	3.8 (2.7–4.7)
≥10	27	3	4.6 (3.0–6.9)	19	2	13.3 (8.1–21.7)
≥100	4	0	24.1 (8.5–68.3)	4	0	24.1 (8.5–68.3)
Prostate baseline measurement						
≥1	231	83	2.9 (1.8–4.6)	205	74	2.7 (1.9–3.9)
≥5	156	56	3.4 (2.5–4.8)	153	55	3.1 (2.3–4.3)
≥10	123	44	3.0 (2.2–4.0)	118	42	3.3 (2.4–4.5)
≥100	35	13	3.7 (2.5–5.6)	31	11	5.5 (3.5–8.4)
Prostate first follow-up measurement						
≥1	207	74	2.3 (1.6–3.4)	165	59	3.9 (2.8–5.5)
≥5	113	40	4.1 (3.0–5.5)	108	38	4.2 (3.1–5.8)
≥10	95	34	3.9 (2.8–5.3)	83	30	4.5 (3.3–6.2)
≥100	29	10	5.2 (3.4–8.1)	23	8	4.0 (2.4–6.6)

^a Cox hazard ratio, ^b confidence interval.

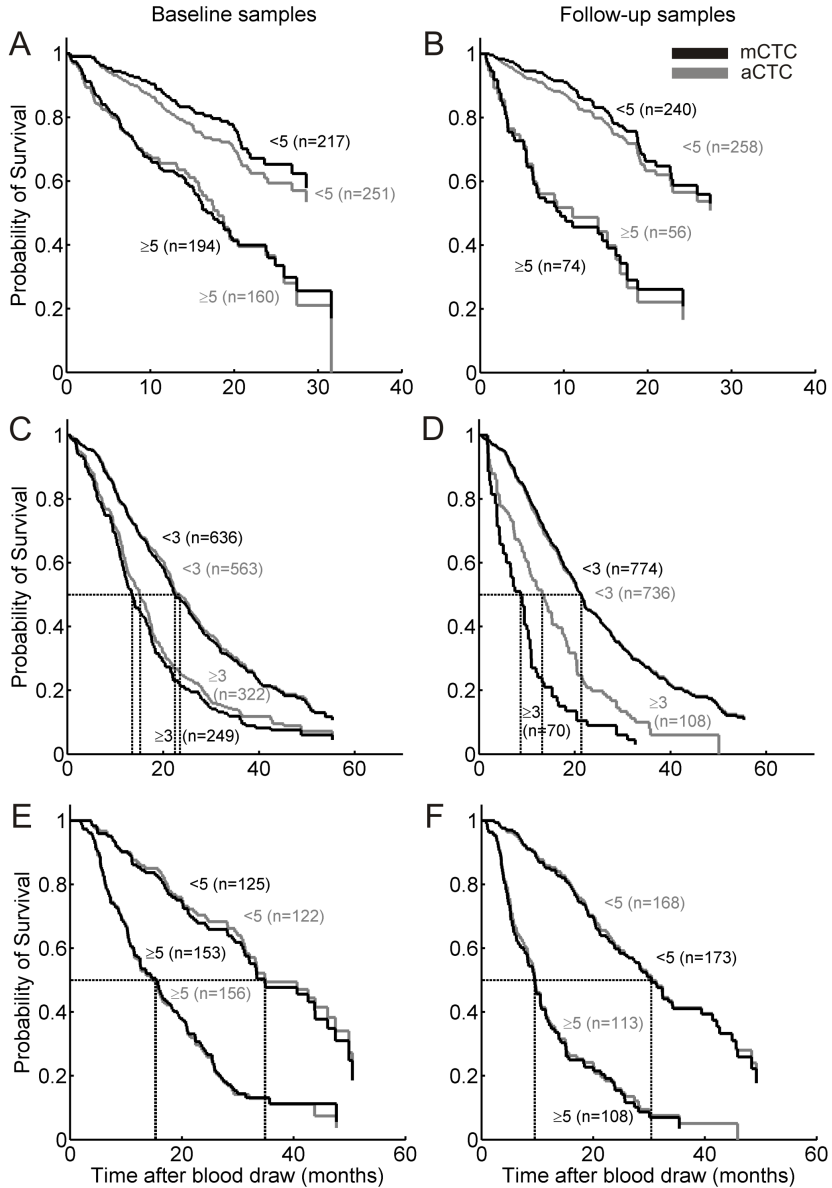


FIGURE 6.5 : Kaplan-Meier plots of baseline and first follow-up samples of the metastatic breast (panel A and B), colorectal (panel C and D), and prostate cancer patients (panel E and F) with favorable and unfavorable mCTC (black lines) and aCTC definition (grey lines). For every type of cancer, patients were pooled from the two studies of each type as shown in table 6.1.

Next, all variables of aCTC were entered in a multivariate Cox regression analysis. mCTC were left out of this analysis, as quantitative parameters are unknown of these cells. Hazard models were calculated using all first follow-up samples for all cancer types together and for each type separately. When all samples were combined, we controlled for cancer type to compensate for general differences in survival. Results are shown in table 6.5B. The optimal model used the number of aCTC, their median roundness per patient and had a Chi-square of 181 with 3 degrees of freedom. The model shows that more objects impose a bigger threat and that objects with a roundness close to one are more dangerous for patients than objects that are for instance elongated. When including only samples from one type of cancer, the median size of the aCTC, and the median NCK-ratio of aCTC had a significant influence on survival in breast cancer samples. HRs smaller than one show that in breast cancer small objects represent greater hazard than large objects. Furthermore, objects with small DAPI signals with respect to their CK signal, i.e. a low NCK-ratio, are more dangerous for patients than cells with high NCK-ratio. Other morphological parameters did not improve the multivariate models.

6.4 DISCUSSION

Recently, we showed that counting CTC by an automated algorithm is preferable to manual counting by a trained reviewer [34]. Our aCTC algorithm performed similar to the human reviewer in terms of prognostic value: similar HRs were found. Additionally, our algorithm performed this task much faster and had 0% variability, against an inter-laboratory variability of 4% to 31% for mCTC (median 14%) reported previously by another group [32]. The aCTC algorithm was trained using samples from prostate cancer patients in the IMMC-38 study and validated on an independent data set of prostate cancer patients that were included in the Abiraterone phase I/II trials [11, 39]. The aCTC algorithm with the highest prognostic value includes objects that contain DNA and CK, rejects CD45 and have a size between $34 \mu\text{m}^2$ and $224 \mu\text{m}^2$.

Here we applied the aCTC algorithm on samples of patients with benign and malignant breast and colorectal cancer, as well as on samples of healthy controls. The upper size limit was increased from $224 \mu\text{m}^2$ to $900 \mu\text{m}^2$, concordant with a recent study showing a larger size of CTC in breast cancer [45]. This change resulted in an aCTC increase by a factor of 1.3 for prostate cancer with less than 15% effect on the HR. The increase in the number of aCTC was a factor 1.4 for colorectal samples and 1.9 for breast cancer samples, and thus indicative for the larger average size of CTC in breast cancer. The frequency distribution of CTC shown in figure 6.1 shows that the aCTC and mCTC counts do not differ greatly. Separation of the results into the different studies however showed a clear discrepancy

TABLE 6.5 : Univariate (panel A) and multivariate (panel B) analysis of variables on follow-up samples. Univariate analysis was performed using all samples, the multivariate analysis on all samples and for each type of cancer separately. The aCTC and mCTC numbers were $10\log$ transformed for improved linearity. Mean and standard deviation of the variable is given for HR interpretation. A $HR < 1$ shows that smaller values of a variable are more dangerous than larger values of that variable. Chi-square is given for comparison of the whole variable or model with the null-hypothesis, degrees of freedom is equal to the number of variables shown.

A: Univariate Analysis				
Variable	mean \pm SD ^a	HR ^b	p	Chi-square
$^{10}\log(\text{aCTC}^c)$	0.7 \pm 0.8	1.7	<0.001	178
$^{10}\log(\text{mCTC}^d)$	0.9 \pm 0.8	1.5	<0.001	203
median size (μm^2)	167 \pm 127	0.98	0.002	9
% speckled	5.9 \pm 17.5	0.996	0.02	6
median NCK-Ratio ^e	0.6 \pm 1.4	0.97	0.3	1
median roundness	1.5 \pm 1.1	0.86	0.001	10
% clusters	6.3 \pm 17.5	0.998	0.1	2
B: Multivariate Analysis				
All first follow-up samples				
$^{10}\log(\text{aCTC})$		2.5	<0.001	180
median roundness		0.9	0.040	
ColorectalControl ^f		1.8	<0.001	
Breast first follow-up samples				
$^{10}\log(\text{aCTC})$		2.9	<0.001	36
median size		0.92	0.018	
median NCK-Ratio		0.5	0.026	
Colorectal first follow-up samples				
$^{10}\log(\text{aCTC})$		2.7	<0.001	50
Prostate first follow-up samples				
$^{10}\log(\text{aCTC})$		2.6	<0.001	42

^a Standard deviation, ^b Cox hazard ratio, ^c automated circulating tumor cells, ^d manual circulating tumor cells, ^e nuclear-to-cytokeratin ratio, ^f 1 if a samples is from a colorectal cancer patient, otherwise 0.

between the counts obtained with samples from the IMMC-01 breast cancer study as compared to all other studies (figure 6.2B). This also resulted in a lower prognostic value of aCTC as shown in figure 6.5A and B, were all the metastatic breast patients were pooled.

The cause of this discrepancy was traced back to the CD45 staining of objects as the CD45 staining in the IMMC-01 study was much higher as compared to the IC2006-04 study. This resulted in the rejection by the algorithm of many objects in the IMMC-01 study. Lowering of the CD45 threshold eliminated the discrepancy between the number of aCTC in the IMMC-01 and other studies (data not shown). Potential causes for the higher CD45 signal are a higher concentration of the CD45-APC fluorochrome, differences in sample preparation or difference in filter cubes giving rise

to more crosstalk of PE signal into the APC channel. Most samples of IMMC-01, 06 and 38 were scanned on the CellSpotter Analyzer instead of the CellTracks Analyzer II used for the other studies, therefore differences between CellTracks and CellSpotter do not offer an explanation for the differences in CD45 signal. The IMMC-01 study was the first large clinical trial run with the CellSearch system, and during this study samples were not automatically processed by the CellSearch Autoprep system. Kaplan-Meier analysis on the IC2006-04 study alone showed that the aCTC and mCTC had the same prognostic value. Therefore, future breast cancer samples, processed using current CellSearch equipment, can be accurately analyzed with our algorithm.

Counts of both aCTC and mCTC in samples from colorectal patients showed very low numbers. We did not find mCTC in 65% of samples and in 48% of the samples, no aCTC were found. In table 6.2 and figure 6.4 it is evident that in the baseline samples the algorithm performs good. The follow-up samples from CAIRO-2 showed that the algorithm counted more aCTC than the operators counted mCTC, but this did not improve the aCTC prognostic value as can be seen in figure 6.5B. The algorithm identified 9 more patients as being unfavorable, but this did not improve the HR and Kaplan-Meier survival estimates: HRs were 1.7/2.0 for aCTC baseline/follow-up samples and 1.9/3.8 for mCTC by stratifying the patient group on the clinically used cut-off of 3 CTC. Prostate samples that were analyzed with our algorithm showed equal prognostic value compare to mCTC: HRs were 3.4/4.1 for aCTC and 3.1/4.2 (cut-off 5 CTC). This was to be expected: the algorithm was trained on these samples and only the upper size limit of aCTC was changed since. However, it is important to notice that this change in the algorithm did not have a significant impact on our initial training set (IMMC-38) and validation set (Abiraterone).

The benign aCTC frequencies were similar to mCTC frequencies, as were the numbers found in control samples. The control samples did not contain more than one aCTC or mCTC, showing that these definitions have very few false positives. It was derived using the aCTC average from the distribution of the control samples that 2 aCTC are expected for 1 in 1750 control samples. Interestingly, benign patients had more aCTC than healthy controls, indicating that in a few of these patients the primary tumor is shedding some cells into the blood. Whether or not these cells are dangerous is currently unclear; it seems plausible that some of these patients were incorrectly diagnosed as being benign and that these aCTC numbers show evidence of progressive disease.

Comparison of aCTC morphological parameters between cancer types shows a larger size, a relatively large nucleus to the cytokeratin size and rounder CTC in breast cancer as compared to prostate and colorectal cancer. The distribution of the size of aCTC is exponential in CTC of both prostate and colorectal cancer which is cut off at the lower end by the CTC algorithm, in contrast with leukocytes that display a more bell-shaped size distribution.

These distributions shows that a wide range of CTC and CTC fragments are present in the blood of patients, as was reported earlier [46]. It was found in earlier work that the number of tumor micro particles (TMPs) is also prognostic for survival, it is therefore interesting to include these TMPs in the analysis, as they are present in abundant numbers [31]. Especially in colorectal cancer, TMPs might reduce the large Poisson error that is currently present by using a clinical cut-off value of three mCTC.

Clusters and speckled cells were found among aCTC, but not among leukocytes. A few exceptional colorectal samples contained a large number of apoptotic cells and clusters, but the presence of these cells or clusters had no detectable impact on survival. Breast cancer aCTC showed the highest number of speckled cells, while prostate cancer aCTC had more clusters, although the differences between cancer types are small. No significant rise or drop in the number of clusters or apoptotic cells was found between baseline and follow-up samples in any patient.

We performed continuous univariate and multivariate survival analysis to investigate the influence of aCTC and its morphological parameters on survival. Median roundness was significant ($p < 0.05$) when we included all cancer types in the multivariate analysis, with rounder cells predicting poorer survival. When we performed multivariate analysis by cancer type, small cells with small nuclei compared to the cytokeratin were significantly ($p < 0.05$) associated with poor survival only for breast cancer samples. Currently, we have no biological explanation for this phenomenon. However, research focused on separating normal blood cells from CTC on the basis of size should consider this observation.

In conclusion, we have validated a CTC automated algorithm, which was trained on prostate cancer samples, on other types of cancer samples and samples from benign patients and healthy controls. We show that differences exist in physical characteristics between CTC from different origin, although their impact on survival is limited. Our algorithm gives reasonable results in all cancer types, but it could be re-trained for colorectal cancer samples to include more objects. We have recently determined that all metastatic patients have CTC [49]. Hence, it is also critical for improving prognostic and predictive value of CTC that larger volumes of blood are analyzed, using for example leukapheresis techniques [50].

6.5 REFERENCES

- [1] T. Ashworth, "A case of cancer in which cells similar to those in the tumours were seen in the blood after death," *Australian Medical Journal*, vol. 14, pp. 146–147, 1869.
- [2] H. C. Engell, "Cancer cells in the circulating blood; a clinical study on the occurrence of cancer cells in the peripheral blood and in venous blood draining the tumour area at operation," *Acta Chirurgica Scandinavica. Supplementum*, vol. 201, pp. 1–70, 1955.

- [3] R. L. Myerowitz, P. A. Edwards, and G. P. Sartiano, "Carcinocythemia (carcinoma cell leukemia) due to metastatic carcinoma of breast - report of a case," *Cancer*, vol. 40, no. 6, pp. 3107–3111, 1977.
- [4] L. T. Yam and A. J. Janckila, "Immunocytodiagnosis of carcinocythemia in disseminated breast-cancer," *Acta Cytologica*, vol. 31, no. 1, pp. 68–72, 1987.
- [5] W. J. Allard, J. Matera, M. C. Miller, M. Repollet, M. C. Connelly, C. Rao, A. G. J. Tibbe, J. W. Uhr, and L. Terstappen, "Tumor cells circulate in the peripheral blood of all major carcinomas but not in healthy subjects or patients with nonmalignant diseases," *Clinical Cancer Research*, vol. 10, no. 20, pp. 6897–6904, 2004.
- [6] S. J. Cohen, C. J. A. Punt, N. Iannotti, B. H. Saidman, K. D. Sabbath, N. Y. Gabrail, J. Picus, M. Morse, E. Mitchell, M. C. Miller, G. V. Doyle, H. Tissing, L. Terstappen, and N. J. Meropol, "Relationship of circulating tumor cells to tumor response, progression-free survival, and overall survival in patients with metastatic colorectal cancer," *Journal of Clinical Oncology*, vol. 26, no. 19, pp. 3213–3221, 2008.
- [7] M. Cristofanilli, G. T. Budd, M. J. Ellis, A. Stopeck, J. Matera, M. C. Miller, J. M. Reuben, G. V. Doyle, W. J. Allard, L. Terstappen, and D. F. Hayes, "Circulating tumor cells, disease progression, and survival in metastatic breast cancer," *New England Journal of Medicine*, vol. 351, no. 8, pp. 781–791, 2004.
- [8] J. S. de Bono, H. I. Scher, R. B. Montgomery, C. Parker, M. C. Miller, H. Tissing, G. V. Doyle, L. Terstappen, K. J. Pienta, and D. Raghavan, "Circulating tumor cells predict survival benefit from treatment in metastatic castration-resistant prostate cancer," *Clinical Cancer Research*, vol. 14, no. 19, pp. 6302–6309, 2008.
- [9] M. G. Krebs, R. Sloane, L. Priest, L. Lancashire, J. M. Hou, A. Greystoke, T. H. Ward, R. Ferraldeschi, A. Hughes, G. Clack, M. Ranson, C. Dive, and F. H. Blackhall, "Evaluation and prognostic significance of circulating tumor cells in patients with non-small-cell lung cancer," *Journal of Clinical Oncology*, vol. 29, no. 12, pp. 1556–1563, 2011.
- [10] S. Matsusaka, K. Chin, M. Ogura, M. Suenaga, E. Shinozaki, Y. Mishima, Y. Terui, N. Mizunuma, and K. Hatake, "Circulating tumor cells as a surrogate marker for determining response to chemotherapy in patients with advanced gastric cancer," *Cancer Science*, vol. 101, no. 4, pp. 1067–1071, 2010.
- [11] G. Attard, A. H. M. Reid, R. A'Hern, C. Parker, N. B. Oommen, E. Folkard, C. Messiou, L. R. Molife, G. Maier, E. Thompson, D. Olmos, R. Sinha, G. Lee, M. Dowsett, S. B. Kaye, D. Dearnaley, T. Kheoh, A. Molina, and J. S. de Bono, "Selective inhibition of cyp17 with abiraterone acetate is highly active in the treatment of castration-resistant prostate cancer," *Journal of Clinical Oncology*, vol. 27, no. 23, pp. 3742–3748, 2009.
- [12] J. S. de Bono, G. Attard, A. Adjei, M. N. Pollak, P. C. Fong, P. Haluska, L. Roberts, C. Melvin, M. Repollet, D. Chianese, M. Connely, L. Terstappen, and A. Gualberto, "Potential applications for circulating tumor cells expressing the insulin-like growth factor-i receptor," *Clinical Cancer Research*, vol. 13, no. 12, pp. 3611–3616, 2007.
- [13] D. F. Hayes, T. M. Walker, B. Singh, E. S. Vitetta, J. W. Uhr, S. Gross, C. Rao, G. V. Doyle, and L. Terstappen, "Monitoring expression of her-2 on circulating epithelial cells in patients with advanced breast cancer," *International Journal of Oncology*, vol. 21, no. 5, pp. 1111–1117, 2002.
- [14] S. Meng, D. Tripathy, S. Shete, R. Ashfaq, H. Saboorian, B. Haley, E. Frenkel, D. Euhus, M. Leitch, C. Osborne, E. Clifford, S. Perkins, P. Beitsch, A. Khan, L. Morrison, D. Herlyn, L. Terstappen, N. Lane, J. Wang, and J. Uhr, "upar and her-2 gene status in individual breast cancer cells from blood and tissues," *Proceedings*

- Of The National Academy Of Sciences Of The United States Of America*, vol. 103, no. 46, pp. 17361–17365, 2006.
- [15] S. D. Meng, D. Tripathy, S. Shete, R. Ashfaq, B. Haley, S. Perkins, P. Beitsch, A. Khan, D. Euhus, C. Osborne, E. Frenkel, S. Hoover, M. Leitch, E. Clifford, E. Vitetta, L. Morrison, D. Herlyn, L. Terstappen, T. Fleming, T. Fehm, T. Tucker, N. Lane, J. Q. Wang, and J. Uhr, “Her-2 gene amplification can be acquired as breast cancer progresses,” *Proceedings Of The National Academy Of Sciences Of The United States Of America*, vol. 101, no. 25, pp. 9393–9398, 2004.
- [16] J. F. Swennenhuis, A. G. Tibbe, R. Levink, R. C. Sipkema, and L. W. Terstappen, “Characterization of circulating tumor cells by fluorescence in situ hybridization,” *Cytometry Part A*, vol. 75, no. 6, pp. 520–7, 2009.
- [17] E. Racila, D. Euhus, A. J. Weiss, C. Rao, J. McConnell, L. Terstappen, and J. W. Uhr, “Detection and characterization of carcinoma cells in the blood,” *Proceedings Of The National Academy Of Sciences Of The United States Of America*, vol. 95, no. 8, pp. 4589–4594, 1998.
- [18] C. Alix-Panabieres, J. P. Brouillet, M. Fabbro, H. Yssel, T. Rousset, T. Maudelonde, G. Choquet-Kastylevsky, and J. P. Vendrell, “Characterization and enumeration of cells secreting tumor markers in the peripheral blood of breast cancer patients,” *Journal of Immunological Methods*, vol. 299, no. 1-2, pp. 177–188, 2005.
- [19] H. J. Kahn, A. Presta, L. Y. Yang, J. Blondal, M. Trudeau, L. Lickley, C. Holloway, D. R. McCready, D. Maclean, and A. Marks, “Enumeration of circulating tumor cells in the blood of breast cancer patients after filtration enrichment: correlation with disease stage,” *Breast Cancer Research and Treatment*, vol. 86, no. 3, pp. 237–247, 2004.
- [20] R. T. Krivacic, A. Ladanyi, D. N. Curry, H. B. Hsieh, P. Kuhn, D. E. Bergsrud, J. F. Kepros, T. Barbera, M. Y. Ho, L. B. Chen, R. A. Lerner, and R. H. Bruce, “A rare-cell detector for cancer,” *Proceedings Of The National Academy Of Sciences Of The United States Of America*, vol. 101, no. 29, pp. 10501–10504, 2004.
- [21] S. Nagrath, L. V. Sequist, S. Maheswaran, D. W. Bell, D. Irimia, L. Ulkus, M. R. Smith, E. L. Kwak, S. Digumarthy, A. Muzikansky, P. Ryan, U. J. Balis, R. G. Tompkins, D. A. Haber, and M. Toner, “Isolation of rare circulating tumour cells in cancer patients by microchip technology,” *Nature*, vol. 450, no. 7173, pp. 1235–U10, 2007.
- [22] K. Pachmann, O. Camara, A. Kavallaris, S. Krauspe, N. Malarski, M. Gajda, T. Kroll, C. Jorke, U. Hammer, A. Altendorf-Hofmann, C. Rabenstein, U. Pachmann, I. Runnebaum, and K. Hoffken, “Monitoring the response of circulating epithelial tumor cells to adjuvant chemotherapy in breast cancer allows detection of patients at risk of early relapse,” *Journal of Clinical Oncology*, vol. 26, no. 8, pp. 1208–1215, 2008.
- [23] G. Vona, A. Sabile, M. Louha, V. Sitruk, S. Romana, K. Schutze, F. Capron, D. Franco, M. Pazzagli, M. Vekemans, B. Lacour, C. Brechot, and P. Paterlini-Brechot, “Isolation by size of epithelial tumor cells - a new method for the immunomorphological and molecular characterization of circulating tumor cells,” *American Journal of Pathology*, vol. 156, no. 1, pp. 57–63, 2000.
- [24] J. P. Gleghorn, E. D. Pratt, D. Denning, H. Liu, N. H. Bander, S. T. Tagawa, D. M. Nanus, P. A. Giannakakou, and B. J. Kirby, “Capture of circulating tumor cells from whole blood of prostate cancer patients using geometrically enhanced differential immunocapture (gedi) and a prostate-specific antibody,” *Lab on a Chip*, vol. 10, no. 1, pp. 27–29, 2010.

- [25] H. B. Hsieh, D. Marrinucci, K. Bethel, D. N. Curry, M. Humphrey, R. T. Krivacic, J. Kroener, L. Kroener, A. Ladanyi, N. Lazarus, P. Kuhn, R. H. Bruce, and J. Nieva, "High speed detection of circulating tumor cells," *Biosensors & Bioelectronics*, vol. 21, no. 10, pp. 1893–1899, 2006.
- [26] A. E. Saliba, L. Saias, E. Psychari, N. Minc, D. Simon, F. C. Bidard, C. Mathiot, J. Y. Pierga, V. Fraasier, J. Salamero, V. Saada, F. Farace, P. Vielh, L. Malaquin, and J. L. Viovy, "Microfluidic sorting and multimodal typing of cancer cells in self-assembled magnetic arrays," *Proceedings Of The National Academy Of Sciences Of The United States Of America*, vol. 107, no. 33, pp. 14524–14529, 2010.
- [27] H. Schwarzenbach, C. Alix-Panabieres, I. Muller, N. Letang, J. P. Vendrell, X. Rebillard, and K. Pantel, "Cell-free tumor dna in blood plasma as a marker for circulating tumor cells in prostate cancer," *Clinical Cancer Research*, vol. 15, no. 3, pp. 1032–1038, 2009.
- [28] *Lab-on-a-chip for the isolation and characterization of circulating tumor cells*, IEEE Engineering in Medicine and Biology Society Conference Proceedings, 2010.
- [29] A. H. Talasaz, A. A. Powell, D. E. Huber, J. G. Berbee, K. H. Roh, W. Yu, W. Z. Xiao, M. M. Davis, R. F. Pease, M. N. Mindrinos, S. S. Jeffrey, and R. W. Davis, "Isolating highly enriched populations of circulating epithelial cells and other rare cells from blood using a magnetic sweeper device," *Proceedings Of The National Academy Of Sciences Of The United States Of America*, vol. 106, no. 10, pp. 3970–3975, 2009.
- [30] S. Y. Zheng, H. K. Lin, B. Lu, A. Williams, R. Datar, R. J. Cote, and Y. C. Tai, "3d microfilter device for viable circulating tumor cell (ctc) enrichment from blood," *Biomedical Microdevices*, vol. 13, no. 1, pp. 203–213, 2011.
- [31] F. A. W. Coumans, C. J. M. Doggen, G. Attard, J. S. de Bono, and L. W. M. M. Terstappen, "All circulating epcam+ck+cd45- objects predict overall survival in castration-resistant prostate cancer," *Annals of Oncology*, vol. 21, no. 9, pp. 1851–7, 2010.
- [32] J. Kraan, S. Sleijfer, M. H. Strijbos, M. Ignatiadis, D. Peeters, J. Y. Pierga, F. Farace, S. Riethdorf, T. Fehm, L. Zorzino, A. G. Tibbe, M. Maestro, R. Gisbert-Criado, G. Denton, J. S. de Bono, C. Dive, J. A. Foekens, and J. W. Gratama, "External quality assurance of circulating tumor cell enumeration using the cellsearch(r) system: A feasibility study," *Cytometry B Clin Cytom*, vol. 80B, no. 2, pp. 112–118, 2010.
- [33] A. G. J. Tibbe, M. C. Miller, and L. Terstappen, "Statistical considerations for enumeration of circulating tumor cells," *Cytometry Part A*, vol. 71A, no. 3, pp. 154–162, 2007.
- [34] S. T. Ligthart, F. A. W. Coumans, G. Attard, A. Mulick Cassidy, J. S. De Bono, and L. W. M. M. Terstappen, "Unbiased and automated identification of a circulating tumour cell definition that associates with overall survival," *Plos One*, 2011.
- [35] D. Marrinucci, K. Bethel, R. H. Bruce, D. N. Curry, B. Hsieh, M. Humphrey, R. T. Krivacic, J. Kroener, L. Kroener, A. Ladanyi, N. H. Lazarus, J. Nieva, and P. Kuhn, "Case study of the morphologic variation of circulating tumor cells," *Human Pathology*, vol. 38, no. 3, pp. 514–519, 2007. ISI Document Delivery No.: 139VG Times Cited: 10 Cited Reference Count: 15 Marrinucci, Dena Bethel, Kelly Bruce, Richard H. Curry, Douglas N. Hsieh, Ben Humphrey, Mark Krivacic, Robert T. Kroener, Joan Kroener, Lindsay Ladanyi, Andras Lazarus, Nicole H. Nieva, Jorge Kuhn, Peter W b saunders co-elsevier inc Philadelphia.
- [36] D. Marrinucci, K. Bethel, D. Lazar, J. Fisher, E. Huynh, P. Clark, R. Bruce, J. Nieva, and P. Kuhn, "Cytomorphology of circulating colorectal tumor cells: A small case series," *Journal of Oncology*, vol. 2010, 2010.

- [37] J. Y. Pierga, F. C. Bidard, F. Andre, T. Petit, F. Dalenc, T. Delozier, G. Romieu, J. Bonnetterre, J. Ferrero, P. Kerbrat, A. Martin, and P. Viens, “Early drop of circulating tumor cells (ctc) and increase of circulating endothelial cells (cec) during neoadjuvant chemotherapy (ct) combined with bevacizumab in her2-negative inflammatory breast cancer (ibc) in multicenter phase ii trial beverly 1,” in *ASCO*, (Chicago, Illinois), p. Abstract no 10510, 2011.
- [38] J. Tol, M. Koopman, M. C. Miller, A. Tibbe, A. Cats, G. J. M. Creemers, A. H. Vos, I. D. Nagtegaal, L. Terstappen, and C. J. A. Punt, “Circulating tumour cells early predict progression-free and overall survival in advanced colorectal cancer patients treated with chemotherapy and targeted agents,” *Annals of Oncology*, vol. 21, no. 5, pp. 1006–1012, 2010.
- [39] G. Attard, A. H. M. Reid, T. A. Yap, F. Raynaud, M. Dowsett, S. Settatree, M. Barrett, C. Parker, V. Martins, E. Folkerd, J. Clark, C. S. Cooper, S. B. Kaye, D. Dearnaley, G. Lee, and J. S. de Bono, “Phase i clinical trial of a selective inhibitor of cyp17, abiraterone acetate, confirms that castration-resistant prostate cancer commonly remains hormone driven,” *Journal of Clinical Oncology*, vol. 26, no. 28, pp. 4563–4571, 2008.
- [40] M. R. de Groot, H. M. Croonen, W. J. T. Mastboom, I. Vermes, A. G. J. Tibbe, H. Tissing, and L. W. M. M. Terstappen, “Circulating tumor cells (ctc) in newly diagnosed breast or colorectal cancer,” *Proc Ann Meet Am Soc Clin Onc*, vol. 25, no. 18s, 2007.
- [41] B. Franken, M. R. De Groot, W. J. B. Mastboom, I. Vermes, J. Van der Palen, A. G. J. Tibbe, and L. W. M. M. Terstappen, “Circulating tumor cells, disease recurrence and survival in newly diagnosed breast cancer,” in *SABC*, (San Antonio, TX), 2011.
- [42] L. Scholten, L. W. M. M. Terstappen, J. van der Palen, W. J. B. Mastboom, A. G. J. Tibbe, I. Vermes, and M. R. de Groot, “Circulating tumor cells as a possible prognostic tool in newly diagnosed nonmetastatic colorectal cancer?,” *Journal of Clinical Oncology*, vol. 30, p. suppl 4; abstr 395, 2012.
- [43] M. Kagan, D. Howard, T. Bendele, J. Mayes, J. Silvia, M. Repollet, J. Doyle, J. Al-lard, N. Tu, T. Bui, T. Russell, C. Rao, M. Hermann, H. Rutner, and L. Terstappen, “A sample preparation and analysis system for identification of circulating tumor cells,” *Journal of Clinical Ligand Assay*, vol. 25, no. 1, pp. 104–110, 2002.
- [44] G. W. Zack, W. E. Rogers, and S. A. Latt, “Automatic-measurement of sister chromatid exchange frequency,” *Journal Of Histochemistry & Cytochemistry*, vol. 25, no. 7, pp. 741–753, 1977.
- [45] S. T. Ligthart, F. C. Bidard, C. Decraene, T. Bachelot, S. Delaloge, E. Brain, M. Campone, P. Viens, J. Y. Pierga, and L. W. M. M. Terstappen, “Unbiased quantitative assessment of her-2 expression of circulating tumor cells in patients with metastatic and non metastatic breast cancer,” *submitted*, 2012.
- [46] C. J. Larson, J. G. Moreno, K. J. Pienta, S. Gross, M. Repollet, S. M. O’Hara, T. Russell, and L. Terstappen, “Apoptosis of circulating tumor cells in prostate cancer patients,” *Cytometry Part A*, vol. 62A, no. 1, pp. 46–53, 2004.
- [47] S. T. Ligthart, J. F. Swennenhuis, J. Greve, and L. W. M. M. Terstappen, “Fish probe counting in circulating tumor cells,” in *Cytokeratins: Tools in Oncology* (G. Hamilton, ed.), pp. 119–134, Rijeka: InTech, 2012.
- [48] H. I. Scher, X. Y. Jia, J. S. de Bono, M. Fleisher, K. J. Pienta, D. Raghavan, and G. Heller, “Circulating tumour cells as prognostic markers in progressive, castration-resistant prostate cancer: a reanalysis of immc38 trial data,” *Lancet Oncology*, vol. 10, no. 3, pp. 233–239, 2009.

- [49] F. Coumans, S. T. Ligthart, and L. W. M. Terstappen, “All patients with metastatic carcinoma have circulating tumor cells,” *submitted*, 2012.
- [50] R. L. Eifler, J. Lind, D. Falkenhagen, V. Weber, M. B. Fischer, and R. Zeillinger, “Enrichment of circulating tumor cells from a large blood volume using leukapheresis and elutriation: Proof of concept,” *Cytometry Part B-Clinical Cytometry*, vol. 80B, no. 2, pp. 100–111, 2011.

INTERPRETATION OF CHANGES IN CIRCULATING TUMOR CELL COUNTS¹

Sjoerd T. Ligthart, Frank A.W. Coumans, and Leon W.M.M. Terstappen

ABSTRACT

Background: The presence of circulating tumor cells (CTC) in the blood of cancer patients may guide the use of therapy. We investigated how to evaluate a reduction in the number of CTC after administration of therapy by using different definitions of CTC. **Methodology:** CTC were enumerated with the CellSearch[®] system in 111 metastatic breast and 185 metastatic prostate cancer patients before start of a new line of chemotherapy and after initiation of therapy. CTC were enumerated by manual review and by an automated algorithm using four selection criteria. Different means to express changes in CTC counts were evaluated with respect to progression free (PFS) and overall survival (OS): by applying a static cut-off, absolute reduction, relative reduction, and a Poisson model. **Results:** It was determined that a static CTC cut-off is the best method to determine whether a therapy is effective. This is exemplified by the highest Cox hazard ratios for OS of 2.1–3.0 and for PFS of 2.0–2.9 for all CTC definitions; three methods to express relative differences performed worse. We furthermore show that a very strict CTC definition is the most useful because of low background in healthy controls. A look-up table is provided from which the significance of a change in CTC can be derived. **Conclusion:** Results show that the aim of therapy should be the elimination of all CTC, and that elimination of CTC in 10–12 weeks after therapy initiation is early evidence of therapy success.

¹Submitted.

7.1 INTRODUCTION

Circulating tumor cells (CTC) leave the tumor site(s), invade the blood stream, and may give rise to distant metastases. The presence of CTC has been associated with poor survival in melanoma, breast, colorectal, prostate, gastric, and non-small cell lung cancer [1, 2, 3, 4, 5, 6]. The association between the presence of CTC and poor prognosis holds true before and at various time intervals after initiation of therapy [3, 4, 7]. Transition of the number of CTC in a patient from unfavorable (≥ 5 CTC/7.5 ml of blood) to favorable (< 5 CTC/7.5 ml of blood) improves survival and as such can be used as a predictive factor for treatment response [2, 3, 4, 7]. This transition can already be measured after the first cycle of therapy, which is considerably earlier compared to treatment response measured by imaging. Moreover, conversion to favorable or unfavorable CTC may be a better indicator of therapy response as compared to imaging [3, 8, 9]. Although conversion to no detectable CTC is most likely the best outcome, the time needed to assure whether or not this target could be reached under the current treatment regimen is unknown. Accurate measurement of a decline in CTC is impeded by the low frequency of CTC in most patients with metastatic disease [10]. Not only the continuous relationship between CTC concentration and survival but also the Poisson sampling error suggests that dichotomization may not be the best approach to detect effective chemotherapy [11, 12, 13].

The goal of this study was to investigate how to optimally define a true decrease in the number of CTC and to determine whether less strict definitions of CTC could aid in the assessment of a CTC reduction. To address these questions we estimated the prognostic power of various methods to define a reduction in the number of CTC.

7.2 MATERIALS AND METHODS

7.2.1 PATIENT DATA

For this study we used stored images of CTC measurements from patients before initiation of a new cycle of cytotoxic chemotherapy (baseline) and at several follow-up time-points in a thirteen week period after initiation of therapy. Samples from patients enrolled into the IMMC-01 metastatic breast cancer [2] and IMMC-38 metastatic prostate cancer [4] studies were used. This study included 111 baseline samples and 265 follow up samples from 111 metastatic breast cancer and 185 baseline and 425 follow-up samples from 185 metastatic prostate cancer patients. 205 samples from healthy controls were available [10, 14]. All patients and healthy controls provided written informed consent. The institutional review boards at each participating center approved the respective study protocols. Table 7.1 shows a summary of patient characteristics.

TABLE 7.1 : Patient characteristics. Numerical data are shown as median (min-max) or % of total N. The sample size N is given in parentheses if different from the total N.

	Breast	Prostate
N	111	185
Age	58 (27-86)	69 (49-92)
OS ^a of Px ^b	16.1 (1.5-48.6)	16.9 (1.9-38.7)
PFS ^c of Px	5.5 (0.7-45.2)	5.0 (0.6-34.5)
Follow-up of censored Px	20.7 (5.5-48.6, N=47)	25.9 (2.4-38.9, N=67)
Types of Tx ^d (1/2/3) ^e	57.8%/5.5%/36.7%	8.2%/91.8%/0%
Tx line (1/2/3+)	43.2%/15.3%/41.5%	67.0%/16.2%/16.8%
Bone mets (yes/no)	90.6%/9.4%	87.4%/12.6%
PSA ^g (ng/ml)	N/A ^f	128 (1.9-17800)
LDH ^h (IU/ml)	N/A	236 (4.7-2364, N=139)
ALP ⁱ (IU/ml)	N/A	126 (3.29-1558, N=179)
Hemoglobin (g/dl)	N/A	12.4 (8.2-15.7, N=182)
Albumin (g/dl)	N/A	3.8 (2.1-41, N=179)
CA15.3 ^j	169.5 (0.9-19315, N=22)	N/A
CA27.29 ^k	96 (9.9-23204, N=65)	N/A
CEA ^l	8.2 (0.5-3920, N=38)	N/A
mCTC ^m	6 (0-10194)	7 (0-5925)

^a Overall survival, ^b patients, ^c progression free survival, ^d treatment, ^e Tx types: 1=chemo+other; 2=chemo, hormone+other; 3=hormone+other, ^f not applicable, ^g prostate specific antigen, ^h lactate dehydrogenase, ⁱ alkaline phosphatase, ^j mucin 1 glycoprotein, ^k milk mucin antigen, ^l carcinoembryonic antigen, ^m manual CellSearch circulating tumor cells.

7.2.2 CTC ENUMERATION

The CellSearch[®] system (Veridex LLC, Raritan, USA) was used for enumeration of CTC. The system consists of a CellTracks Autoprep[®] for sample preparation and a CellTracks Analyzer II[®] for sample analysis [10]. The CellTracks Autoprep immuno-magnetically enriches epithelial cells from 7.5 ml of blood using ferrofluids conjugated to the epithelial cell adhesion molecule (EpCAM). The enriched sample is stained with phycoerythrin-conjugated (PE) antibodies directed against cytokeratins (CK) 8, 18 and 19, an allophycocyanin-conjugated (APC) antibody to CD45 and the nuclear dye 4',6-diamidino-2-phenylindole (DAPI). The enriched sample (~300 μ l in volume) is transferred to a cartridge placed in a CellTracks Magnest[®]. Inside the MagNest magnetically labeled objects are pulled towards the imaging surface. The surface is imaged by the CellTracks Analyzer II which records digital images for four different fluorescent dyes, using a 10 \times /0.45NA objective. Per cartridge, 144–180 four-layer tiff images are recorded. To arrive at a CTC count, objects that are CK+DAPI+ are selected by a computer algorithm and presented in a gallery to a trained reviewer. The reviewer manually scores events as CTC (mCTC) when the objects are

TABLE 7.2 : Characteristics of the mCTC and aCTC definitions A through D used to test different reduction criteria. Thresholds for each parameter were defined elsewhere [15]. Size is shown as a diameter assuming a round cell. Log-rank $p < 0.0001$ for all HRs

CTC definition	CK ^b	DNA	CD45	Size (μm)	Median count		HR ^a	
					BL ^c	FU1 ^d	BL	FU1
mCTC	+	+	-	>4	7	1	2.6	3.8
aCTC A	>50	>170	<60	6–16	2	1	2.3	3.0
aCTC B	>80	>120	<160	3–30	10	5	3.2	3.7
aCTC C	>40	>80	<180	2–30	22	13	3.0	3.4
aCTC D	>20	any	any	6–30	68	40	2.6	2.8

^a Hazard ratio, ^b cytokeratins 8, 18, and 19, ^c baseline measurement, ^d first follow-up measurement.

CD45 negative, larger than 4 μm and have a cell-like morphology.

For performing automated CTC (aCTC) enumeration, the same 144–180 tiff images are used as input to count aCTC by a fully automated algorithm [15]. This algorithm selects potential aCTC by object segmentation in the CK channel and measures various properties of the object. Next, the algorithm applies criteria to arrive at a final aCTC count. For this study four different definitions of aCTC were developed using overall survival (OS) as training parameter. We defined a classifier as a combination of selection criteria; by changing these criteria we can change the sensitivity and specificity of the classifier. The characteristics of the four aCTC definitions -together with the mCTC definition- used in this study are provided in table 7.2. These definitions were used to gauge whether one of these definitions was particularly suited for assessment of a change in CTC number.

7.2.3 POISSON MODEL FOR REDUCTION IN CTC COUNT

The three major sources of variation in CTC counting are [12]: i) the Poisson distribution of the number of CTC present in a randomly drawn sample; ii) the variability in the enrichment efficiency, and iii) the intra- and inter-reader variability. The latter source of error disappears when counting with an automated algorithm. With $80 \pm 15\%$ assay efficiency [12], the Poisson error is dominant up to 44 CTC and only this error will be taken into account. When taking a sample from a large blood volume with an average concentration of λ objects/sample volume, the probability that when we take a sample we have k objects in that sample is given by the Poisson distribution:

$$P(X = k) = \frac{e^{-\lambda} \lambda^k}{k!} \text{ for } k = 0, 1, 2, \dots \quad (7.1)$$

The probability that a pair of successive measurements represents a reduction in the true mean number of CTC was determined in Matlab 2009a

(Mathworks, Natick, MA). Briefly, for two measurements k_1 and k_2 the probability was simulated that λ_2 is smaller than λ_1 , as is represented in equation 7.2:

$$P(\lambda_2 < \lambda_1) = \int_0^\infty \left(P(X = k_2) \int_{\lambda_1 \geq \lambda_2}^\infty P(X = k_1) d\lambda_1 \right) d\lambda_2 \quad (7.2)$$

In which k_1, λ_1 represent the first and k_2, λ_2 the second CTC measurement. This equation compares two distributions of underlying means λ that were derived from two measured k 's. First, for the two k 's, the probability is simulated that it originates from a certain mean λ , this distribution is normalized to 1. Second, for every λ from k_2 the probability is determined that it is lower than the λ 's from k_1 . This total is summed to determine the total probability that k_2 came from a lower true mean than k_1 . There are an infinite number of λ 's were a CTC measurement could originate from, in our simulation we chose a λ distribution with 0.01 CTC spacing. Using equation 7.2, we can determine the confidence for a CTC reduction given any pair of measurements. For each combination of baseline and follow-up measurements ranging from 0–50, the confidence that a true reduction occurred was calculated. For aCTC definitions with high background noise, the range where the Poisson error is dominant is reduced.

7.2.4 CTC REDUCTION CRITERIA

Several criteria to express CTC reduction were tested for their ability to predict favorable outcome. To allow use of the same patients for each reduction criterion, only patients with ≥ 5 mCTC at baseline were used for this analysis (N=164). The traditional CTC numbers and the four different aCTC definitions were used in this comparison. Tested criteria were:

1. Conversion to favorable group (from $\geq x$ to $< x$ CTC): setting a static cut-off.
2. Reduction confidence larger than $x\%$: according to equation 2.
3. Measurement 2 is x smaller than measurement 1: setting a absolute reduction cut-off.
4. Measurement 2 is x times smaller than measurement 1: setting a proportional reduction cut-off.

The parameter x was determined for each reduction criterion and CTC definition as described in statistical analysis.

7.2.5 STATISTICAL ANALYSIS

The Cox proportional hazard ratio (HR) for each aCTC classifier was determined by dichotomizing on the median of the number of aCTC found at baseline. For reference, the HR for mCTC was also determined in the same way. We did not apply the cut-off of 5 (which is clinically used for mCTC) to the aCTC definitions, because this would distort the relative size of favorable and unfavorable groups.

For each criterion, the HR was determined for OS unless otherwise stated. A method which delivers a high HR, but distinguishes very few metastatic patients is of no practical use. To determine how well different criteria stratify patients into two groups, the relative size of the two groups resulting from each criterion must be balanced. To achieve this, the x in the criterion 1 through 4 was set such that the percentages of patients in the favorable and unfavorable groups were similar to the clinically validated mCTC cutoff in criterion 1. The percentage of patients who remained unfavorable was found to be 59% using criterion 1 for $x = 5$ mCTC. Thus for each of the mCTC and aCTC definitions from table 7.2, the variable x was set for each criterion to approximate the unfavorable group at 59%.

OS was measured from the time of blood draw to time of death from any cause. Patients were censored at last follow-up if progression or death had not occurred. Survival curves were compared with the use of log-rank testing. Progression free survival (PFS) was defined as the elapsed time between the date of the blood draw and the date of progression by either CT scans and/or clinical signs and symptoms or death.

7.3 RESULTS

7.3.1 OVERALL SURVIVAL AS A FUNCTION OF CTC NUMBER

mCTC were enumerated with the CellSearch system in 296 blood samples at the first follow-up after initiation of a new line of therapy in 111 metastatic breast and 185 metastatic prostate cancer patients. Patients were divided into those with 0 mCTC/7.5 ml of whole blood ($N=123$, 42%), those between 1 and 4 mCTC ($N=67$, 23%), those between 5 and 24 mCTC ($N=55$, 19%) and those with ≥ 25 mCTC ($N=51$, 17%).

Figure 7.1 shows the Kaplan-Meier plot of these 296 patients. Log-rank p for a comparison between the survival curves of patients with 0 versus 1–4 mCTC was $p=0.044$, 1–4 mCTC versus 5–24 mCTC $p=0.002$, and 5–24 mCTC versus ≥ 25 mCTC $p=0.003$, all other $p < 0.0001$.

7.3.2 CHANGES IN THE NUMBER OF CTC AND OVERALL SURVIVAL

Metastatic breast and prostate cancer patients were separated into 164 (55%) patients with unfavorable mCTC (≥ 5 mCTC) and those 132 (45%) with favorable mCTC (< 5 mCTC) at baseline. To illustrate the relationship

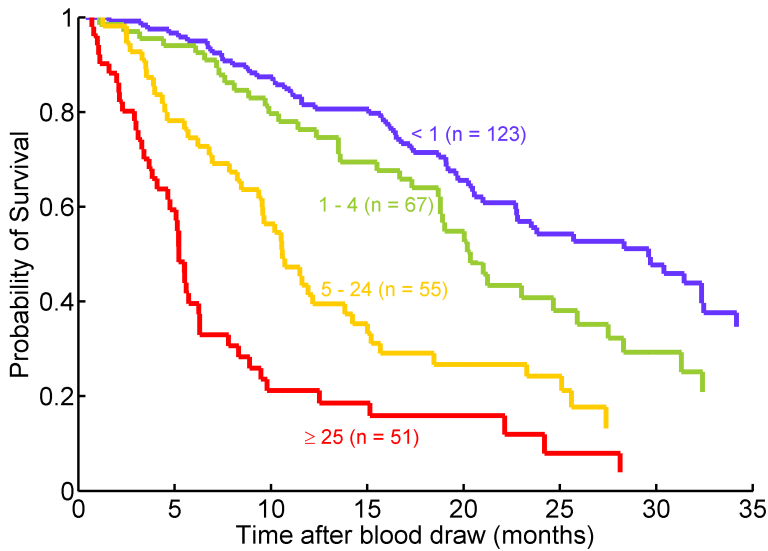


FIGURE 7.1 : Overall survival time of metastatic breast and prostate cancer patients was calculated from the date of the first follow-up blood draw after initiation of a new line of therapy. Kaplan-meier plots of the probability of overall survival for 123 patients with <1 mCTC, 67 patients with 1–4 mCTC, 55 patients with 5–24 mCTC and 51 patients with ≥ 25 mCTC is shown.

between changes in mCTC count and overall survival, patients were further subdivided into groups that survived 1–6 months, 6–12 month, 12–24 month and 24–36 months. From the 296 patients 690 blood samples were taken in 4–6 week intervals after initiation of therapy. Smoothed spline fits were used to show the relationship between the median number of mCTC and time for the different survival groups.

Figure 7.2 shows the median number of mCTC as a function of time for patients with unfavorable baseline mCTC in panel A and favorable baseline mCTC in panel B. The patients that survived for only 1–6 months showed a clear trend of increasing mCTC for both the patients groups with unfavorable and with favorable baseline mCTC. In absolute terms, the changes are less severe for patients with <5 mCTC at baseline. In relative terms they are comparable: a 3–4 fold increase in mCTC over a three month period. For patients surviving 6–12 months a trend of slight mCTC decrease followed by an increase was observed for patients with unfavorable baseline mCTC and a trend of steadily increasing numbers for those with favorable baseline mCTC. For patients surviving 12–24 months and 24–36 months a clear trend of decreasing mCTC was observed that reached 0 mCTC after 10–12 weeks of therapy.

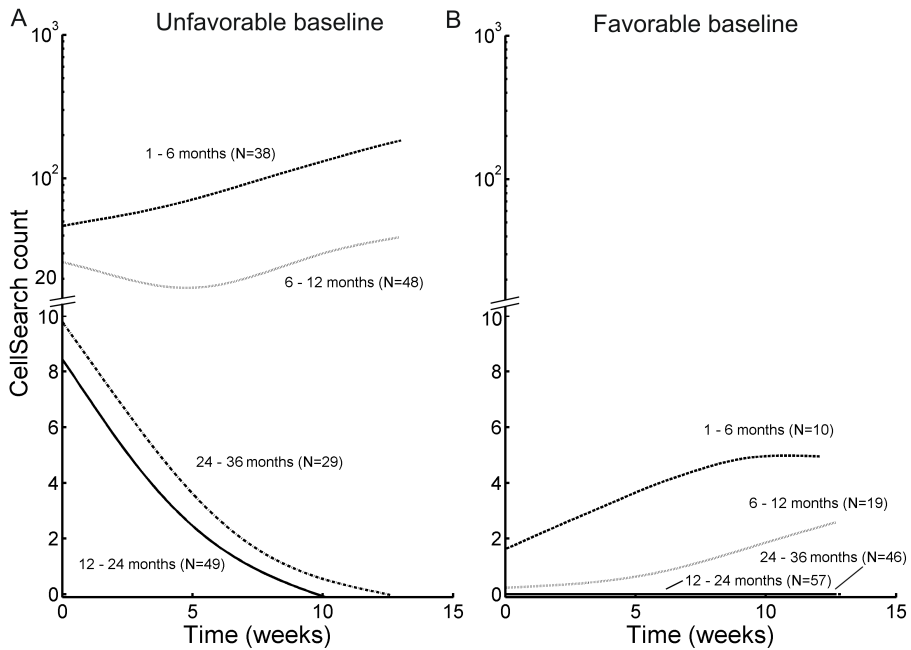


FIGURE 7.2 : mCTC trends for patients that survived 1–6, 6–12, 12–24 and 24–36 months after initiation of therapy. mCTC trends were obtained by smoothed spline fits to the number of follow-up mCTC at several time points. Panel A shows 164 baseline and 369 follow-up samples from patients with unfavorable mCTC before initiation of therapy and Panel B shows 132 baseline and 312 follow-up samples from patients with favorable mCTC before initiation of therapy.

7.3.3 TRUE CTC CHANGES DETERMINED USING A POISSON MODEL

Measurements on samples taken from a patient at successive time points may show a decline in CTC. For low CTC number this decline may reflect a true decline in the number of CTC in the patient's blood, but it may also be due to the Poisson sampling error. A look-up table shown in figure 7.3 was created that can be used as a reference to determine confidence for a true CTC reduction in the range of 0–50 CTC.

For example, in case the baseline measurement is 7 CTC: detecting 0–1 CTC at follow-up provides a CTC reduction confidence of >95% (see the lower arrow in figure 7.3), whereas for a CTC follow-up count of 5 to 9 this confidence reduces to 25–75%. A follow-up measurement of 10–14 CTC gives confidence of 5–25% and for a follow up count above 14 CTC confidence of reduction is less than 5%, or 95% confidence of an increase (see the upper arrow in figure 3). For larger CTC numbers at baseline and follow-up, the relative width of the central confidence windows becomes narrower, making it easier to detect a true change in the CTC number.

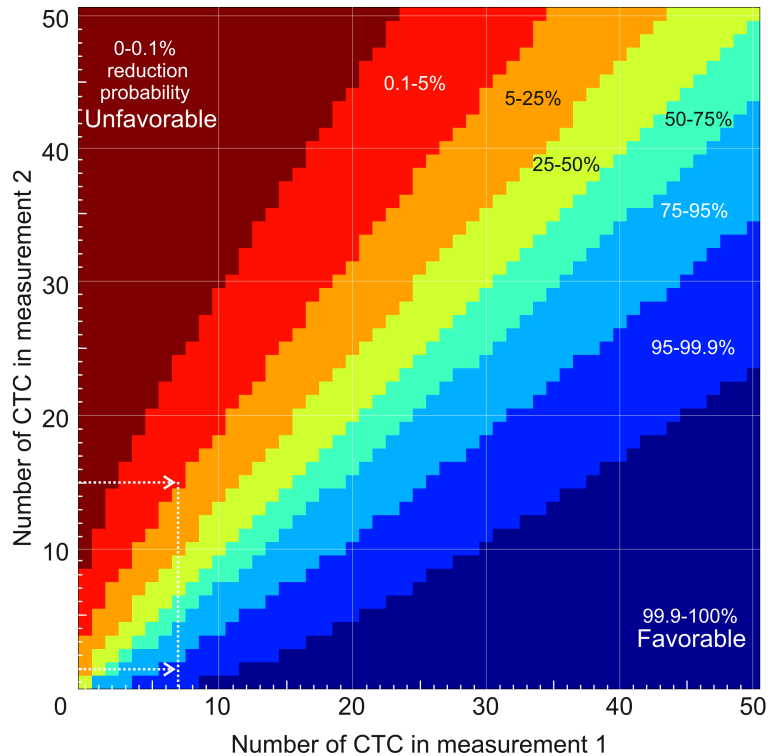


FIGURE 7.3 : Look-up table for the probability of a true reduction in CTC count between two measurements. A true reduction occurs when the true average present in a patient at the time of the second measurement is lower than the true average at the time of the first measurement.

7.3.4 RELATION BETWEEN CTC DEFINITIONS AND CLINICAL OUTCOME

The look-up table in figure 7.3 illustrates that low CTC numbers result in a large uncertainty when a change in the true average CTC number present in a patient is determined. To evaluate whether it is possible to reduce this uncertainty by applying less stringent definitions of CTC- aCTC were identified by a computer algorithm and classified according to different criteria as shown in table 7.2. The table also shows the median CTC count for each CTC definition at baseline and at the first follow-up measurements and the resulting HR. All CTC definitions resulted in a significant difference in survival of both groups ($p < 0.0001$ for all) and HRs ranged between 2.3 (95% CI 1.7–3.2) and 3.2 (95% CI 2.3–4.3) at baseline and between 2.8 (95% CI 2.0–3.8) and 3.8 (95% CI 2.8–5.1) at first follow-up.

TABLE 7.3 : HRs using OS/PFS (% unfavorable patients) using different criteria for CTC reduction for mCTC and aCTC definitions (N=164, all patients ≥ 5 mCTC at baseline). Bold HRs had a logrank $p < 0.05$ for both OS and PFS.

	Static cut-off	Reduction confidence	Relative Reduction	Absolute Reduction
CTC definition	$FU1^a < x$	$> \% \text{ certain}$	$FU1 < BL^b / x$	$FU1 < BL - x$
mCTC	2.1/2.5 (59%)	0.8/0.7 (62%)	1.5/1.9 (59%)	0.7/0.8 (58%)
aCTC A	2.7/2.8 (49%)	1.0/1.1 (61%)	1.5/1.5 (60%)	0.9/0.8 (60%)
aCTC B	3.0/2.9 (58%)	1.0/1.0 (60%)	1.5/1.5 (58%)	0.8/0.9 (59%)
aCTC C	2.6/2.4 (59%)	1.0/1.0 (53%)	1.4/1.4 (57%)	0.8/0.8 (59%)
aCTC D	2.3/2.0 (60%)	1.5/1.5 (38%)	1.3/1.4 (61%)	0.9/1.1 (59%)

^a First follow-up measurement, ^b baseline measurement.

7.3.5 CORRELATION OF REDUCTION CRITERIA WITH SURVIVAL

Four different criteria for CTC reduction were evaluated together with the different CTC definitions. We tested a static cut-off, a confidence of reduction, a proportional reduction, and an absolute reduction. The boundaries for each criterion when applied to the mCTC definition are shown in figure 7.4. The figure shows the mCTC number at baseline and first-follow-up of 164 of 296 (65%) patients with unfavorable mCTC at baseline. For the standard mCTC definition the static cut-off for unfavorable mCTC is 5 CTC, the 99.9% probability for mCTC reduction was based on the Poisson model, the minimum relative reduction was set at five fold and the minimum absolute reduction was set at 13 mCTC.

Using these specific values in each approach resulted in a dichotomization that had 59% of the patients in the unfavorable group, thus the group sizes remained constant. For illustration purposes, two arrows are shown in figure 7.4 to exemplify the implications of applying these criteria. Arrow 1 is a patient that started with 11 mCTC, which reduced to 2 mCTC at first follow-up. According to static mCTC cut-off and relative mCTC reduction this patient has a favorable prognosis, but using the 99.9% mCTC reduction confidence or the absolute mCTC decrease this patient has an unfavorable prognosis. The opposite case is indicated by arrow 2, where a patient had 39 mCTC at baseline and 9 mCTC at first follow-up; static mCTC cut-off and relative mCTC reduction give this patient an unfavorable prognosis but 99.9% mCTC reduction confidence and absolute mCTC decrease indicate a favorable prognosis.

The HRs for OS and PFS for the different CTC definitions and methods to express CTC changes for patients with unfavorable CTC at baseline (≥ 5 mCTC) are shown in table 7.3. The static cut-off predicted a significant difference in OS and PFS for all definitions ($p < 0.001$ for all HRs). The aCTC definitions A and B gave the highest predictive power using the static cut-off method, but 95% confidence intervals overlapped for all definitions

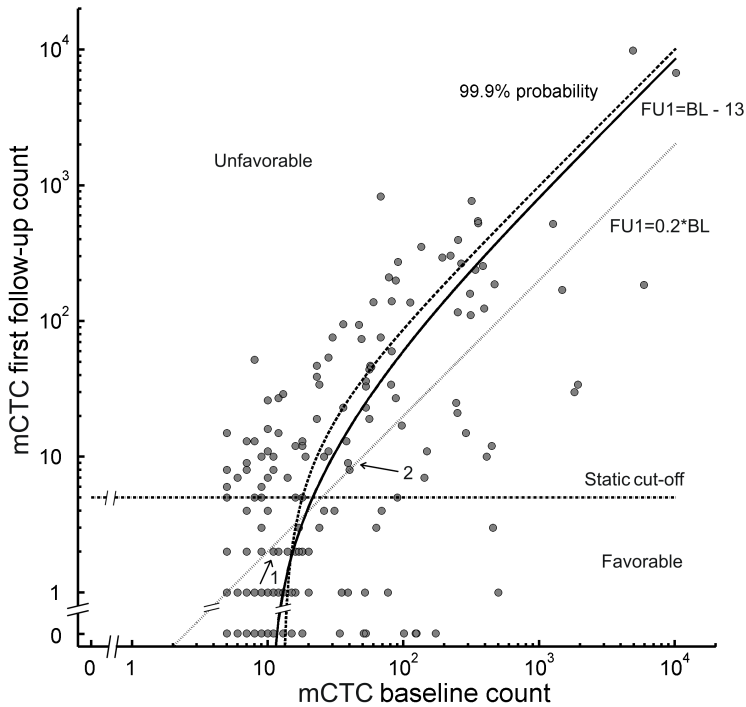


FIGURE 7.4 : Conversion from unfavorable to favorable mCTC using the classic mCTC definition of 164 metastatic breast and prostate cancer patients that had ≥ 5 mCTC at baseline. Four different criteria of assessing a CTC reduction are indicated with different lines. Arrows indicate patients described in more detail in the text. BL = number of mCTC at baseline, FU1 = number of mCTC at first follow-up.

(data not shown). Using the confidence reduction method, only the aCTC D definition was significant ($p \leq 0.05$) for OS/PFS ($p = 0.029/0.001$), although the group size could not match that of the mCTC definition. Applying the relative model showed that only aCTC definitions A ($p = 0.046/0.035$) and B ($p = 0.045/0.040$) predicted significant differences in OS/PFS, while the absolute method did not predict a significant difference in survival for all definitions ($p > 0.06$ for all HRs). Confidence of reduction (with mCTC) was at least 50% for 74% of patients, and at least 95% confidence for 56% of patients.

7.3.6 APPLYING CRITERIA FOR MCTC CHANGE TO MULTIPLE TIME POINTS

In order to create clear-cut rules for deciding whether or not a certain therapy is working, survival of the group of unfavorable patients at baseline

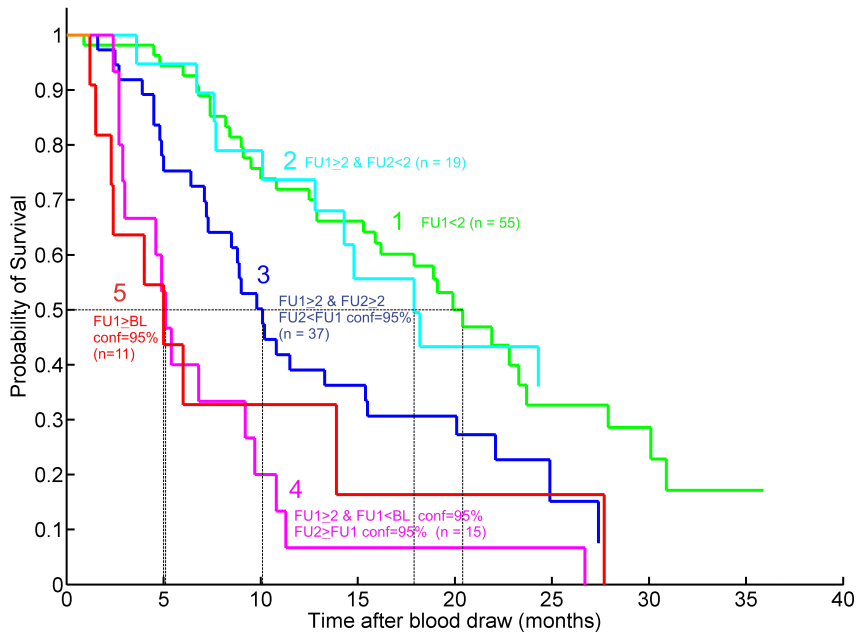


FIGURE 7.5 : Kaplan-Meier plot of patients with ≥ 2 aCTC at baseline who were monitored at two subsequent follow-up time points ($N=137$). Patients were grouped (see also text) accordingly to a drop below a cut-off of 2 mCTC, 95% confidence of a rise, or in between both. BL = baseline measurement; FU1 = first follow-up measurement; FU2 = second follow-up measurement.

was assessed for two subsequent follow-up measurements: the first between 2–5 weeks and the second between 6–8 weeks after initiation of therapy. Figure 7.1 and 7.2 show that it is important to treat patients until the number of CTC becomes zero. It was found that the number of mCTC and aCTC in healthy controls ranged from 0–1 mCTC: one of 205 healthy controls had one mCTC (0.5%), six controls had one aCTC (2.9%). Therefore, the cut-off for deciding if a patient was unfavorable was set to ≥ 2 mCTC. This reduction of the cut-off from five to two aCTC allowed for 137 patients to be included in the analysis.

The Kaplan-Meier survival plot of this analysis is shown in figure 7.5, for which the patients were divided into five groups: (1) patients who changed to favorable (< 2 mCTC) at first follow-up (green line in figure 7.5); (2) patients who stayed unfavorable at first follow-up, but changed to favorable at second follow-up (cyan line); (3) patients who stayed unfavorable at first and second follow-up, but did not show a significant rise in the number of mCTC (blue line); (4) patients who stayed unfavorable at first follow-up without significant rise that change to a significant rise at second follow-up (magenta line); (5) patients who showed a significant rise at first follow-up

(red line). In these cases, a significant rise was defined from figure 7.2, panel A: the unfavorable group with survival from 1–6 months showed a relative rise of 50% (or factor 1.5 increase). This increase represents 95% confidence of a true change as calculated by our model. For patients that had lower than 20 mCTC, the factor of 1.5 does not give 95% confidence due to the higher Poisson noise at these lower numbers: a rise from two to three mCTC would then already be significant. For these patients, the look-up table showed in figure 7.3 was used to determine if a rise was significant. If two subsequent measurements are located in the bright or dark red part of this look-up table, the rise was at least 95% confident. Median survival for groups 1–5 was 20.4, 17.9, 10.1, 5.1, and 5.0, respectively. Log-rank ps for comparison between the survival curves of patients of group 1 and group 2, and between groups 4 and 5 were $p > 0.8$; 2 and 3 had $p = 0.07$. Comparison of group 3 to group 5 (group 5 is closer to 3 than to group 4) resulted in $p = 0.17$.

7.4 DISCUSSION

Evidence is increasing that CTC are an independent prognostic and predictive biomarker for patients treated for metastatic carcinomas [2, 3, 4, 5, 6, 7, 8, 9, 16, 17, 18, 19, 20]. Whether an early switch of treatment based on persistence of CTC after the first cycle of therapy can prolong survival is still being investigated in ongoing clinical studies (SWOG 0500 - NCT00382018). A difficulty that arises in applying the results of a CTC assay to evaluate therapy response is the low number of CTC that are typically found in patients. Using the CellSearch system -which has been validated clinically for CTC enumeration- patients are divided into two groups: those with favorable CTC (< 5 CTC/7.5 ml of blood) and those with unfavorable CTC (≥ 5 CTC/7.5 ml of blood). However, examination of patient survival as a function of the number of CTC clearly shows a significant correlation between the CTC load and survival prospects as shown in figure 7.1 for breast and prostate cancer patients. CTC of metastatic breast and prostate cancer patients were combined for this analysis to illustrate that interpretation of CTC results can be made regardless of origin of the disease or treatment.

Elimination of all CTC clearly is the most desired outcome, the time needed for a CTC reduction may however vary (figure 7.2) and it is of importance to have the proper tools to evaluate changes in CTC numbers. Using Poisson statistics, a measure was given of the confidence that two CTC measurements represent a true change in the underlying average number of CTC found in patients. A look-up table (figure 7.3) was created to assist physicians in determining whether or not a change in CTC is significant. This look-up table shows that for 0–10 CTC, most possible combinations of CTC at baseline and follow-up have confidence less than 95% for determining decrease or increase. At baseline, 57% of patients in

prostate and 66% in breast have 0–10 mCTC. For 56% of patients in our data set, the observed reduction is significant with at least 95% confidence. However, this reduction does not result in an improvement of OS and PFS unless the absolute number of mCTC is reduced below 5 and preferably approaches 0. We attribute this inconsequential reduction to the cytotoxic chemotherapy, which kills cells in most patients, but probably does not target all malignant cells in those patients with brief survival. This also suggests that when CTC are found after therapy was initiated, these CTC may be resistant to the current therapy and might be characterized using this knowledge.

In patients where the number of mCTC is reduced but survival does not improve, therapy possibly affects only part of the metastases that give rise to CTC, or alternative therapy is given at insufficient dose or frequency to achieve a durable effect. This is reflected in the changes in CTC counts for patients that survived 1–6, 6–12, 12–24 and 24–36 months in figure 7.2 panel A. The group with 6–12 month survival showed some decline of CTC in response to therapy in the first five weeks after initiation of therapy, but deteriorated at subsequent time points. This figure further shows that a period of 10–12 weeks may be needed before definitive conclusions can be drawn whether a count is going to zero. Figure 7.5 further confirms this conclusion, in which patient samples of two time points were used. The cut-off for unfavorable patients was lowered to 2 mCTC for this analysis, as (i) we showed that we only found up to one mCTC in healthy controls [10, 14], and (ii) to include the maximum number of patients of whom multiple time points were available. It can be seen that patients who stay unfavorable at first follow-up may improve at second follow-up, which is reflected in their survival chances. Patients who have a rise in number of mCTC with at least 95% confidence deteriorate rapidly.

The therapeutic regimen of the majority of these patients included chemotherapy and the rate of CTC decline may be dependent on the type of therapy. Figure 7.2, panel B shows that also for those patients with <5 mCTC at baseline a strong upward trend is an indicator for poor prognosis, even though the majority of patients with 6–12 month survival have fewer than the clinically used cutoff of 5 mCTC after 12 weeks. It is unknown whether these patients have a subtype of cancer which sheds fewer mCTC into the blood, or whether they rapidly deteriorated after the last follow-up measurement.

To determine the best definition of a change in CTC count, several reduction criteria were tested. Four criteria were used to measure a reduction using modeled confidence of reduction, static, absolute and proportional cut-offs. Next, using our automated classifiers [15], different definitions of what constitutes an aCTC were applied, each having significant impact on survival of patients as measured by HRs at baseline, shown in table 7.2. The different definitions ranged from strict to loose, where the strictest definition excluded most false positives and part of the true positives, while

the most inclusive definition included most true positives and some false positives. This allowed us to explore the hypothesis that using more objects from a patient improves the ability to measure a reduction in aCTC count. This hypothesis was generated after the observation that tumor micro particles (EpCAM+CK+CD45-, $<4 \mu\text{m}$ diameter), present at a 20 fold higher frequency as compared to mCTC, are equally prognostic as the CellSearch manually counted mCTC [21]. These different CTC definitions were combined with different criteria to measure CTC reduction. For patients who started with five or more mCTC, a count below a static cut-off after 6 weeks of therapy remains the best indicator of treatment success for all CTC definitions. Combined with figure 7.1, we can again conclude that treatment should focus on getting the number of CTC to zero.

Table 7.3 shows that HRs were comparable for all definitions, indicating that the reduction in Poisson sampling noise with looser CTC definitions is roughly offset by an increase in background due to inclusion of more false positives. The aCTC D definition encompasses tumor fragments; the presence of similar particles in healthy controls however hampers their use as a surrogate for detection of the presence of CTC. Reduction of the background of aCTC D may improve the ability to use these tumor fragments as a means to measure the presence of CTC. This however will require alteration of the assay, such as addition of an extra fluorescent marker, which may result in such reduction of the numbers that it defeats the purpose. Improving the yield of CTC with the CellSearch will not significantly increase the numbers as $\sim 80\%$ of tumor cells spiked in blood are already recovered by the system [10]. Chasing CTC phenotypes that are currently not detected by the CellSearch system is likely to identify subgroups of patients in whom no CTC are detected, but will not significantly increase the number of CTC in patients, in whom few CTC are detected. Moreover, clinical studies will have to be conducted to demonstrate the prognostic and predictive value of CTC that do not express EpCAM and cytokeratins 8, 18 or 19.

The approach that truly will reduce the Poisson error is to sample a larger blood volume. Extrapolation of the CTC frequency in patients with metastatic breast and colon cancer showed that by increasing the blood volume to 5 liters CTC will be detected in all patients [13]. This however can only be achieved through in-vivo measurements [22, 23] or apheresis-based systems [24]. Using the current blood volume of 7.5 ml, the automated aCTC A definition is the preferred method of measuring if a patients CTC number approaches zero, because it eliminates reviewer variability. When reviewers assess two subsequent measurements, their variability in determining a change will propagate by the square root of the sum of the variability of the individual measurements squared.

We conclude that (i) CTC are best measured by means of an automated CTC count with a low background (aCTC A), (ii) aim of treatment should be the eliminating of all CTC. (iii) while reduction to 0 CTC can be seen

after 4–6 weeks, to reach this aim 10–12 weeks of therapy may be needed for some patients, and (iv) if the number does not decrease within this time span, treatment is not effective.

7.5 REFERENCES

- [1] C. Rao, T. Bui, M. Connelly, G. Doyle, I. Karydis, M. R. Middleton, G. Clack, M. Malone, F. A. W. Coumans, and L. W. M. M. Terstappen, “Circulating melanoma cells and survival in metastatic melanoma,” *International Journal of Oncology*, vol. 38, no. 3, pp. 755–760, 2011.
- [2] M. Cristofanilli, G. T. Budd, M. J. Ellis, A. Stopeck, J. Matera, M. C. Miller, J. M. Reuben, G. V. Doyle, W. J. Allard, L. Terstappen, and D. F. Hayes, “Circulating tumor cells, disease progression, and survival in metastatic breast cancer,” *New England Journal of Medicine*, vol. 351, no. 8, pp. 781–791, 2004.
- [3] S. J. Cohen, C. J. A. Punt, N. Iannotti, B. H. Saidman, K. D. Sabbath, N. Y. Gabrail, J. Picus, M. Morse, E. Mitchell, M. C. Miller, G. V. Doyle, H. Tissing, L. Terstappen, and N. J. Meropol, “Relationship of circulating tumor cells to tumor response, progression-free survival, and overall survival in patients with metastatic colorectal cancer,” *Journal of Clinical Oncology*, vol. 26, no. 19, pp. 3213–3221, 2008.
- [4] J. S. de Bono, H. I. Scher, R. B. Montgomery, C. Parker, M. C. Miller, H. Tissing, G. V. Doyle, L. Terstappen, K. J. Pienta, and D. Raghavan, “Circulating tumor cells predict survival benefit from treatment in metastatic castration-resistant prostate cancer,” *Clinical Cancer Research*, vol. 14, no. 19, pp. 6302–6309, 2008.
- [5] S. Matsusaka, K. Chin, M. Ogura, M. Suenaga, E. Shinozaki, Y. Mishima, Y. Terui, N. Mizunuma, and K. Hatake, “Circulating tumor cells as a surrogate marker for determining response to chemotherapy in patients with advanced gastric cancer,” *Cancer Science*, vol. 101, no. 4, pp. 1067–1071, 2010.
- [6] M. G. Krebs, R. Sloane, L. Priest, L. Lancashire, J. M. Hou, A. Greystoke, T. H. Ward, R. Ferraldeschi, A. Hughes, G. Clack, M. Ranson, C. Dive, and F. H. Blackhall, “Evaluation and prognostic significance of circulating tumor cells in patients with non-small-cell lung cancer,” *Journal of Clinical Oncology*, vol. 29, no. 12, pp. 1556–1563, 2011.
- [7] D. F. Hayes, M. Cristofanilli, G. T. Budd, M. J. Ellis, A. Stopeck, M. C. Miller, J. Matera, W. J. Allard, G. V. Doyle, and L. Terstappen, “Circulating tumor cells at each follow-up time point during therapy of metastatic breast cancer patients predict progression-free and overall survival,” *Clinical Cancer Research*, vol. 12, no. 14, pp. 4218–4224, 2006.
- [8] G. T. Budd, M. Cristofanilli, M. J. Ellis, A. Stopeck, E. Borden, M. C. Miller, J. Matera, M. Repollet, G. V. Doyle, L. Terstappen, and D. F. Hayes, “Circulating tumor cells versus imaging - predicting overall survival in metastatic breast cancer,” *Clinical Cancer Research*, vol. 12, no. 21, pp. 6403–6409, 2006.
- [9] U. De Giorgi, V. Valero, E. Rohren, M. Mego, G. V. Doyle, M. C. Miller, N. T. Ueno, B. C. Handy, J. M. Reuben, H. A. Macapinlac, G. N. Hortobagyi, and M. Cristofanilli, “Circulating tumor cells and bone metastases as detected by fdg-pet/ct in patients with metastatic breast cancer,” *Annals of Oncology*, vol. 21, no. 1, pp. 33–39, 2010.
- [10] W. J. Allard, J. Matera, M. C. Miller, M. Repollet, M. C. Connelly, C. Rao, A. G. J. Tibbe, J. W. Uhr, and L. Terstappen, “Tumor cells circulate in the peripheral blood

- of all major carcinomas but not in healthy subjects or patients with nonmalignant diseases,” *Clinical Cancer Research*, vol. 10, no. 20, pp. 6897–6904, 2004.
- [11] H. I. Scher, X. Y. Jia, J. S. de Bono, M. Fleisher, K. J. Pienta, D. Raghavan, and G. Heller, “Circulating tumour cells as prognostic markers in progressive, castration-resistant prostate cancer: a reanalysis of immc38 trial data,” *Lancet Oncology*, vol. 10, no. 3, pp. 233–239, 2009.
- [12] A. G. J. Tibbe, M. C. Miller, and L. Terstappen, “Statistical considerations for enumeration of circulating tumor cells,” *Cytometry Part A*, vol. 71A, no. 3, pp. 154–162, 2007.
- [13] F. Coumans, S. T. Ligthart, and L. W. M. Terstappen, “All patients with metastatic carcinoma have circulating tumor cells,” *submitted*, 2012.
- [14] M. R. de Groot, H. M. Croonen, W. J. T. Mastboom, I. Vermes, A. G. J. Tibbe, H. Tissing, and L. W. M. M. Terstappen, “Circulating tumor cells (ctc) in newly diagnosed breast or colorectal cancer,” *Proc Ann Meet Am Soc Clin Onc*, vol. 25, no. 18s, 2007.
- [15] S. T. Ligthart, F. A. W. Coumans, G. Attard, A. Mulick Cassidy, J. S. De Bono, and L. W. M. M. Terstappen, “Unbiased and automated identification of a circulating tumour cell definition that associates with overall survival,” *Plos One*, 2011.
- [16] S. Riethdorf, H. Fritsche, V. Muller, T. Rau, C. Schindibeck, B. Rack, W. Janni, C. Coith, K. Beck, F. Janicke, S. Jackson, T. Gornet, M. Cristofanilli, and K. Pantel, “Detection of circulating tumor cells in peripheral blood of patients with metastatic breast cancer: A validation study of the cellsearch system,” *Clinical Cancer Research*, vol. 13, no. 3, pp. 920–928, 2007.
- [17] H. Yagata, S. Nakamura, M. Toi, H. Bando, S. Ohno, and A. Kataoka, “Evaluation of circulating tumor cells in patients with breast cancer: multi-institutional clinical trial in japan,” *International Journal of Clinical Oncology*, vol. 13, no. 3, pp. 252–256, 2008.
- [18] F. C. Bidard, C. Mathiot, A. Degeorges, M. C. Etienne-Grimaldi, R. Delva, X. Pivot, C. Veyret, L. Bergougnoux, P. de Cremoux, G. Milano, and J. Y. Pierga, “Clinical value of circulating endothelial cells and circulating tumor cells in metastatic breast cancer patients treated first line with bevacizumab and chemotherapy,” *Annals of Oncology*, vol. 21, no. 9, pp. 1765–1771, 2010.
- [19] D. R. Shaffer, M. A. Leversha, D. C. Danila, O. Lin, R. Gonzalez-Espinoza, B. Gu, A. Anand, K. Smith, P. Maslak, G. V. Doyle, L. Terstappen, H. Lilja, G. Heller, M. Fleisher, and H. I. Scher, “Circulating tumor cell analysis in patients with progressive castration-resistant prostate cancer,” *Clinical Cancer Research*, vol. 13, no. 7, pp. 2023–2029, 2007.
- [20] M. C. Liu, P. G. Shields, R. D. Warren, P. Cohen, M. Wilkinson, Y. L. Ottaviano, S. B. Rao, J. Eng-Wong, F. Seillier-Moiseiwitsch, A. M. Noone, and C. Isaacs, “Circulating tumor cells: A useful predictor of treatment efficacy in metastatic breast cancer,” *Journal of Clinical Oncology*, vol. 27, no. 31, pp. 5153–5159, 2009.
- [21] F. A. W. Coumans, C. J. M. Doggen, G. Attard, J. S. de Bono, and L. W. M. M. Terstappen, “All circulating epcam+ck+cd45- objects predict overall survival in castration-resistant prostate cancer,” *Annals of Oncology*, vol. 21, no. 9, pp. 1851–7, 2010.
- [22] Y. C. Chang, J. Y. Ye, T. P. Thomas, Z. Y. Cao, A. Kotlyar, E. R. Tkaczyk, J. R. Baker, and T. B. Norris, “Fiber-optic multiphoton flow cytometry in whole blood and in vivo,” *Journal of Biomedical Optics*, vol. 15, no. 4, 2010.

- [23] E. I. Galanzha, E. V. Shashkov, T. Kelly, J. W. Kim, L. L. Yang, and V. P. Zharov, “In vivo magnetic enrichment and multiplex photoacoustic detection of circulating tumour cells,” *Nature Nanotechnology*, vol. 4, no. 12, pp. 855–860, 2009.
- [24] R. L. Eifler, J. Lind, D. Falkenhagen, V. Weber, M. B. Fischer, and R. Zeillinger, “Enrichment of circulating tumor cells from a large blood volume using leukapheresis and elutriation: Proof of concept,” *Cytometry Part B-Clinical Cytometry*, vol. 80B, no. 2, pp. 100–111, 2011.



UNBIASED QUANTITATIVE
ASSESSMENT OF HER-2
EXPRESSION OF CIRCULATING
TUMOR CELLS IN PATIENTS WITH
METASTATIC AND
NON-METASTATIC BREAST
CANCER¹

Sjoerd T. Ligthart, François-Clément Bidard, Charles Decraene, Thomas Bachelot, Suzette Delaloge, Etienne Brain, Mario Campone, Patrice Viens, Jean-Yves Pierga, and Leon W.M.M. Terstappen.

ABSTRACT

Circulating tumor cells (CTC) can provide the basis for a real-time liquid biopsy. Assessment of biomarkers such as Her-2 status of CTC detected in breast cancer patients may guide the use of targeted therapies. However, reported techniques are either observer-dependent and/or unfit for large scale screening of patients. We report on unbiased quantification of Her-2 protein expression of CTC in 103 metastatic (M1) and 88 non-metastatic (M0) breast cancer patients. Digital images recorded by the CellSearch[®] system were processed by an automated algorithm, which determined a Her-2 positivity threshold in each sample using the Her-2-fluoroisothiocyanate

¹Submitted.

fluorescence of leukocytes within each sample as internal control. This automated method distinguished Her-2 positive from Her-2 negative breast cancer cell lines. Her-2 expression of CTC varied greatly between and within patients, suggesting a possible mechanism of treatment resistance. However, in M1 patients, a 75% threshold of Her-2 positive CTC in patients with ≥ 5 CTC showed a relatively low discrepancy rate between the primary tumor and CTC Her-2 status. Applying this threshold, 2% (95%CI:[0-11%]) of M1 patients with Her-2 negative primary tumors had Her-2 positive CTC status and 29% (95%CI:[12-55%]) of M1 patients with Her-2 positive primary tumors had Her-2 negative CTC status. Interestingly, no Her-2 discrepancy was observed between CTC and primary tumor in M0 patients with ≥ 5 CTC, suggesting that Her-2 status shift could appear at metastatic relapse or treatment resistance. These findings demonstrate the feasibility of real-time quantitative and reproducible assessment of treatment targets on CTC, opening a path towards personalized treatment.

8.1 INTRODUCTION

Treatment of metastatic breast cancer took a leap forwards with the introduction of trastuzumab, which targets the human epidermal growth-factor-receptor type 2 (Her-2) gene [1, 2]. Her-2 is amplified and over-expressed in $\sim 15\%$ of invasive breast cancers, and is usually assessed using immunohistochemistry and/or in situ hybridization on tissue obtained from the primary tumor, following current guidelines [3]. At the metastatic stages of this disease, Her-2 status is considered to be globally stable, but numerous reports have shown discrepancies in the Her-2 status of patients between the primary tumor and the following metastatic relapses [4, 5]. It is now considered relevant to perform a tissue biopsy at a first metastatic relapse to reassess the therapeutic targets expressed by the tumor. However, obtaining tissue biopsies from metastatic sites is not always feasible, and is associated with discomfort and risk to the patient.

Detection and characterization of tumor cells, circulating in the blood at the time of treatment, may resolve this issue. Several studies have now established that the presence of circulating tumor cells (CTC) in metastatic (M1) breast-cancer patients is associated with a poor outcome, and that their perseverance after the first cycles of therapy indicate an even worse prognosis [6, 7, 8, 9, 10]. Beyond the quantitative analysis of CTC in blood, the second interest of CTC detection is that molecular targets can be assessed on CTC [11, 12, 13, 14]. However, classification of CTC and quantification of treatment targets, such as Her-2, for CTC can be subjective as they are morphologically heterogeneous [15, 16].

At present, expression of Her-2 on CTC using the clinically validated CellSearch[®] system for CTC enumeration relies on visual inspection of auto-scaled Her-2-fluorescein isothiocyanate (FITC) images from CTC [17, 18]. However, this method is poorly reproducible and the auto-scaled images can mislead the human reviewer. To overcome these challenges, we automated

the classification of CTC and the quantification of their Her-2-expression levels.

8.2 MATERIALS AND METHODS

8.2.1 PATIENTS

Patients with M1 breast cancer were part of the multicenter prospective IC 2006-04 study (NCT00898014). From 06/2007 to 09/2009, 267 patients were included in the IC 2006-04 study [10]. In the first period of patient recruitment, only patients with a known primary Her-2-positive status were screened for Her-2 expression on CTC. However, starting from 04/2009, all patients, irrespective of primary-tumor status, were screened. Among the 267 patients included at baseline, 174 patients had ≥ 1 CTC detected by manual review (mCTC) in 7.5 ml of blood; of these, 103 patients were further screened for Her-2 expression on detected CTC and were part of this analysis.

Patients with non metastatic (M0) breast cancer were part of two multicenter prospective phase II sister studies assessing the neoadjuvant treatment efficacy on inflammatory (T4dNxM0) breast cancer either Her-2 negative (BEVERLY 01, NCT00820547) or positive (BEVERLY 02, NCT00717405). From 12/2008 to 09/2010, 101 patients with Her-2-negative inflammatory M0 breast cancer were included in the BEVERLY 01 study, and the systematic screening of Her-2 expression on CTC at baseline started in 12/2009 [19]. 38 patients were evaluable at baseline for both CTC detection and Her-2 expression on CTC, and were part of this analysis. From 10/2008 to 10/2009, 52 patients with Her-2-positive inflammatory M0 breast cancer were included in the BEVERLY 02 study; of these, 50 patients were evaluable at baseline for both CTC detection and Her-2 expression on CTC [20], and were part of this analysis. The IC 2006-04, BEVERLY 01 & 02 studies were approved by the national ethics board and all patients provided written informed consent.

8.2.2 HER-2 ASSESSMENT OF THE PRIMARY TUMOR

Patients had the Her-2 status of their primary tumor determined using immunohistochemistry, at each cancer center, according to the current recommendations [3]. Her-2 intensity was scored as being null (0+), weak (1+), intermediate (2+), or strong (3+). Tissues that were 0+ and 1+ were considered as Her-2-negative, and 3+ as Her-2-positive. In doubtful cases (2+), FISH was performed to determine Her-2-amplification status (Her-2/CEP17 ratio > 2.2) [3]. Her-2 status of the primary tumor was mandatory in BEVERLY studies. In a few patients included in the IC 2006-04 study, Her-2 status remained unknown due to lack of tumor material.

8.2.3 MANUAL CTC ENUMERATION AND HER-2 ASSESSMENT

The CellSearch system (Veridex LLC, Raritan, NJ) was used to enumerate CTC [10, 21]. The system consists of a CellTracks Autoprep[®] for sample preparation, and a CellTracks Analyzer II[®] for sample analysis. The CellTracks Autoprep immuno-magnetically enriches epithelial cells from 7.5 ml of blood using ferrofluids conjugated to epithelial cell-adhesion molecule antibodies (EpCAM). The enriched sample is stained with phycoerythrin-conjugated (PE) antibodies C11 and A53-B/A2 directed against cytokeratins (CK) 8, 18, and 19 respectively, an allophycocyanin-conjugated (APC) antibody HI30 to CD45, and nuclear dye 4',6-diamidino-2-phenylindole (DAPI). FITC-conjugated antibody HER81 (Veridex LLC [11]) was added to the staining cocktail to identify Her-2. The enriched and stained samples were transferred to a cartridge, which was placed in between two magnets that distributed the cells over the analysis surface. The cartridge was placed in the CellTracks Analyzer II; this four-color semi-automated fluorescence microscope captured and stored 8-bit digital images of the four different fluorescent dyes. The microscope was equipped with a 10 \times /0.45NA objective and a CCD camera, using 6.7 \times 6.7 μ m pixels.

To ascertain the number of CTC in a sample, objects that were positive for DNA and CK were selected by the system and were shown, in a gallery, to a trained reviewer. The reviewer marked DNA+, CK+, and CD45- objects that were larger than 4 \times 4 μ m, and had a cell-like morphology, as manual CTC (mCTC). A review of the thumbnails of the Her-2 staining of CTC was used to assign the mCTC as either Her-2-negative or Her-2-positive. This was done by visually comparing images obtained with spiked cell lines of known Her-2 status, as already described by other groups [17, 18]. All images from the samples were saved to a CD or DVD. Six breast cancer cell lines were used: SKBR3, BT474, MDA-MB361, ZR75-1, BT20, and MCF7. Early passage cells were used from these cell lines that were originally obtained from the American Type Culture collection (Manassas, VA, USA). 105 cells of each cell lines were cytopun onto slides and immunostaining was performed with the Her-2 specific antibody CB11 to confirm their Her-2 immunocytostaining. SKBR3, BT474, MDA-MB361, ZR75-1, BT20, and MCF7 were scored as 3+, 3+, 2+, 1+, 0+ and 0+ respectively, by senior cytologist Dr J. Klijanienko at the Institut Curie (images not shown). These cell lines were spiked into blood from healthy donors for the Cellsearch analysis.

8.2.4 AUTOMATED CTC ENUMERATION AND HER-2 ASSESSMENT

The data from the CDs and DVDs that contained the archived images of the 191 samples were copied onto a central hard drive. Objects were detected and classified using an automated algorithm developed in Matlab 2009a (Mathworks, Natick, MA), using the DIPimage toolbox (Delft University

of Technology, The Netherlands, <http://www.diplib.org>). This algorithm was optimized using survival times from prostate-cancer patients: details of this algorithm have been reported previously [22]. In short, 140–180 four-channel tiff images of a sample from each patient were uploaded onto the computer. First, detection of the border of the sample cartridge was performed via edge detection in the FITC images. The true imaging area, where all the objects were located, could then be determined.

Next, a dynamic threshold was determined using the CK-PE channel-image histogram of the selected imaging area. This threshold method determines one threshold value for the whole cartridge. Applying this threshold to the CK-PE images gave the outline, size, and location of objects. In the next step, the standard deviation of the CK-PE channel and the peak values of both the DNA-DAPI and CD45-APC channels were measured on every object using locations and outlines revealed by the thresholding procedure. Finally, classification of every object was performed and the objects were counted as automated CTC (aCTC), where: (i) the CK-PE standard deviation was >50 counts; (ii) it was sized between 75 and 2000 pixels ($34\text{--}898\ \mu\text{m}^2$); (iii) the DNA-DAPI peak value was >170 counts, and (iv) CD45-APC peak was <60 counts. The criterion for aCTC size was raised from 500 to 2000 pixels to accommodate for the larger size of CTC from breast-cancer patients compared to CTC from prostate-cancer patients.

For every patient's sample, the objects assigned as aCTC were totaled to give a final aCTC count per patient. The image-analysis algorithm does not use the FITC channel for aCTC recognition, and thus permitted determination of Her-2-FITC expression on the outline covering the area where the aCTC were present. The mean value, summed with two times the standard deviation of the value of the Her-2-FITC channel within the outline of each CTC, was measured to quantify Her-2 expression. Using this combination of two measures, cells with an irregular or speckled staining pattern, and a low mean value, could be included as positive Her-2 aCTC because the standard deviations of their staining intensity would be high. The coefficient of variation (CV) of Her-2 FITC fluorescence intensity was defined as the standard deviation divided by the mean and was determined for cell lines and patient samples that had equal or more than 5 aCTC.

8.2.5 DETERMINATION OF A THRESHOLD FOR AUTOMATED HER-2 ASSESSMENT

To arrive at a threshold for Her-2-FITC staining within each sample, Her-2-FITC staining of leukocytes present in the enriched sample was used as an internal control. To identify leukocytes in the sample, thresholding of the CD45-APC channel was performed to select the outline of the leukocytes. A leukocyte classifier was created that included objects with a maximum value, in CK-PE, of 100 counts, with a size range between 50 and 500

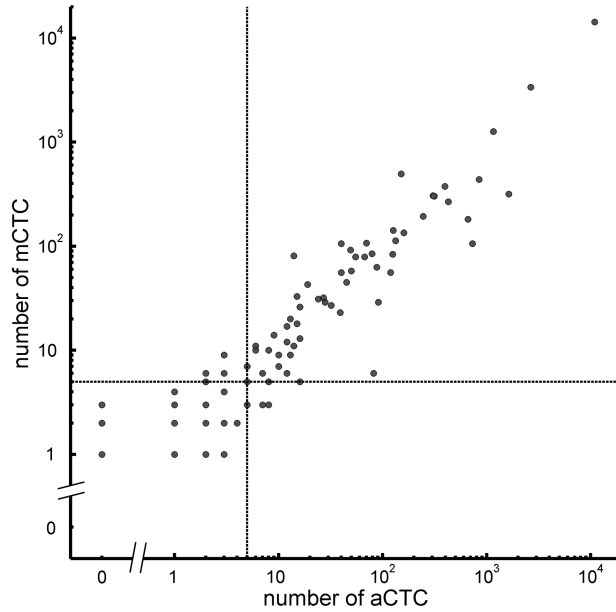


FIGURE 8.1 : Scatter plot of measured aCTC versus manually scored mCTC in 103 samples of M1 patients (IC 2006-04 study). The dashed lines represent the clinical threshold of ≥ 5 CTC.

pixels ($22\text{--}224\ \mu\text{m}^2$), a DNA-DAPI peak value of at least 100 counts, and a CD45-APC peak value of at least 100 counts. The distribution of Her-2-FITC signals in leukocytes within each sample was used to determine an appropriate threshold percentile for Her-2-staining of aCTC.

8.3 RESULTS

8.3.1 IDENTIFICATION OF CTC AND LEUKOCYTES IN M1 BREAST-CANCER PATIENTS

Leukocytes and aCTC from 103 patients with M1 breast-cancer who had at least one mCTC detected and the Her-2 CTC staining assessed were identified in the stored images generated by the CellSearchTM system, before the start of first-line chemotherapy. The number of leukocytes identified in the 103 samples ranged from 38–12541, with a mean of 1298, a median of 371, and a SD of 2432. The number of aCTC identified by the algorithm ranged from 0–10902, with a median of 8, a mean of 217, and a SD of 1178. In comparison, manual reviews (performed using CellSearch instructions) resulted in 1–14274 mCTC, with a median of 9, a mean of 232, and a SD of 1443. Figure 8.1 shows the correlation between the number of aCTC

versus the number of mCTC in the samples from the 103 patients. The horizontal and vertical lines were drawn at the clinical cut-off numbers of five aCTC and five mCTC, to create four quadrants. The R^2 between aCTC and mCTC counts was 0.978 (slope = 1.28, intercept = -44.85). The upper left and lower right quadrants, in Figure 8.1, show that seven patients (7%) had discordant results between manual and automated counts, according to the clinically defined ≥ 5 CTC cut-off point. Note that samples with zero mCTC where CTC could have been potentially detected by the automated algorithm, were not part of this analysis.

8.3.2 HER-2 EXPRESSION BREAST-CANCER CELL LINES

Cells from the six breast cancer cell lines SKBR3, BT474, MDA-MB361, ZR75-1, MCF7, and BT20 were each spiked in blood samples from healthy donors and processed with the CellSearch system using Her-2-FITC as an additional reagent. Tumor cells were identified as aCTC in the stored images generated by the CellSearch system. The median and 95th percentile of the Her-2 signals in each of the cell lines are shown in Panel A of Figure 8.2. As expected, the Her-2(3+) SKBR3 and BT474 cells showed the highest Her-2 levels, followed by the Her-2(2+) MDA-MB361 cells; Her-2 was expressed lower in Her-2(0 or 1+) ZR75-1, MCF7, and BT20 cells.

8.3.3 HER-2 STAINING OF LEUKOCYTES AS INTERNAL CONTROL

To arrive at a threshold for Her-2 expression in aCTC, the Her-2 expression of leukocytes within each sample was investigated. Her-2 intensity, in the areas where leukocytes were present, was determined for all patients. Figure 8.3 shows a histogram of the mean intensity, plus two times the standard deviation of the Her-2 signal, for 6379 leukocytes detected in one sample. The distribution of Her-2 signals in leukocytes suggests the presence of two populations. The population with the lowest fluorescence can likely be contributed to the autofluorescence of the lymphocytes carried over through the enrichment procedure. The other population either non-specifically bound to the Her-2 FITC antibody or exhibited a higher autofluorescence, and can likely be contributed to granulocytes. We assumed that both populations were quasi-normally distributed (by means of a Poisson-like distribution), as illustrated by the solid and dashed black lines in figure 8.3.

It was determined that $\sim 9\%$ of leukocytes had the higher fluorescence signals. The Her-2-FITC threshold was set at 99.9% of the low intensity leukocyte population (vertical dashed grey line in figure 8.3), which was the 91st percentile of the total distribution of leukocytes from each separate sample. The R^2 of this Gaussian fit of two distributions was 0.982. In cases where the leukocyte count was low, this 91st percentile can have a confidence interval that is excessively large. By means of a simulation in Matlab, it was established that, in order to generate a 95% confidence interval from 0.82-1,

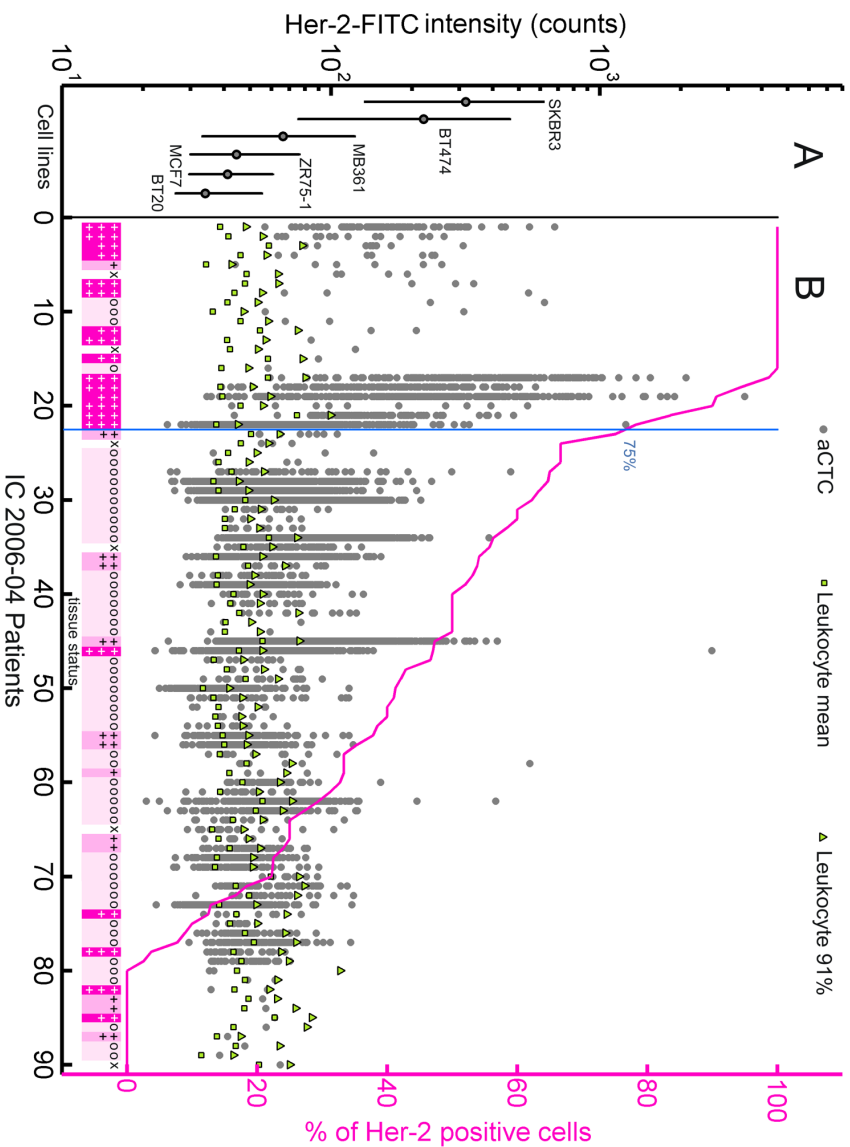


FIGURE 8.2 : Panel A shows median Her-2 expression with the 95th percentile of SKBR3(3+), BT474(3+), MIB361(2+), ZR75-1(1+), MCF7(0+) and BT20(0+) cell lines. Panel B shows a scatter plot of aCTC and median leukocyte Her-2-FITC signal intensities, together with the percentage of aCTC that were above 91% of the baseline leukocyte Her-2 expression levels for 90 MI patients that had at least one aCTC (pink line). At the bottom of the figure, the primary tumor-tissue status of each patient is plotted (bright pink = Her-2-positive (3+ or 2+/FISH+), other colors = Her-2-negative (0+, 1+ or 2+/FISH-), x=unknown).

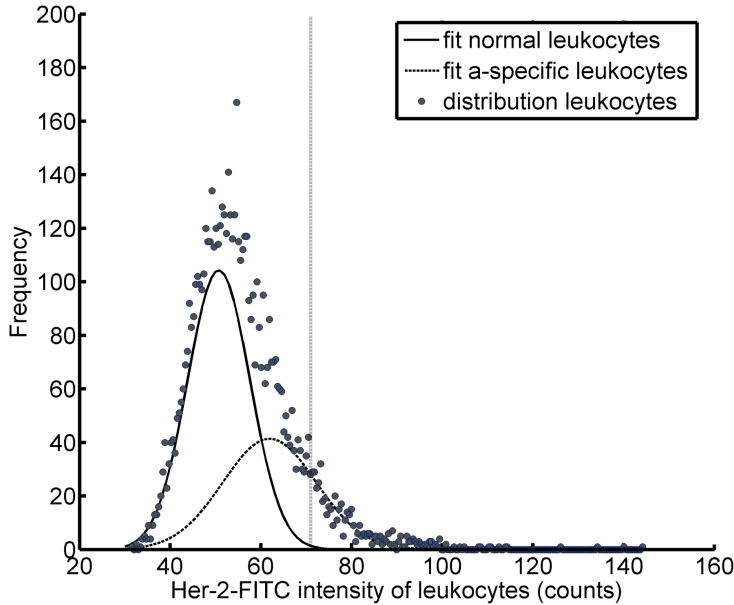


FIGURE 8.3 : Histogram of Her-2-FITC signals from 6,379 leukocytes detected in a patient sample. Two populations of leukocytes are shown: one with no fluorescence in the FITC channel (solid black line) and the other with a low a-specific fluorescence (dashed black line). The vertical dashed grey line indicates the 99.9% level of the leukocyte population with no fluorescence.

around the 91st percentile, at least 75 leukocytes needed to be included in each sample to define a robust internal Her-2-positivity threshold. Among the 191 samples, only 2 samples were below the threshold of 75 leukocytes and were inspected to assure the 91st percentile was set appropriately.

Median and 91st percentile of the Her-2-FITC intensity from the leukocyte populations in the samples of each of the 90 M1 breast cancer patients are displayed in panel B of figure 8.2. The median fluorescence intensity of these internal controls shows a 4.0 fold variation and highlights the importance of an internal control for assessment of Her-2 expression. Comparison of Her-2-FITC intensity of breast cancer cell lines showed that Her-2 signals from Her-2(3+)SKBR3 and BT474 cells were clearly above the background Her-2 staining of leukocytes; those from Her-2(2+) MDA-MB361 cells were just above the background level in most samples, whereas those from the Her-2(1+ or 0+) ZR75-1, MCF7, and BT20 cells could not be discriminated from the background Her-2 staining of leukocytes.

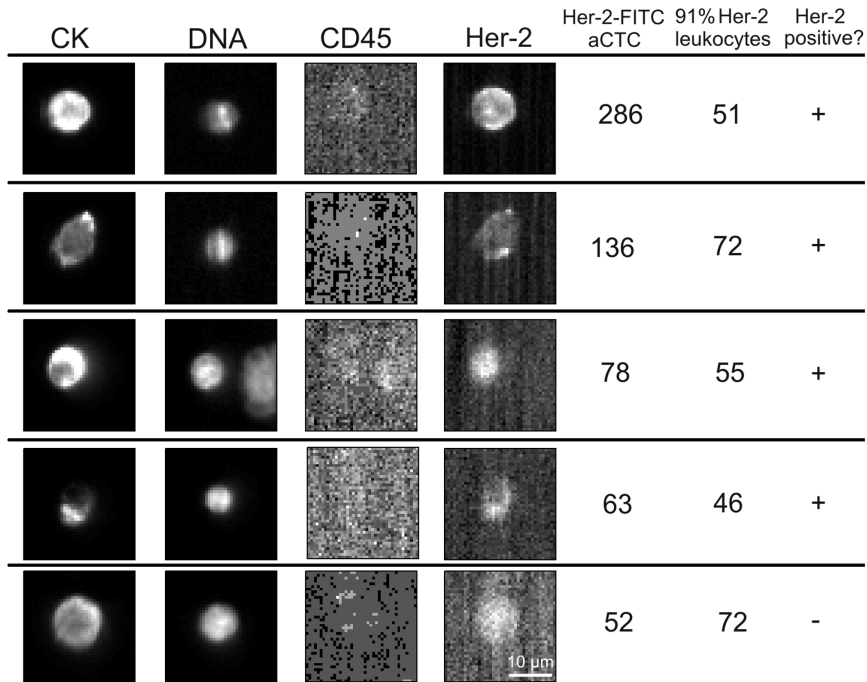


FIGURE 8.4 : Examples of CTC and the Her-2 expression of five CTC from five different patients. The last three columns show intensity values for aCTC Her-2, the 91% expression level of leukocytes within the same sample, and the final Her-2 status of the cell. The scale bar applies to all images.

8.3.4 HER-2 EXPRESSION OF CTC FROM M1 BREAST-CANCER PATIENTS

At least one aCTC was identified in 90 of the 103 IC 2006-04 patients, and Her-2 expression of aCTC in these patients is shown in panel B of figure 8.2. Her-2 expression of each aCTC is given on the left y-axis and the percentage of Her-2-positive aCTC (i.e. aCTC expressing Her-2 above the background staining of leukocytes) is on the right y-axis. Samples were sorted, firstly on the percentage of positive Her-2 aCTC, and secondly on the number of aCTC. The correlation between the number of Her-2-positive CTC retrieved in each sample by automated and manual assessments was good, with an R^2 between automated and manual Her-2 assessment of 0.979 (slope = 1.78, intercept = -30.64).

Figure 8.4 shows examples of the cytokeratin, DAPI, CD45 and Her-2 images of five aCTC from different samples, together with the relative fluorescence Her-2 intensity and the background Her-2 intensity of the leukocytes in the samples. Interestingly, the heterogeneity of Her-2 expression, reflected by the spread of Her-2 intensities among the different CTC re-

trieved in a sample of the same patient, was higher for clinical samples than for the spiked cell lines, as revealed by Her-2 FITC CV. Median CV for the cell lines samples was 0.07 (range 0.06-0.08) against a median CV of 0.29 for aCTC from patient samples (range 0.20-1.63). Leukocytes in these patient's samples had a median CV of 0.26 (range 0.15-2.29). When comparing these distributions, p-values <0.001 were found between aCTC and cell lines, and between leukocytes and cell lines (Mann-Whitney non-parametric U-test). The spread of Her-2 signal between leukocytes and aCTC was not significantly different ($p=0.12$).

8.3.5 PRIMARY TISSUE HER-2 EXPRESSION VERSUS CTC HER-2 EXPRESSION IN M1 PATIENTS

The primary tumor was Her-2 positive in 28 patients (27%), Her-2 negative in 69 patients (67%), and Her-2 status was unknown in six patients (6%). The Her-2 status of tissue from metastatic patients in whom aCTC were detected, is indicated at the bottom of figure 8.2B. This figure strongly suggests that a threshold of 75% of the aCTC expressing Her-2 above the background staining of leukocytes may accurately identify the Her-2 status of the primary tumor of the patients.

We further investigated whether the number of aCTC detected and/or the primary tumor Her-2 staining had an impact on this Her-2 status assessment. Figure 8.5 A and B show percentage of patients (y-axis), separated into three groups according to the Her-2 staining of the primary tumor versus the percentage (x-axis) of Her-2-positive aCTC. M1 patients with a low aCTC count (i.e. 1-4 aCTC, panel A) were plotted apart from those with high aCTC counts (≥ 5 aCTC, panel B). The figures show a clear relation between the percentage of Her-2-positive aCTC and the Her-2 status of the primary tumor. For patients with known Her-2-negative (0, 1+, or 2+/FISH-) tumors, Her-2-positive rates of aCTC declined steadily with increased percentage (threshold) of Her-2-positive aCTC.

Secondly, comparison between panel A and B shows that this observation was most clear in the subgroup of patients that had ≥ 5 aCTC. This result gives support to the concept that, when assessing the Her-2 status of patients using CTC, the largest possible number of CTC should be analyzed. Focusing on patients with ≥ 5 aCTC the threshold of $>75\%$ Her-2 positive aCTC resulted in 29% (95%CI=[12%-55%]) of M1 patients with Her-2 positive primary tumor and ≥ 5 aCTC being Her-2 negative according to their CTC. Inversely, 2% (95%CI=[0%-11%]) of M1 patients with Her-2 negative primary tumor and ≥ 5 aCTC were Her-2 positive according to their CTC.

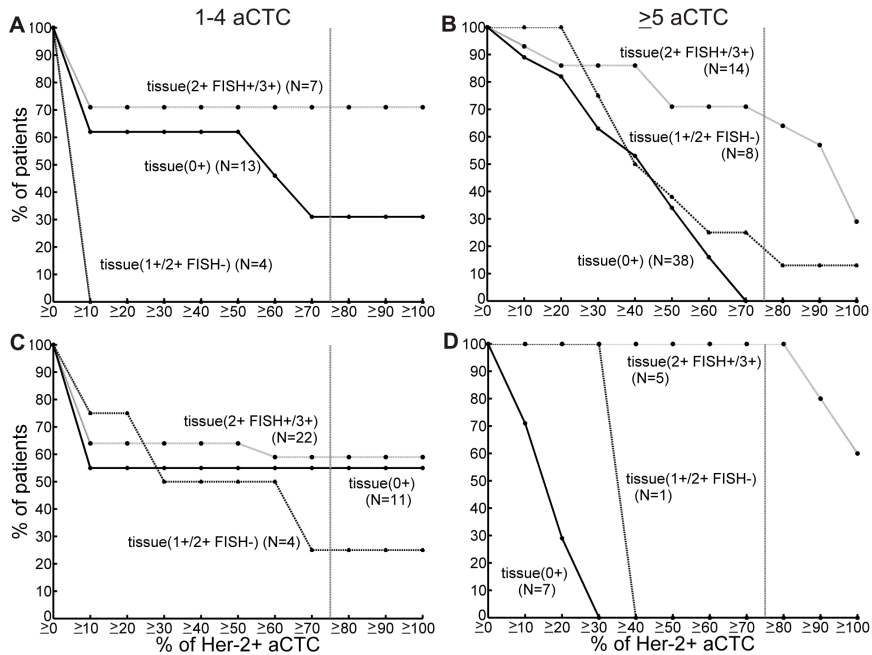


FIGURE 8.5 : Comparison of Her-2 expression in tissue and aCTC. Panel A and B show comparisons for M1 patients with 1-4 and ≥ 5 aCTC, respectively (IC 2006-04 study). Panel C and D show comparisons for M0 patients with 1-4 and ≥ 5 aCTC, respectively (BEVERLY 01 and 02 studies pooled). The dotted line indicates the threshold for patients with $>75\%$ Her-2-positive aCTC.

8.3.6 VALIDATION IN M0 BREAST CANCER PATIENTS

Biologically, it is expected that the discrepancy between the primary breast tumor and CTC is lower with CTC retrieved at M0 stage (i.e. CTC likely released by the primary tumor) than with CTC retrieved at M1 stage (CTC derived from metastatic masses made of highly selected tumor cells). Therefore, we analyzed the images of the 88 M0 patients included in BEVERLY 01 (Her-2-negative primary tumors) and 02 (Her-2-positive primary tumors) studies. In this inflammatory non metastatic (T4dM0) breast cancer population, 55% of patients had ≥ 1 and 15% ≥ 5 aCTC/7.5ml detected before the start of neoadjuvant chemotherapy. Correlation between mCTC and aCTC was good with a R^2 of 0.979 (slope = 1.41, intercept = -1.57).

Figure 8.6A and 8.6B show the aCTC count and the Her-2 fluorescence intensity for each aCTC detected in patients from BEVERLY 01 and 02, respectively: several patients in both studies seemed to have a discrepancy of CTC and primary tumor Her-2 status. However, Figure 8.5C and 8.5D shows that high discrepancy rates were observed only in patients with 1-4

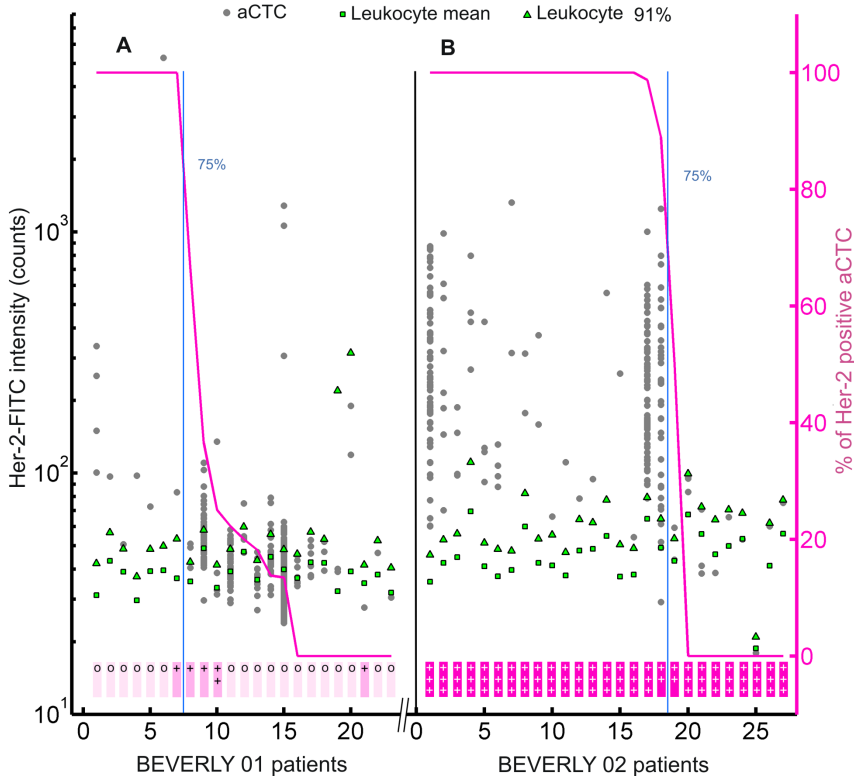


FIGURE 8.6 : Panel A shows a scatter plot of aCTC and median leukocyte Her-2-FITC signal intensities, together with the percentage of aCTC that were above 91% of the baseline leukocyte Her-2 expression levels for 23 M0 patients that had at least one aCTC (pink line) and negative Her-2 tissue status. At the bottom of the figure, the primary tumor-tissue status of each patient is plotted (bright pink = Her-2-positive (3+ or 2+/FISH+), other colors = Her-2-negative (0+, 1+ or 2+/FISH-)). Panel B shows 27 M0 patients that had at least one aCTC and positive Her-2 tissue status.

aCTC; no discrepancy between Her-2 status of the primary tumor and their corresponding aCTC was seen in M0 patients with ≥ 5 aCTC, giving support to the validity of our Her-2 assessment method.

8.3.7 PROCESSING TIME AND REPRODUCIBILITY OF THE ALGORITHM

The processing time required for automated identification and presentation of the images for manual reviewing were similar. However, mCTC counts and Her-2-intensity grading by a human operator typically took an additional 8 minutes per sample (median 5, range 1–39, SD 8 min). The variability of counting aCTC using the algorithm was 0%.

Determining tumor Her-2 status in patients with breast cancer is needed to provide tailored treatment. The ability to use CTC as a surrogate for a tissue biopsy would enable personalized treatment for these cancer patients. CTC are however not always detected in patients with metastatic breast cancer, and when present, Her-2 expression is heterogeneous and its determination is encumbered by inter- and intra-laboratory variations by human reviewers. Feasibility of Her-2 assessment by FISH on CTC has been reported [12, 13, 23], but it is much more expensive and labor intensive than immunocytostaining. The difficulties a reviewer faces when judging if a CTC is Her-2-positive by immunofluorescence include estimating Her-2 contrast, and determining whether or not it is above the staining intensity of, for example, a Her-2-expressing cell line. Furthermore, the reviewer has to judge whether or not there is variation in the intensity based on differences in Her-2 staining or sensitivity of the microscope. The latter can for example be influenced by the time the mercury arc lamp has been used.

Previous reports on Her-2 staining of CTC from breast-cancer patients have mostly relied on visual estimates of the “global level” of Her-2-FITC fluorescence from the “pool” of CTC detected in each patient [18]. Beyond the technical issues listed above, this approach implicitly prevents any further study on the clinical implications of Her-2 heterogeneity within the pool of CTC regarding their responses to anti-Her-2 therapies. A H-score-based approach has also been proposed to provide a weighted score, based on the proportion of CTC with a given level of fluorescence [24]. The H-score is routinely used to analyze the immunohistochemistry of tumor tissues, but it is based on the assumption that a large number of cancer cells (typically $>10^2$ – 10^3 cells) are assessed. This scoring system should therefore be used with care for CTC analysis, as visual estimates of proportions may be poor for most samples due to the low number of detected CTC.

A third method has been proposed, based on the signal/background-ratio calculation using image-analysis software, which may offer an objective estimate of Her-2 intensity for each CTC detected [17]. This method uses the low-resolution “capture screen” function of the computer to manually export each CTC picture from the CellSearch system to the image-analysis software, and is not suitable for samples that contain a large number of CTC. Moreover, this method is heavily impacted by the automatic scaling of the image intensity done by the CellSearch system, especially if a bright object is located near to the cell of interest.

Here, we have presented a new method to quantify Her-2 status of CTC in images generated by the CellSearch system, using an automated algorithm that was tested in a large cohort of patients. A computer program was used to classify CTC in the recorded CellSearch images. This CTC classifier was optimized using clinical outcome of images from prostate-cancer patients [22]. The outline of the CTC was used to determine the mean and standard

deviation of the Her-2-FITC signal for all CTC in every patient.

A good correlation was found between manual and automated CTC counts (Figure 8.1). The variability of counting aCTC using the algorithm was 0%, whereas a previous report by our group and others showed that inter-reader variability ranged from 4–31% for mCTC counts (median 14%) [16]. When evaluating Her-2-FITC staining, the variability of automated evaluation was also 0%. Although this variability has not yet been statistically compared to human readers, the variability is likely significantly higher for human readers. In our quantitative measurements of Her-2 on aCTC within individual patients, we observed great heterogeneity in Her-2 intensity levels. Interestingly, the spread of Her-2-intensity (CV) was much higher in patients than in cell lines, suggesting the presence of several sub-clones with different Her-2 expression levels in the patients' blood. Such heterogeneity is known in primary tumor, but has not been measured in CTC yet. It may play a role in treatment resistance, and is a further demonstration that cell lines do not recapitulate the heterogeneity of real cancer cells. It also indicates that, in order to derive from these CTC data a global Her-2 positive or Her-2 negative status, thresholds need to be set for both Her-2-intensity level and the proportion of Her-2-positive CTC.

First, a quantitative Her-2 threshold was set for CTC in each sample, based on Her-2 staining of the leukocytes carried over from the enrichment procedure in each sample, which acted as an internal control. Figure 8.4 illustrates the importance of an internal control: if a fixed-intensity value is set, then the CTC from row 4 could be counted as Her-2-negative because of its low intensity value, whereas Her-2-intensity of the CTC is actually well above the leukocyte threshold. In contrast, the CTC in row 5 could have been judged as Her-2-positive, whereas its Her-2 intensity is below that of the leukocytes in the sample. This threshold allowed determination of Her(2+) and Her-2(3+) cell lines as Her-2-positive, but not Her-2(0+) and Her-2(1+) cell lines.

Second, we set the percentage of Her-2-positive aCTC at >75% to arrive at a minimal discrepancy between Her-2 status of the primary tumor tissue and aCTC in the training cohort consisting of M1 breast cancer patients, before the start of first line chemotherapy. Interestingly, the global discrepancy rate decreased when taking into account patients with ≥ 5 aCTC detected. To our opinion, this result is not sustained by any biological phenomenon and directly suggests that Her-2 positivity percentage has to be assessed in a significant number of CTC (here, no other threshold than ≥ 5 CTC was tested). We finally conclude, with our method, that CTC and primary tumor have discordant Her-2 statuses in 29% (95%CI=[12%-55%]) of the M1 patients with ≥ 5 CTC count and Her-2 positive primary tumor and only 2% (95%CI=[0%-11%]) of the patients with ≥ 5 CTC count and Her-2 negative primary tumor.

Our aCTC Her-2 scoring method was then validated in M0 patients. At M0 stage, any detected CTC should derive directly from the primary

tumor: no major molecular discrepancy is expected between CTC and the primary tumor; at least these discrepancies should be less frequent than in M1 patients. In BEVERLY 01 and 02 studies, patients were considered non-metastatic (M0) after an exhaustive radiological screening but had inflammatory primary tumors (T4dNxM0), which correspond to rare but aggressive breast cancers. Our results showed high Her-2 discrepancy rates when all patients were analyzed, in the range of those reported in a previous publication by German groups in the neoadjuvant GEPARQUATTRO trial [18].

However, when looking at the M0 patients with ≥ 5 CTC count, we observed no discrepancy between CTC and primary tumor Her-2 status, in line with our previous hypothesis. This also suggests that discrepancy reports in M0 patients may be flawed by the low number of CTC assessable in this setting. These findings are in line with the seminal case-report study by Meng et al [13], who showed that Her-2 status change are mostly detectable in advanced M1 breast cancer patients and are possibly associated with or driven by treatment resistance. However, the number of patients with M0 disease and ≥ 5 CTC is limited ($<15\%$ even in case of T4d tumor).

Finally, this study, based on immunofluorescence analysis of a total of 191 M1 and M0 patients, provides a quantitative, reproducible, and fast method to determine Her-2 status of CTC. Our method eliminates intra- and inter-reviewer, as well as inter-laboratory variations, and enables standardization of Her-2 assessment in CTC. With the availability of such a reliable tool, further interventional studies -e.g. DETECT III trial [25]- need to be conducted to determine in M1 patients whether Her-2 status changes in CTC are clinically relevant when deciding to add (or remove) Her-2-targeted therapies, such as trastuzumab or lapatinib.

ACKNOWLEDGEMENTS

The authors wish to thank Dr Claire Mathiot, Isabelle Vaucher, and Mustapha Khazour, from the CTC laboratory at the Institut Curie, for their contributions. The IC 2006-04 study was promoted by the Institut Curie. The BEVERLY 01 study was promoted by UNICANCER / Fédération Française des Centres de Lutte Contre le Cancer (FNCLCC). The BEVERLY 02 study was promoted by Roche (Neuilly sur Seine, France).

8.5 REFERENCES

- [1] L. Coussens, T. L. Yangfeng, Y. C. Liao, E. Chen, A. Gray, J. McGrath, P. H. Seeburg, T. A. Libermann, J. Schlessinger, U. Francke, A. Levinson, and A. Ullrich, "Tyrosine kinase receptor with extensive homology to egf receptor shares chromosomal location with neu oncogene," *Science*, vol. 230, no. 4730, pp. 1132–1139, 1985.
- [2] D. J. Slamon, B. Leyland-Jones, S. Shak, H. Fuchs, V. Paton, A. Bajamonde, T. Fleming, W. Eiermann, J. Wolter, M. Pegram, J. Baselga, and L. Norton, "Use

- of chemotherapy plus a monoclonal antibody against her2 for metastatic breast cancer that overexpresses her2,” *New England Journal of Medicine*, vol. 344, no. 11, pp. 783–792, 2001.
- [3] A. C. Wolff, M. E. H. Hammond, J. N. Schwartz, K. L. Hagerty, D. C. Allred, R. J. Cote, M. Dowsett, P. L. Fitzgibbons, W. M. Hanna, A. Langer, L. M. McShane, S. Paik, M. D. Pegram, E. A. Perez, M. F. Press, A. Rhodes, C. Sturgeon, S. E. Taube, R. Tubbs, G. H. Vance, M. V. de Vijver, T. M. Wheeler, and D. F. Hayes, “American society of clinical oncology/college of american pathologists guideline recommendations for human epidermal growth factor receptor 2 testing in breast cancer,” *Journal of Clinical Oncology*, vol. 25, no. 1, pp. 118–145, 2007.
- [4] A. Fabi, A. Di Benedetto, G. Metro, L. Perracchio, C. Nistico, F. Di Filippo, C. Ercolani, G. Ferretti, E. Melucci, S. Buglioni, I. Sperduti, P. Papaldo, F. Cognetti, and M. Mottolese, “Her2 protein and gene variation between primary and metastatic breast cancer: Significance and impact on patient care,” *Clinical Cancer Research*, vol. 17, no. 7, pp. 2055–2064, 2011.
- [5] C. Simmons, N. Miller, W. Geddie, D. Gianfelice, M. Oldfield, G. Dranitsaris, and M. J. Clemons, “Does confirmatory tumor biopsy alter the management of breast cancer patients with distant metastases?,” *Annals of Oncology*, vol. 20, no. 9, pp. 1499–1504, 2009.
- [6] M. Cristofanilli, G. T. Budd, M. J. Ellis, A. Stopeck, J. Matera, M. C. Miller, J. M. Reuben, G. V. Doyle, W. J. Allard, L. Terstappen, and D. F. Hayes, “Circulating tumor cells, disease progression, and survival in metastatic breast cancer,” *New England Journal of Medicine*, vol. 351, no. 8, pp. 781–791, 2004.
- [7] M. C. Liu, P. G. Shields, R. D. Warren, P. Cohen, M. Wilkinson, Y. L. Ottaviano, S. B. Rao, J. Eng-Wong, F. Seillier-Moiseiwitsch, A. M. Noone, and C. Isaacs, “Circulating tumor cells: A useful predictor of treatment efficacy in metastatic breast cancer,” *Journal of Clinical Oncology*, vol. 27, no. 31, pp. 5153–5159, 2009.
- [8] S. Nakamura, H. Yagata, S. Ohno, H. Yamaguchi, H. Iwata, N. Tsunoda, Y. Ito, N. Tokudome, M. Toi, K. Kuroi, and E. Suzuki, “Multi-center study evaluating circulating tumor cells as a surrogate for response to treatment and overall survival in metastatic breast cancer,” *Breast Cancer*, vol. 17, no. 3, pp. 199–204, 2010.
- [9] F. Nole, E. Munzone, L. Zorzino, I. Minchella, M. Salvatici, E. Botteri, M. Medici, E. Verri, L. Adamoli, N. Rotmensch, A. Goldhirsch, and M. T. Sandri, “Variation of circulating tumor cell levels during treatment of metastatic breast cancer: prognostic and therapeutic implications,” *Annals of Oncology*, vol. 19, no. 5, pp. 891–897, 2008.
- [10] J. Y. Pierga, D. Hajage, T. Bachelot, S. Delaloge, E. Brain, M. Campone, V. Dieras, E. Rolland, L. Mignot, C. Mathiot, and F. C. Bidard, “High independent prognostic and predictive value of circulating tumor cells compared with serum tumor markers in a large prospective trial in first-line chemotherapy for metastatic breast cancer patients,” *Annals of Oncology*, 2011.
- [11] D. F. Hayes, T. M. Walker, B. Singh, E. S. Vitetta, J. W. Uhr, S. Gross, C. Rao, G. V. Doyle, and L. Terstappen, “Monitoring expression of her-2 on circulating epithelial cells in patients with advanced breast cancer,” *International Journal of Oncology*, vol. 21, no. 5, pp. 1111–1117, 2002.
- [12] S. Meng, D. Tripathy, S. Shete, R. Ashfaq, H. Saboorian, B. Haley, E. Frenkel, D. Euhus, M. Leitch, C. Osborne, E. Clifford, S. Perkins, P. Beitsch, A. Khan, L. Morrison, D. Herlyn, L. Terstappen, N. Lane, J. Wang, and J. Uhr, “upar and her-2 gene status in individual breast cancer cells from blood and tissues,” *Proceedings Of The National Academy Of Sciences Of The United States Of America*, vol. 103, no. 46, pp. 17361–17365, 2006.

- [13] S. D. Meng, D. Tripathy, S. Shete, R. Ashfaq, B. Haley, S. Perkins, P. Beitsch, A. Khan, D. Euhus, C. Osborne, E. Frenkel, S. Hoover, M. Leitch, E. Clifford, E. Vitetta, L. Morrison, D. Herlyn, L. Terstappen, T. Fleming, T. Fehm, T. Tucker, N. Lane, J. Q. Wang, and J. Uhr, "Her-2 gene amplification can be acquired as breast cancer progresses," *Proceedings Of The National Academy Of Sciences Of The United States Of America*, vol. 101, no. 25, pp. 9393–9398, 2004.
- [14] F. C. Bidard, C. Mathiot, A. Degeorges, M. C. Etienne-Grimaldi, R. Delva, X. Pivot, C. Veyret, L. Bergougnoux, P. de Cremoux, G. Milano, and J. Y. Pierga, "Clinical value of circulating endothelial cells and circulating tumor cells in metastatic breast cancer patients treated first line with bevacizumab and chemotherapy," *Annals of Oncology*, vol. 21, no. 9, pp. 1765–1771, 2010.
- [15] F. A. W. Coumans, C. J. M. Doggen, G. Attard, J. S. de Bono, and L. W. M. M. Terstappen, "All circulating epcam+ck+cd45- objects predict overall survival in castration-resistant prostate cancer," *Annals of Oncology*, vol. 21, no. 9, pp. 1851–7, 2010.
- [16] J. Kraan, S. Sleijfer, M. H. Strijbos, M. Ignatiadis, D. Peeters, J. Y. Pierga, F. Farace, S. Riethdorf, T. Fehm, L. Zorzino, A. G. Tibbe, M. Maestro, R. Gisbert-Criado, G. Denton, J. S. de Bono, C. Dive, J. A. Foekens, and J. W. Gratama, "External quality assurance of circulating tumor cell enumeration using the cellsearch((r)) system: A feasibility study," *Cytometry B Clin Cytom*, vol. 80B, no. 2, pp. 112–118, 2010.
- [17] M. Ignatiadis, F. Rothe, C. Chaboteaux, V. Durbecq, G. Rouas, C. Criscitiello, J. Metallo, N. Kheddoumi, S. K. Singhal, S. Michiels, I. Veys, J. Rossari, D. Lar-simont, B. Carly, M. Pestrin, S. Bessi, F. Buxant, F. Liebens, M. Piccart, and C. Sotiriou, "Her2-positive circulating tumor cells in breast cancer," *Plos One*, vol. 6, no. 1, 2011.
- [18] S. Riethdorf, V. Muller, L. L. Zhang, T. Rau, S. Loibl, M. Komor, M. Roller, J. Huober, T. Fehm, I. Schrader, J. Hilfrich, F. Holms, H. Tesch, H. Eidtmann, M. Untch, G. von Minckwitz, and K. Pantel, "Detection and her2 expression of circulating tumor cells: Prospective monitoring in breast cancer patients treated in the neoadjuvant gearquattro trial," *Clinical Cancer Research*, vol. 16, no. 9, pp. 2634–2645, 2010.
- [19] J. Y. Pierga, F. C. Bidard, F. Andre, T. Petit, F. Dalenc, T. Delozier, G. Romieu, J. Bonnetterre, J. Ferrero, P. Kerbrat, A. Martin, and P. Viens, "Early drop of circulating tumor cells (ctc) and increase of circulating endothelial cells (cec) during neoadjuvant chemotherapy (ct) combined with bevacizumab in her2-negative inflammatory breast cancer (ibc) in multicenter phase ii trial beverly 1," in *ASCO*, (Chicago, Illinois), p. Abstract no 10510, 2011.
- [20] J. Y. Pierga, F. C. Bidard, T. Petit, T. Delozier, J. Ferrero, M. Campone, J. Gligorov, F. Lerebours, H. Roche, S. Kraemer, C. Mathiot, and P. Viens, "Monitoring circulating tumor cells (ctc) and circulating endothelial cells (cec) during neoadjuvant combination of trastuzumab and bevacizumab with chemotherapy in her2 overexpressing inflammatory breast cancer (ibc): An ancillary study of beverly 2 multicenter phase ii trial," in *SABCS*, (San Antonio Texas), pp. Abstract no 70:PD04–07, 2010.
- [21] M. Kagan, D. Howard, T. Bendele, J. Mayes, J. Silvia, M. Repollet, J. Doyle, J. Al-lard, N. Tu, T. Bui, T. Russell, C. Rao, M. Hermann, H. Rutner, and L. Terstappen, "A sample preparation and analysis system for identification of circulating tumor cells," *Journal of Clinical Ligand Assay*, vol. 25, no. 1, pp. 104–110, 2002.
- [22] S. T. Ligthart, F. A. W. Coumans, G. Attard, A. Mulick Cassidy, J. S. De Bono, and L. W. M. M. Terstappen, "Unbiased and automated identification of a circulating tumour cell definition that associates with overall survival," *Plos One*, 2011.

- [23] J. F. Swennenhuis, A. G. Tibbe, R. Levink, R. C. Sipkema, and L. W. Terstappen, "Characterization of circulating tumor cells by fluorescence in situ hybridization," *Cytometry Part A*, vol. 75, no. 6, pp. 520–7, 2009.
- [24] E. A. Punnoose, S. K. Atwal, J. M. Spoerke, H. Savage, A. Pandita, R. F. Yeh, A. Pirzkall, B. M. Fine, L. C. Amler, D. S. Chen, and M. R. Lackner, "Molecular biomarker analyses using circulating tumor cells," *Plos One*, vol. 5, no. 9, 2010.
- [25] C. Melcher, W. Janni, U. Ortmann, B. Jager, B. Rack, V. Majller, A. Schneewiess, K. Pantel, E. Solomayer, and T. Fehm, "Detect iii - a multicenter, randomized, phase iii study to compare standard therapy alone versus standard therapy plus lapatinib in patients with initially her2-negative metastatic breast cancer and her2-positive circulating tumor cells," in *San Antonio Breast Cancer Symposium*, (San Antonio, TX), pp. OT1–02–06, 2011.

SUMMARY

CONCLUSIONS

The detection of circulating tumor cells (CTC) is a difficult task. CTC are very heterogeneous in size, shape, antigen expression, and are present at very low frequencies. Missing or misjudging a few events may have great consequences, as the clinically used cut-off for determining favorable or unfavorable prognosis for a patient is currently at three or five CTC in 7.5 ml of blood, depending on the cancer type. It is thus of utmost importance that all objects that are a threat for a patient are counted in reliable way. In this thesis, methods to quantify CTC and their features are presented.

In **chapter 1**, background is provided on cancer and the role of CTC in this disease; the CellSearch system developed for enumeration of CTC is explained. The CellSearch system was validated on a large number of clinical studies and has proven its value. It is however unknown if other definitions of CTC exist that are even more relevant for patient survival and if these definitions can improve the detection of a rise or drop in the number of CTC in patients over time. Therefore, there is great need for automated, quantitative methods: to remove current variability in counting CTC, to perform a fair comparison between measurements on different systems, and to investigate quantitative biomarker levels.

In **chapter 2**, different methods of spectral imaging are compared by simulation and a method of microscope calibration is shown which enables fair comparison between systems by using a well-defined light source. We show that a tailored method of spectral imaging, using dichroic mirrors that are specific for a set of fluorochromes, performs the task of imaging an assay of quantum dots a factor 15 faster than other spectral imaging methods. When using a LED-board for microscope calibration, the differences between real life setups are smaller: a factor 3 to 10, depending on the wavelength used. It furthermore becomes apparent that different cameras, although employing the same CCD chip, have very different characteristics. It is thus very important to use such calibration tools in the future: no fair comparison between measurements is otherwise possible.

In **chapter 3**, an algorithm is shown that counts fluorescent in situ hybridization (FISH) probes in CTC. FISH probes label specific regions on individual chromosomes, if a cell has more than two copies of a certain

probe, its either aneuploid or amplified with a high probability of being cancerous. We show that counting of chromosomes by humans using FISH is hampered by intra- and inter-reviewer variations. When reviewing CTC, variation between reviewers was measured as high as 33.2%, thus showing the need for proper standardization of this technique. A probe-counting algorithm was trained on leukocyte samples and tested on CTC samples. We show comparable results between reviewers and the algorithm when counting CTC. However, reproducibility of the algorithm is 100%, so when the computer makes mistakes, at least they will be consistent.

In **chapter 4**, we modified the algorithms for finding chromosomal aberrations to detect tumor particles and cells. Technical details of the algorithm that detects objects in the images recorded by the CellSearch system of the blood of prostate cancer patients are provided. To select the most hazardous objects for patients, we selected objects by means of certain parameters and measured their influence on survival of these patients. Parameters of objects were ranked based on their ability to dichotomize the patient group into favorable and unfavorable groups. Furthermore, the algorithm was tuned such that the number of objects that were found in the control group of healthy subjects were kept as low as possible. The strongest definition of automated CTC required that objects had a high standard deviation of the cytokeratin signal, a high maximum value of the DNA signal, a low maximum value of the CD45 signal, and a size range between 34 and 224 μm^2 . Stratifying patients based on these parameters resulted in comparable Cox hazard ratios (HRs) as compared to the manual selected cells by human reviewers. However, the computer algorithm is much faster and has 0% variability. Of note, the optimal definition required objects to contain DNA and not contain CD45, thereby showing that the optimal definition dismisses small tumor particles.

In **chapter 5** the CTC definition defined in chapter 4 was validated on an independent data set from prostate cancer patients; our group was blinded to the survival data of these patients. Again, our definition showed equal prognostic value as compared to the manual CellSearch definition, being the first automated algorithm to count CTC reliably. We furthermore show that when the DNA requirement in the definition was left out, HRs dropped by much. Leaving out the CD45 requirement did not influence the HR much, but allowed for a lot of objects to be counted in healthy controls, which is undesirable.

In **chapter 6**, we applied the automated counting algorithm on images from samples of breast and colon cancer patients. Additionally, we verified if size, nuclear-to-cytokeratin ratio, roundness, clusters and speckled CTC had impact on survival of patient. Although correlation between counts using the manual and automated count was low in some studies, prognostic value was similar using both methods. We show that morphological parameters of CTC differ between different cancer types. Multivariate analysis showed furthermore that size and nuclear-to-cytokeratin ratio had significant impact

on survival in breast cancer: small cells with small nuclei compared to their cytokeratin signal are associated with unfavorable prognosis.

In **chapter 7**, we investigated the possibility of measuring an actual drop in CTC count after administration of therapy. An actual drop indicates that therapy has an effect. However, counts are usually low in patients and thus the Poisson error becomes large. We applied different criteria to measure a drop on patient data at multiple time points and show that treatment should focus on getting the number of CTC to zero: a relative or absolute reduction still may leave patients with CTC and has low impact on survival. We furthermore show that a very strict CTC definition is the most useful because of low background in healthy controls. Other definitions, allowing for example tumor micro particles to be counted, did not improve our ability to count a change in the number of CTC.

In our final **chapter 8**, we enabled the quantitative power of our new CTC definition to measure an extra biomarker on CTC. We used the CellSearch control channel to image cells with Her-2 conjugated antibodies in samples of metastatic and non-metastatic breast cancer patients. We show again good correlation between manual and automated count, and a very large heterogeneity of Her-2 expression on CTC within patients. By applying a 75% threshold of Her-2 positive CTC in patients with ≥ 5 CTC, a relatively low discrepancy rate of 2% of metastatic patients with Her-2 negative primary tumors that had Her-2 positive CTC status was found; 29% of patients with Her-2 positive primary tumors had Her-2 negative CTC status. Of note, no Her-2 discrepancy was found between CTC and primary tumor in non-metastatic patients with ≥ 5 CTC, suggesting that Her-2 status shift could appear at metastatic relapse or treatment resistance.

OUTLOOK

The results shown in this thesis suggest that samples of other types of cancer should be investigated with our automated algorithm, for example samples from lung cancer patients. Samples from other patients may also benefit from the lower variability and fast processing using our algorithm. However, we also show that the low frequencies of CTC using the current system pose a real challenge, which cannot be resolved by counting more objects in the same sample. Larger blood volumes should be processed to circumvent this issue, as our group already showed that every metastatic patient has CTC in 5 liters of blood. Nevertheless, a wide range of extra biomarkers may be tested using this algorithm next to Her-2, to determine quantitative expression on CTC, such as IGF-1 or uPAR. Measuring expression of these receptors on CTC may allow for a true quantitative liquid biopsy. Such a biopsy, which is less tedious and hazardous to collect from patients than a traditional biopsy, could improve cancer diagnostics. Hopefully, it can bring cancer closer to the status of a chronic, controllable disease.

SAMENVATTING

CONCLUSIE

Het opsporen van circulerende tumor cellen (CTC) in het bloed van patiënten met kanker is een zware opgave. CTC lopen zeer uiteen in grootte, vorm, antigeen expressie en komen voor in zeer lage aantallen. Indien er een paar CTC verkeerd beoordeeld of over het hoofd gezien worden, dan kan dit grote gevolgen hebben, daar de huidige klinische grens voor een goede of slechte prognose van een patiënt bij drie of vijf CTC in 7.5 ml bloed ligt, afhankelijk van het type kanker. Het is daarom van zeer groot belang dat alle objecten die een gevaar opleveren voor een patiënt betrouwbaar geteld kunnen worden. In dit proefschrift worden daarom methoden behandeld om metingen aan CTC te automatiseren en te kwantificeren.

In **hoofdstuk 1** wordt achtergrond gegeven bij de ziekte kanker, de rol van CTC in het ziektebeeld en het CellSearch systeem om CTC op te sporen. Dit systeem is gevalideerd in een groot aantal klinische studies en heeft zijn waarde bewezen. Echter, het is op dit moment onduidelijk of er andere definities van CTC bestaan die nog meer impact kunnen hebben op de levensverwachting van patiënten, en of deze definities beter kunnen vaststellen of het aantal CTC in patiënten stijgt of daalt. Er is daarom grote behoefte aan geautomatiseerde, kwantitatieve definities en methoden: om de huidige variabiliteit in CTC karakterisatie te elimineren, om metingen tussen systemen accuraat te kunnen vergelijken en om kwantitatief naar biomarkers te kunnen kijken.

In **hoofdstuk 2** worden verschillende methoden om spectrale beelden op te nemen vergeleken door simulatie, en een methode wordt getoond om systemen eerlijk te kunnen vergelijken met behulp van een goed gedefinieerde lichtbron. We laten zien dat het spectraal opnemen van beelden met een toegesneden methode -gebruikmakende van dichroïsche spiegels- resulteert in een factor 15 snelheidswinst wanneer een assay van quantum dots gemeten wordt. Wanneer een LED kalibratiebron wordt gebruikt in een echt systeem, dan blijkt dit verschil kleiner: een factor 3 tot 10, afhankelijk van de gebruikte golflengte. Tevens wordt duidelijk dat verschillende camera's met dezelfde CCD chip sterk afwijkende eigenschappen vertonen. Het is daarom erg belangrijk om in de toekomst systemen te blijven kalibreren met een degelijk instrument: anders zijn metingen moeilijk te vergelijken.

Hoofdstuk 3 beschrijft een algoritme waarmee het mogelijk is fluorescente DNA probes in CTC te tellen. Met dergelijke probes kunnen specifieke chromosoomgebieden gekleurd worden. Indien een cel meer dan 2 kopieën heeft van een degelijk probe, dan is de cel aneuploid en hoogstwaarschijnlijk een kankercel. We laten zien dat het tellen van deze probes door experts geplaagd wordt door inter- en intra-expert variaties. Gemeten variaties tussen experts liepen op tot 33.2%, waardoor duidelijk werd dat deze techniek beter gestandaardiseerd moet worden. Een computer algoritme werd getraind op beelden van leukocyten en getest op CTC. We laten zien dat dit algoritme gelijkwaardige aantallen chromosomen telt vergeleken met experts, het algoritme doet dit echter 100% reproduceerbaar. Als de PC fouten maakt, dan doet hij dit in ieder geval altijd hetzelfde.

In **hoofdstuk 4** hebben we het bovenstaande algoritme aangepast om tumor deeltjes en CTC te detecteren. Technische details worden gegeven over de detectie van objecten in het bloed van prostaatkanker patiënten zoals afgebeeld door het CellSearch systeem. Om de gevaarlijkste objecten voor patiënten te selecteren, werden parameters van deze objecten gecorreleerd met de tijd dat patiënten overleefden. De parameters werden gerangschikt op geschiktheid om de patiënten groep te scheiden op relatief goede en slechte prognose. Verder werd het algoritme aangepast om zo weinig mogelijk vals-positieve objecten te tellen in gezonde controle personen. De beste CTC definitie liet slechts objecten toe met een hoge standaard deviatie van het cytokeratine signaal, een hoge maximale waarde van het DNA signaal, een lage maximale waarde van het CD45 signaal en een grootte tussen de 34 en 224 μm^2 . Het moet hierbij opgemerkt worden dat kleine tumor deeltjes niet inbegrepen zijn in deze definitie, daar zij geen DNA bevatten. Het scheiden van de patiënten groep met behulp van deze optimale definitie resulteerde in vergelijkbare Cox hazard ratio's (HR's) zoals bepaald met de huidige CellSearch definitie. Echter, de PC vervulde deze taak veel sneller en zonder variabiliteit.

In **hoofdstuk 5** hebben we het algoritme van hoofdstuk 4 extern gevalideerd op een onafhankelijk data set van prostaatkanker patiënten, zonder dat we de overlevingsdata van de patiënten hadden. De door het algoritme opgelegde definitie had wederom dezelfde prognostische waarde als de CellSearch definitie; ons algoritme is hierdoor het eerste algoritme dat betrouwbaar en snel CTC kan tellen. We laten verder zien dat als we de DNA parameter uit onze definitie negeren, dat dan de HR sterk daalt. Deze HR werd niet sterk beïnvloed door de CD45 parameter, maar deze parameter zorgt ervoor dat het aantal objecten in gezonde proefpersonen laag bleef.

In **hoofdstuk 6** passen we het algoritme vervolgens toe op bloed samples van borst- en darmkanker patiënten. Verder onderzochten we of de parameters grootte, kern-cytokeratine ratio, rondheid, aantal clusters en aantal gespikkelde CTC van invloed was op de overleving van de patiënten. De correlatie tussen de geautomatiseerde en CellSearch definitie bleek laag in sommige studies, de prognostische waarde was echter vergelijkbaar bij

beide methoden. We laten zien dat er morfologische verschillen bestaan tussen CTC afkomstig van verschillende typen kanker. Met een multivariabele analyse laten we verder zien dat grootte en kern-cytokeratine ratio van CTC van significante invloed is op overleving van borstkanker patiënten. Kleine cellen met een kleine kern in vergelijking tot hun cytotkeratine signaal geven een slechte prognose voor de patiënt.

In **hoofdstuk 7** bestuderen we de mogelijkheid om een daling in het aantal CTC van patiënten te meten na het starten van therapie: een daling geeft aan dat de therapie werkt. Echter, het aantal CTC in patiënten is vaak zo laag dat dit resulteert in een grote Poisson fout. Daarom passen we verschillende criteria toe om een daling te meten op verschillende tijdstippen. We tonen aan dat therapie moet focussen op het elimineren van alle CTC in patiënten. Een relatieve of absolute daling waarbij nog steeds CTC overblijven in de patiënt heeft weinig invloed op de overleving. Verder laten we zien dat een strikte CTC definitie in de kliniek het nuttigst is, daar deze weinig achtergrond objecten in gezonde proefpersonen telt. Andere definities die bijvoorbeeld kleine tumor deeltjes in het bloed ook meetellen verbeteren niet de mogelijkheid om een verandering in het aantal CTC gedurende therapie te tellen.

Hoofdstuk 8 laat tenslotte de kwantitatieve kracht van ons nieuwe algoritme zien door het meten van een extra biomarker in CTC. Gebruikmakende van het controle kanaal in het CellSearch systeem hebben we beelden opgenomen van cellen gelabeld met antilichamen gekoppeld aan het Her-2 eiwit in samples van metastatische en niet-metastatische patiënten. De correlatie tussen CellSearch en geautomatiseerde aantallen is sterk, we laten verder zien dat er een grote variatie aan Her-2 expressie binnen patiënten is. Door het toepassen van een criterium dat 75% van de CTC Her-2 positief moet zijn in patiënten met ≥ 5 CTC laten we zien dat in 2% van de patiënten met een Her-2 negatief primair tumor Her-2 positieve CTC status gevonden wordt. In 29% van de patiënten met een Her-2 positief primair tumor vinden we een negatieve Her-2 CTC status. Hierbij moet opgemerkt worden dat we in niet-metastatische patiënten geen discrepantie tussen primair tumor en CTC Her-2 status vonden. Deze resultaten suggereren dat een verandering in Her-2 status in CTC op zou kunnen treden als de patiënt een metastatische terugval heeft of doordat CTC resistent worden voor de therapie.

VOORUITZICHTEN

De resultaten in dit proefschrift suggereren dat samples van patiënten met anderen typen kankers onderzocht moeten worden met behulp van ons algoritme, zoals bijvoorbeeld longkanker samples. Deze patiënten zouden ook kunnen profiteren van de lagere variabiliteit in getelde aantallen en een sneller telproces. Echter laten we ook zien dat de lage aantallen CTC een

grote uitdaging vormen, een uitdaging die niet aangegaan kan worden door meer objecten te tellen in onze huidige samples. Onze onderzoeksgroep heeft al aangetoond dat alle metastatische patiënten CTC in hun bloed hebben: grotere bloedvolumes moeten daarom bekeken worden, willen we deze CTC te pakken krijgen. Desalniettemin kan in de huidige samples met ons algoritme kwantitatieve expressie van vele andere biomarkers onderzocht worden. Het meten van dergelijke expressie kan uiteindelijk leiden tot een heus vloeibaar biopt van een patiënt. Het verkrijgen van een vloeibaar biopt is eenvoudiger en levert minder gevaar op voor de patiënt dan het afnemen van huidige biopten. Het zou de kwaliteit van de diagnose van kankerpatiënten aanzienlijk kunnen verbeteren. In de toekomst zou kanker met behulp van de technieken uit dit proefschrift hopelijk dichterbij de status van een chronische, controleerbare ziekte kunnen komen.

PUBLICATIONS

JOURNAL ARTICLES AND BOOK CHAPTERS

T. M. Scholtens, F. Schreuder, S. T. Ligthart, J. F. Swennenhuis, A. G. J. Tibbe, J. Greve, L. W. M. M. Terstappen, “CellTracks TDI: an Image Cytometer for Cell Characterization,” *Cytometry Part A*, vol. 79, pp 203–213, 2011.

S. T. Ligthart, F. A. W. Coumans, G. Attard, A. Mulick Cassidy, J. S. de Bono, L. W. M. M. Terstappen, “Unbiased and Automated Identification of a Circulating Tumour Cell Definition that Associates with Overall Survival,” *PLoS ONE* vol. 6, no. 11, DOI:10.1371/2011/journal.pone.0027419, 2011.

T. M. Scholtens, F. Schreuder, S. T. Ligthart, J. F. Swennenhuis, J. Greve, L. W. M. M. Terstappen, “Automated Identification of Circulating Tumor Cells by Image Cytometry,” *Cytometry Part A*, vol. 81A, pp 138–148, 2011.

S. T. Ligthart, J. F. Swennenhuis, J. Greve, L. W. M. M. Terstappen, “FISH Probe Counting in Circulating Tumor Cells,” in: *Cytokeratins, Tool in Oncology*, Editor: Gerhard Hamilton, ISBN 978-953-51-0047-8, Imtech, 2012.

F. A. W. Coumans, S. T. Ligthart, L. W. M. M. Terstappen, “Identification of Circulating Tumor Cells,” in: *Biofunctional Surface Engineering in Medicine*, Nanobiotechnology series, Pan Stanford Publishing, in press, 2012.

S. T. Ligthart, F. C. Bidard, C. Decraene, T. Bachelot, S. Delaloge, E. Brain, M. Campone, J. Y. Pierga, L. W. M. M. Terstappen, “Unbiased quantitative assessment of HER2 expression on Circulating Tumor Cells in patients with metastatic and non-metastatic breast cancer,” submitted.

F. A. W. Coumans, S. T. Ligthart, L. W. M. M. Terstappen, “All patients with metastatic carcinoma have circulating tumor cells,” submitted.

S. T. Ligthart, F. A. W. Coumans, L. W. M. M. Terstappen, “Interpretation of changes in Circulating Tumor Cell counts,” submitted.

S. T. Ligthart, F. A. W. Coumans, F.-C. Bidard, L. H. J. Simkens, C. J. A. Punt, M. de Groot, G. Attard, J. S. de Bono, J.-Y. Pierga, and L. W. M. M. Terstappen, “Morphology of Circulating Tumor Cells and survival in breast, colorectal and prostate cancer,” submitted.

CONFERENCE CONTRIBUTIONS (ORAL)

S. T. Ligthart, F. Schreuder, J. Greve, A. G. J. Tibbe, L. W. M. M. Terstappen, "Theoretical Comparison of Spectral Imaging Methods for measuring Aneuploidy," Focus on Microscopy: Valencia, Spain, April 10–13, 2007.

S. T. Ligthart, J. Greve, L. W. M. M. Terstappen, "Simulation of spectral imaging methods to measure aneuploidy," International Society for analytical Cytometry XXIV: Budapest, Hungary May 17–21, 2008.

S. T. Ligthart, J. F. Swennenhuis, R. Sipkema, L. W. M. M. Terstappen, "Automatic FISH Probe Counting in Circulating Tumor Cells," Annual Dutch Meeting on Molecular and Cellular Biophysics: Veldhoven, the Netherlands, September 28-29, 2008.

S. T. Ligthart, L. W. M. M. Terstappen, "Comparison of Spectral Imaging Methods by simulation, for measuring Aneuploidy with Qdot labeled FISH probes," Focus on Microscopy: Krakow, Polen, April 5–8, 2009.

S. T. Ligthart, M. Cloosterman, J. Hilderink, M. Snijder-van As, C. Ungureanu, "Finding Advanced Markers For Tissue Ischemia: Developing an optical pen to monitor ischemia markers in vivo," MIRA Event Top Technology for Patients: Hengelo, The Netherlands, March 31 - April 1, 2010.

L. W. M. M. Terstappen, G. van Dalum, F. A. W. Coumans, A. Hoepener, S. T. Ligthart, J. F. Swennenhuis, European Multidisciplinary Cancer Congress on Integrating Basic and Translational Science, Surgery, Radiotherapy, Medical oncology, Advocacy and Care: Stockholm, Sweden, September 23-27, 2011

CONFERENCE CONTRIBUTIONS (POSTER)

S. T. Ligthart, F. Schreuder, J. Greve, L. W. M. M. Terstappen, A. G. J. Tibbe, "Fast Tumor Cell Detection by scanning for Chromosomal Aberrations," Summerschool Molecular Imaging: Heidelberg, Germany, April 4–8 2006.

F. Schreuder, T. M. Scholtens, S. T. Ligthart, A. G. J. Tibbe, L. W. M. M. Terstappen, J. Greve, "A fast imaging cytometer based on Time Delay Integration," International Society for analytical Cytometry XXIII: Quebec City, Canada, May 20-24, 2006.

F. Schreuder, T. M. Scholtens, S. T. Ligthart, A. G. J. Tibbe, J. F. Swennenhuis, J. Greve, L. W. M. M. Terstappen, "Characterization of circulating tumor cells by image cytometry," International Society for analytical Cytometry XXIV: Budapest, Hungary May 17–21, 2008.

S. T. Ligthart, J. F. Swennenhuis, R. Sipkema, J. Greve, A. G. J. Tibbe, L. W. M. M. Terstappen, "Automatic FISH Probe Counting on Circulating Tumor Cells," International Society for analytical Cytometry XXIV: Budapest, Hungary May 17–21, 2008.

S. T. Ligthart, J. F. Swennenhuis, R. Sipkema, L. W. M. M. Terstappen, "Automatic FISH Probe Counting in Circulating Tumor Cells," BMTI symposium: Enschede, the Netherlands, November 20, 2008.

S. T. Ligthart, F. A. W. Coumans, G. Attard, J. S. de Bono, L.W.M.M. Terstappen, "Automatic classification of EpCAM+ CD45- cytokeratin+ objects versus survival in castration resistant prostate cancer," MIRA Event Top Technology for Patients: Hengelo, The Netherlands, March 31 - April 1, 2010.

S. T. Ligthart, F. A. W. Coumans, G. Attard, J. S. de Bono, L.W.M.M. Terstappen, "Automatic classification of EpCAM+ CD45- Cytokeratin+ objects versus survival in castration resistant prostate cancer," the 2010 meeting of the American Association of Cancer Research: Washington DC, U.S.A, April 17-21, 2010.

S. T. Ligthart, J. F. Swennenhuis, J. Greve, L. W. M. M. Terstappen, "Development of automatic FISH probe counting in CTC," the 2010 meeting of the American Association of Cancer Research: Washington DC, U.S.A, April 17-21, 2010.

T. M. Scholtens, F. Schreuder, S. T. Ligthart, J. F. Swennenhuis, A. G. J. Tibbe, J. Greve, L. W. M. M. Terstappen, "Celltracks TDI: An Image Cytometer for Cell Characterization," Lorentz meeting circulating tumor cell isolation and diagnostics: Leiden, the Netherlands, February 7-11, 2011.

S. T. Ligthart, F. A. W. Coumans, G. Attard, A. Cassidy, J. S. de Bono, L. W. M. M. Terstappen, "Automated identification of Circulating Tumor Cell," Lorentz meeting Circulating tumor cell isolation and diagnostics: Leiden, the Netherlands, Februari 7-11, 2011.

S. T. Ligthart, F. A. W. Coumans, G. Attard, A. Mulick Cassidy, J. S. de Bono, and L. W. M. M. Terstappen, "Automated classification of circulating tumor cells optimized using clinical outcome of castration resistant prostate cancer patients," the 2011 meeting of the American Association of Cancer Research: Orlando FL, U.S.A., April 2-6, 2011.

F. C. Bidard, S. T. Ligthart, C. Decraene, T. Bachelot, S. Delalogue, E. Brain, M. Campone, J. Y. Pierga, L. W. M. M. Terstappen, "Automated Quantitative Assessment of HER2 Expression of Circulating Tumor Cells (CTC) in Metastatic Breast Cancer (IC 2006-04 Study)," The 2011 CTRC-AACR San Antonio Breast Cancer Symposium, San Antonio, TX, USA, December 6-10, 2011.

F. C. Bidard, S. T. Ligthart, C. Decraene, T. Bachelot, S. Delalogue, E. Brain, M. Campone, J. Y. Pierga, L. W. M. M. Terstappen, "Unbiased quantitative assessment of HER2 expression on Circulating Tumor Cells in patients with metastatic and non-metastatic breast cancer," the 8th European Breast Cancer Conference, Brussels, Belgium, March 20-24, 2012.

J. Y. Pierga, S. T. Ligthart, C. Decraene, T. Bachelot, S. Delalogue, E. Brain, M. Campone, F. C. Bidard, L. W. M. M. Terstappen, "Unbiased quantitative assessment of HER2 expression on Circulating Tumor Cells in patients with metastatic and non-metastatic breast cancer," the 2012

meeting of the American Association of Cancer Research: Chicago IL, U.S.A., March 30 - April 4, 2012.

F. A. W. Coumans, S. T. Ligthart, L. W. M. M. Terstappen “All patients with metastatic breast, colorectal and prostate carcinoma have circulating tumor cells,” the 2012 meeting of the American Association of Cancer Research: Chicago IL, U.S.A., March 30 - April 4, 2012.

ACKNOWLEDGMENTS

Finally, after 4++ years of struggling with cells, microscopes, and computers, I am done. It was not an easy journey. It all started when Leon came to me with the topic for a PhD project: use a computer to find valuable information in the images from the CellSearch clinical studies. Of course I did not accept his offer, as spending hours and hours behind a PC seemed rather boring. I wanted to build stuff! So I spend the year working as a researcher, until Leon came with a second proposal: build a setup and develop an assay for counting all chromosomes in circulating tumor cells. This subject was right for me, as so I thought at the end of 2006.

Obviously, I started with the best intentions, simulating stuff, designing other stuff, and labeling some more stuff for months. Until after a considerable amount of time, we assessed the situation and concluded that it was never going to work. Luckily, I had done some work on a computer algorithm for counting chromosomes in cells, so Leon suggested that this algorithm might also be good for a data set Frank Coumans had worked on: some clinical CellSearch data. My algorithm turned out to work well on this new data set, and all was good. I was a happy camper behind my PC until the horrible truth hit me: I was actually performing the work I had originally rejected...

Rest assured, I know that I could not have done it without the people around me. This project would never have been launched if it wasn't for the financial support from Veridex LLC. Thank you Mike, Bob, Jan, Chandra, Madeline, and all other people for your support. At the Veridex Enschede office, I spent many hours messing up work benches. Thank you Arjan and crew while you were still at the BTC. Leon: thank you for starting the project, keeping confident and for the tons of support, especially in the last year. Your clinical view on things has set me straight more times than I can count. Frank: your input in my project was invaluable. Thanks for being a technical sparring partner and a razor sharp critic. Joost: thanks for your help on FISHing and always being positive. Jan: thanks for sharing your experience. Ceas: your sharp view on spectral issues is greatly appreciated. Of course I cannot forget Ingrid: thank you for keeping me up to scratch. This work was also not possible without the help from our collaborators abroad. Johann, Gert, and Amy from Royal Marsden: thank you for sharing your data, for the discussions on how to combine technical and clinical data

and for the possibility of validating. Francois-Clement, Jean-Yves, and all the others at Institut Curie for letting us validate our algorithm on Her-2. Thanks to the people at MST and UMC Radboud hospitals for sharing data and research. At the TUD, I performed measurements using an exciting interferometer: thank you Ted and Ronald.

In my years as a PhD student, not many students were enthusiastic on working with Matlab in cancer research. Luckily I had two of the best to help me: thanks Edzo and Bernd for your efforts and the company when the Carre was still empty (and hot!). Furthermore, I was lucky enough to be accompanied by very enthusiastic people during my stay in de Hogekamp, Zuidhorst, and Carre buildings. Thank you Bart, for the being a Matlab tutor and all the fun (and hits hits hits!) we had. Thanks to Christian for the technical assistance and great company. Thank you Tycho, Erik, Babu, Vinay, Remco, Vivek, Niels, Khalid, Liesbeth, Astrid, Kiki, Niels, and Barbara for sharing coffee, pc (and PhD) frustrations. Next to these people, I have drank coffee and had great discussions with too many people to put down on paper. So to all the people of the former BPE group (now in MCBP, NBP, and BMPI): thanks!

Reflecting on my time in the Zuidhorst building, I will always remember the great ease to meet and have fun with people from other groups. My total negligence on the subject of polymers did not stand in the way of making friends in the (former) PBM group. Lucky for me, these guys knew a thing or two about good beers, throwing parties and making legendary videos. Thank you Andries, Hans, Niels, Sytze, and all the other (PhD) students at PBM, but also at TR, BSS, and BW! We spend some great time sailing/playing poker/dancing and screaming/making sushi/organizing triathlons/trying every single special beer that was available in Twente.

It was a pleasure to work on a technical project, but a conversation about totally nothing sometimes works very refreshing. More than five years ago, a poker club was founded that fulfilled that need excellently. I really enjoyed the lively discussions and crazy hands we've played. Thanks to all people who joined our crazy evenings, who are still there and those who will stop by in the future. Thanks to all my friends in Enschede, Amsterdam, Gouda, and Utrecht, for great conversations and parties. Willem: thanks for all the UAZ's, woestijntjes, and of course why?!

So it has been a long journey, full of interesting topics and great people. Finally, I'd like to thank my paranims Andries and Michiel to be brave enough to join me on the podium, and my parents. Mom, dad, thanks for your unrelenting support, even if things take a bit longer than expected.

Enschede, 11 april 2012

ABOUT THE AUTHOR

Sjoerd Theodorus Ligthart was born on Februari 4th 1980, in Hoogkarspel, the Netherlands. He graduated from high school S.G. Werenfridus in Hoorn in 1998, and started studying applied physics at the University of Twente in Enschede that same year. During his studies, he organized a study tour to Switzerland and Italy, and was in the board of study association Arago. After an intership at Immunicon Corporation in 2004, he became focused on biophysics and circulating tumor cells. In 2005 he graduated on “Improved enumeration of Circulating Tumor Cells by 3D image restoration in Celltracks” in the BioPhysical Engineering group of Prof. Dr. Jan Greve. He continued his research on circulating tumor cells in the Medical Cell BioPhysics group of Prof. Dr. Leon Terstappen, which led to this thesis.