

Article

Classification of Cells in CTC-Enriched Samples by Advanced Image Analysis

Sanne de Wit ^{1,†} , Leonie L. Zeune ^{1,2,†} , T. Jeroen N. Hiltermann ³, Harry J. M. Groen ³, Guus van Dalum ⁴ and Leon W. M. M. Terstappen ^{1,*} 

¹ Department of Medical Cell BioPhysics, University of Twente, 7522 NH Enschede, The Netherlands; s.dewit@utwente.nl (S.W.); l.l.zeune@utwente.nl (L.L.Z.)

² Department of Applied Mathematics, University of Twente, 7522 NH Enschede, The Netherlands

³ Department of Pulmonary Diseases, University Medical Center Groningen, University of Groningen, 9713 GZ Groningen, The Netherlands; t.j.n.hiltermann@umcg.nl (T.J.N.H.); h.j.m.groen@umcg.nl (H.J.M.G.)

⁴ Department of General, Visceral and Pediatric Surgery, University Hospital of the Heinrich-Heine-University Düsseldorf, 40225 Düsseldorf, Germany; Guus.vanDalum@med.uni-duesseldorf.de

* Correspondence: l.w.m.m.terstappen@utwente.nl; Tel.: +31-(053)-489-2425

† Both authors contributed equally.

Received: 21 August 2018; Accepted: 8 October 2018; Published: 10 October 2018



Abstract: In the CellSearch[®] system, blood is immunomagnetically enriched for epithelial cell adhesion molecule (EpCAM) expression and cells are stained with the nucleic acid dye 4'6-diamidino-2-phenylindole (DAPI), Cytokeratin-PE (CK), and CD45-APC. Only DAPI+/CK+ objects are presented to the operator to identify circulating tumor cells (CTC) and the identity of all other cells and potential undetected CTC remains unrevealed. Here, we used the open source imaging program Automatic CTC Classification, Enumeration and PhenoTyping (ACCEPT) to analyze all DAPI+ nuclei in EpCAM-enriched blood samples obtained from 192 metastatic non-small cell lung cancer (NSCLC) patients and 162 controls. Significantly larger numbers of nuclei were detected in 300 patient samples with an average and standard deviation of $73,570 \pm 74,948$, as compared to 359 control samples with an average and standard deviation of 4191 ± 4463 ($p < 0.001$). In patients, only $18\% \pm 21\%$ and in controls $23\% \pm 15\%$ of the nuclei were identified as leukocytes or CTC. Adding CD16-PerCP for granulocyte staining, the use of an LED as the light source for CD45-APC excitation and plasma membrane staining obtained with wheat germ agglutinin significantly improved the classification of EpCAM-enriched cells, resulting in the identification of $94\% \pm 5\%$ of the cells. However, especially in patients, the origin of the unidentified cells remains unknown. Further studies are needed to determine if undetected EpCAM+/DAPI+/CK-/CD45- CTC is present among these cells.

Keywords: circulating tumor cells; CellSearch[®]; EpCAM; leukocytes; ACCEPT; deep Learning; classification; image analysis; non-small cell lung cancer

1. Introduction

Circulating tumor cells (CTC) are cells disseminated from the primary or metastatic tumor site into the bloodstream. In several cancer types, including non-small and small cell lung, prostate, breast, colon, bladder, gastric carcinoma, and melanoma, the presence of CTC in circulation is associated with a poor survival rate of the patients [1–9]. With many blood cells circulating in our cardiovascular system, CTC are rare events. A concentration of one CTC per mL of blood is already high, while the concentration of white blood cells is around 5×10^6 cells and red blood cell concentrations reach up to 5×10^9 cells per mL. When detecting CTC, it is therefore of utmost importance to locate all CTC

present and have an extremely sensitive and specific assay to discriminate between the tumor cells and blood cells. In the CellSearch[®] system, epithelial cell adhesion molecule (EpCAM) antibodies coupled to ferrofluids are used for immunomagnetic enrichment of cells thereby removing the bulk of leukocytes. After this, CTC are fluorescently labeled with: 4',6-diamidino-2-phenylindole (DAPI) to stain the nucleus; phycoerythrin (PE) labeled antibodies recognizing cytokeratins (CK) 4–6, 8, 10, 13, 18, and 19, which are present in the majority of cells derived from epithelial cancers, and allophycocyanin (APC) labeled cluster of differentiation (CD) 45, which is present on the cell surface of leukocytes [10]. Images covering the surface of the CellTracks cartridge are acquired by the CellTracks Analyzer II, and all events that are positive for both DAPI and CK are presented as thumbnail images to the operator, who will decide which cells follow the criteria for CTC assignment. The identity and number of all other nucleated cells are not provided, thereby obscuring any potential cells of interest. These cells can now be explored using the open source image analysis program Automatic CTC Classification, Enumeration and PhenoTyping (ACCEPT) (download on <https://github.com/LeonieZ/ACCEPT>). The ACCEPT program is an open source toolbox designed to analyze and characterize all objects present in the fluorescent images.

With this program, we analyzed 659 EpCAM-enriched blood samples, obtained from 162 healthy donors or patients with benign disease and 192 metastatic non-small cell lung cancer (NSCLC) patients. We pursued several methods to determine the extent and identity of all cell populations present, in order to assess the number of cells that potentially are CTC.

We observed that the number of cells identified as leukocytes in the EpCAM-enriched fraction was low and showed a large variance across the samples. We investigated several possible causes: A possible correlation with sample or patient attributes, insufficiencies in the image analysis, and insufficiencies in the staining of the leukocytes or its detection. To improve this last option, we explored several avenues. We showed previously that addition of CD16, recognizing the FcγIIIa, to the staining cocktail improved the identification of nuclei as granulocytes since they express CD45 at a lower level compared to lymphocytes and monocytes [11]. We now applied this antibody in a large cohort of patient samples. We also explored an alternative LED excitation light source to improve the excitation efficiency of the APC fluorochrome, as opposed to the mercury arc lamp which is used as a light source in the CellTracks [12]. In addition, wheat germ agglutinin (wga), a lectin that binds to sialic acid and N-acetylglucosamine residues, was explored because it is present on cellular plasma membranes [13–18]. We examined the presence of a cell membrane by using wga conjugated to AlexaFluor-488 and adding it to the CellSearch immunostaining cocktail.

In summary, we present a method to closely examine every cell present in the cartridge after EpCAM enrichment. CTC subpopulations expressing low or no CK are currently not examined, and their potential clinical utility remains unexploited. Here we describe several methods to discover the extent of the anonymous cell population that could harbor these unidentified CTC.

2. Results

2.1. Enumeration of Nucleated Cells in the EpCAM-Enriched Cells

The image analysis program ACCEPT was used to enumerate the total number of nucleated cells in the image files after EpCAM enrichment. Figure 1a–c shows an example of the ACCEPT analysis in which the nucleated cells (blue dots) are identified among the other objects in the images (grey dots). To identify nucleated cells in the cartridge images, showing the following criteria for DNA staining were used: 1. Mean fluorescence intensity ≥ 50 (panel a); 2. standard deviation of the fluorescence intensity ≥ 20 ; 3. Size $< 500 \mu\text{m}^2$ (panel a); and 4. Roundness < 0.95 , where 0 is a perfect circle (panel b). To identify single cells, a perimeter to area (P2A) of < 1.5 was used, for doublets a P2A between ≥ 1.5 and < 2.5 , for small clusters a P2A between ≥ 2.5 and < 4 , and for large clusters a P2A ≥ 4 (panel b). For enumeration of nucleated cells, the doublets and clusters were counted as one. An example of the single cells, double cells, and clusters that were identified in this manner are illustrated in Figure 1d–g.

Figure 2 shows the nucleated cell count from 300 samples obtained from 192 NSCLC patients (mean $73,570 \pm 74,948$, median 47,192) and from 359 samples obtained from 162 healthy donors and patients with benign disease (mean 4191 ± 4463 , median 2729). The number of nucleated cells was significantly larger in the patient samples as compared to the samples from healthy donors and patients with benign disease (Mann-Whitney U test: $p < 0.001$).

For a complete analysis, we determined the association between the number of nucleated cells and the sample type (“patients” or “controls”). For this analysis, 300 samples from metastatic NSCLC patients and 359 samples from healthy volunteers and patients with benign disease were used.

First, several factors were investigated that could influence the nucleated cell count since these factors were present in a limited set of patients or controls only. As the first factor, we evaluated the influence of the assay itself by determining the nucleated cell count from 30 controls with benign disease from which four blood tubes were obtained and processed simultaneously, see Figure S1. It indicated that a coefficient of variation of less than 10.7% or more than 40.2% (mean of $25.6 \pm 14.9\%$ (1 standard deviation)) of the nucleated cell count cannot be attributed to assay variation with a 97% certainty. This was only the case for six (20%) controls. The influence of the assay itself was therefore excluded from further analysis. Second, undergoing treatment as a contributing factor was considered. In 64 NSCLC patients, we observed that 95% of the patients showed an increase or decrease in nucleated events of more than 10% during treatment, compared to the nucleated events at the start of treatment, see Figure S1. We considered this a significant change and therefore this factor was included in further analysis. The third factor investigated was gender. No influence of the gender (133 male patients, 127 female patients, and 40 patients of unknown gender) on the nucleated cell count was observed with the Mann-Whitney U test ($p = 0.237$), and the factor was therefore excluded from further analysis.

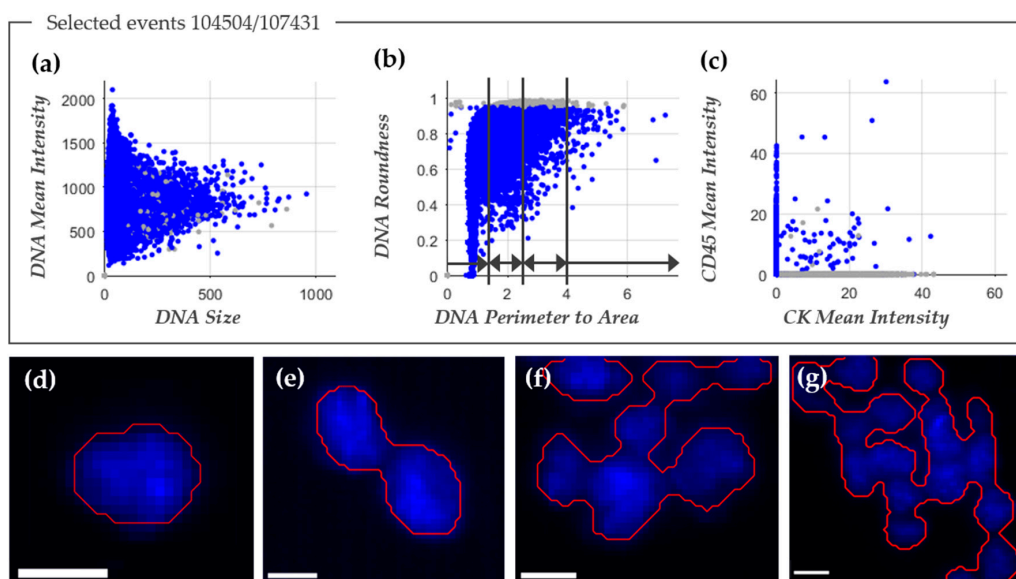


Figure 1. ACCEPT identification of nucleated cells. The three scatter plots in panels (a–c) were used to define nucleated cells (depicted in blue) by ACCEPT. In total, 104,504 events were detected and 107,431 of them were classified as nucleated cells, whereas the other events are depicted as grey dots. The division into single cells, doublets, small, and large clusters was based on DNA perimeter to area ratio, as illustrated in panel (b). In the scatter plot in panel (c), the mean fluorescence intensity of Cytokeratin (CK)-phycoerythrin (PE) (CK mean intensity) is plotted against the mean fluorescence intensity of CD45-allophycocyanin (APC) (CD45 mean intensity). In panels (d–g) typical examples of the segmentation (red line) around the nucleus of a single cell are illustrated in panel (d); of a doublet in panel (e); a small cluster in panel (f); and a large cluster in panel (g). The white scale bar represents 10 pixels, which corresponds to a size of $6.4 \mu\text{m}$.

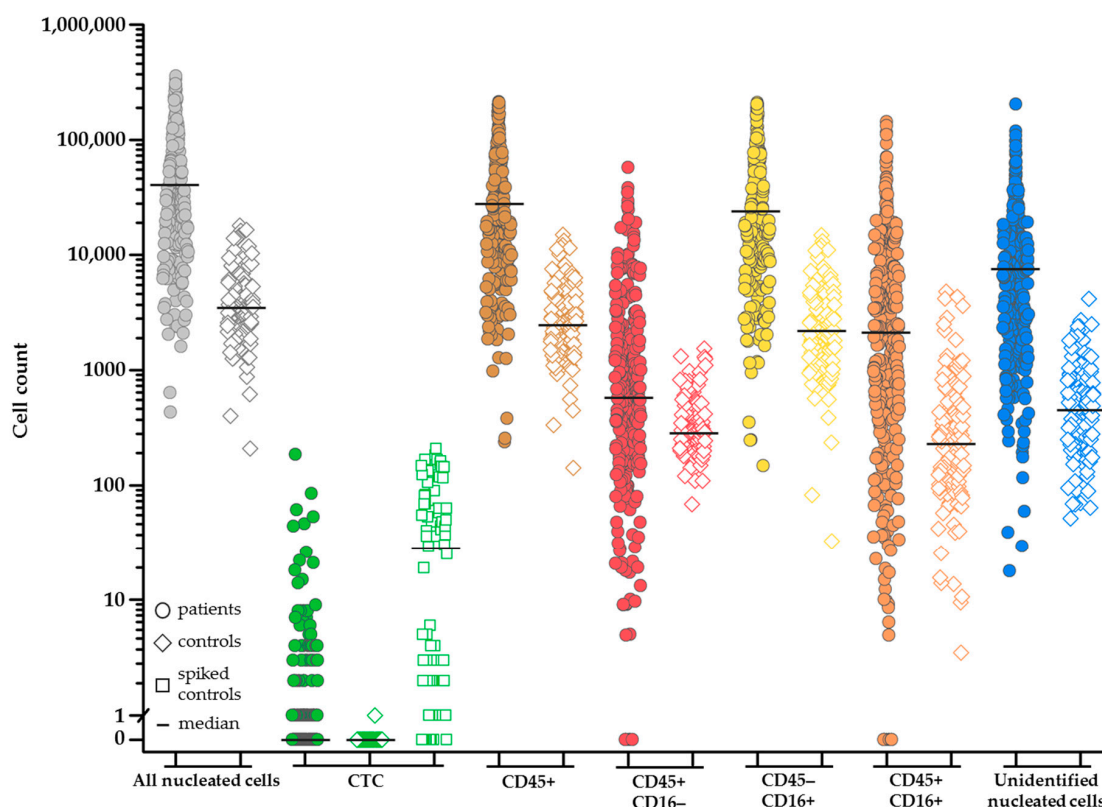


Figure 2. Frequency distribution of the cell populations detected in CellSearch[®] and with the addition of the CD16 classification after epithelial cell adhesion molecule (EpCAM) immunomagnetic enrichment of 300 blood samples from 192 non-small cell lung cancer (NSCLC) patients (filled circle) and 127 blood samples from 20 healthy volunteers (open diamond). Healthy volunteer samples spiked with cell line cells ($n = 88$) are indicated with an open square. Cell populations: All nucleated cells (grey), circulating tumor cells (CTC) (green), CD45+ leukocytes (brown), CD45+/CD16- leukocytes (red), CD45-/CD16+ leukocytes (yellow), CD45+/CD16+ leukocytes (orange), and unidentified cells (blue). The differences in cell count between the NSCLC patients and healthy volunteers in the same cell classification is significant for all subclasses ($p < 0.001$). Samples from healthy volunteers were not spiked (diamond symbol, $n = 39$; median 0 CTC) or spiked (square symbol) with cancer cell lines PC3 ($n = 47$; median 63 CTC; 1.0×10^4 EpCAM antigens) and NCI-H460 ($n = 41$; median 1 CTC; 1.4×10^2 EpCAM antigens).

In the crude multiple regression analysis for all 300 NSCLC patients and 359 controls, the relation between sample types “patient” and “control” was established. The result of this multiple regression analysis is presented in Table 1. This was followed by an analysis to correct for patient age (younger vs. older than 55 years). As a third analysis, we corrected for the patient variables: CTC count (continuous variable), treatment (during treatment vs. before or no treatment) and sample age at the time of processing with CellSearch (1 day vs. 2–4 days).

Table 1. Multiple regression analysis for association between sample type and nucleated events corrected for several patient variables.

Analysis Type	Regression Coefficient	95% Confidence Interval	<i>p</i> -Value
Crude analysis	17.0	14.1–20.6	<0.001
Corrected analysis for <i>age</i>	19.2	15.3–24.2	<0.001
Corrected analysis for <i>age</i> , and patient variables (<i>sample age</i> , <i>CTC count</i> and <i>treatment</i>)	12.4	9.4–16.4	<0.001

Each variable was analyzed for its confounding effect on the number of nucleated events. This showed that age of the patient, undergoing treatment, and CTC count had little confounding effect (below 5%). The sample age showed a confounding effect of 11.7%.

2.2. Improved Image Analysis

The number of nucleated events present in some cartridges can be enormous (over 100,000 for 27% of the patients), covering the whole detection surface, while some cartridges barely hold any cells (less than 5000 for 6% of the patients). The high number of cells creates a large area of packed cells, which makes it difficult to define borders for each cell, either manually or by image analysis algorithms. The segmentation quality of the ACCEPT toolbox was influenced by the number of cells present in the sample, an illumination bias, and fragmented staining. This is illustrated in detail in Figure S2. These problems motivated the use of a segmentation approach based on deep learning. As opposed to the model-based approaches currently implemented in the ACCEPT toolbox, deep learning approaches do not make any model assumptions about the signal (e.g., a nearly homogenous background and cell intensity), but learn a model based on only the data used for the training.

2.3. Assignment of Nucleated Cells to a Cell Lineage

In the CellSearch system, CD45-APC is used to identify leukocytes and Cytokeratin-PE to identify CTC among the nucleated cells. Nucleated cells expressing both CD45-APC and Cytokeratin-PE are assigned as cells binding non-specifically to at least one of the antibodies. The number of CTC identified by the classical CellSearch method in the 300 samples from 192 NSCLC patients ranged from 0–186 (mean 2; median 0) and represent only a very small portion of the nucleated cells identified in the samples from these patients. To determine whether the identity of the majority of nucleated cells could be assigned to the hematopoietic cell lineage by their expression of CD45, the thumbnails of the nucleated cells were put into a trained deep learning network for segmentation and the segmented images were fed back into the ACCEPT toolbox. Using the marker characterization processor tool, the results could be visualized and analyzed. The improved segmentation allowed for a relatively simple definition of cell subpopulations and in most cases a simple yes or no could be used to define the presence or absence of a marker. The actual definitions of the different cell populations are listed in Table 2. In this manner, nucleated cells were classified as CD45+, CD45- and CK- or CK+ in 300 samples from 192 NSCLC patients and 127 samples from 20 healthy volunteers.

Table 2. The definition of cell populations used to classify cells in ACCEPT after deep learning segmentation in 300 non-small cell lung cancer (NSCLC) patient samples and 127 samples from healthy volunteers.

Cell Population		Mean Intensity Value					Overlay Nucleus			Size (μm^2)
		DAPI	CD45	CK	CD16 ¹	wga ¹	with CK	with CD45	with CD16	CK
CTC	CK+/CD45-	>0	0	≥ 50	0	>0	>0.4	ND	ND	>9
	CD45+/CD16-	>0	>0	0	0	>0	ND	>0	ND	ND
Leukocytes	CD45+/CD16+	>0	>0	0	>0	>0	ND	>0	>0	ND
	CD45-/CD16+	>0	0	0	>0	>0	ND	ND	>0	ND
Nucleus	Bare nucleus	>0	0	0	0	0	ND	ND	ND	ND
	Unstained cell	>0	0	0	0	>0	ND	ND	ND	ND

¹ Only for those samples in which CD16 or wga was added to the immunostaining cocktail. Abbreviations used: DAPI: 4'6-diamidino-2-phenylindole; CD: cluster of differentiation; CK: cytokeratins; wga: wheat germ agglutinin.

As described in the method section, for this analysis, 10% of the nucleated cells found by the ACCEPT toolbox, up to a maximum number of 5000 cells, were analyzed and the results were extrapolated to the full number of nucleated events to arrive at the cell numbers. The number of nucleated cells that could be identified as leukocytes in CellSearch was only 18% \pm 21% (range 0–89%; median 9%) in patients and 23% \pm 15% (range 2–68%; median 19%) in healthy volunteers, see Figure 3.

To be able to assign the cell lineage of origin of the nucleated cells on which no CD45- or CK-expression could be detected, several approaches were pursued.

Distribution of cell populations in (a) patients and (b) controls

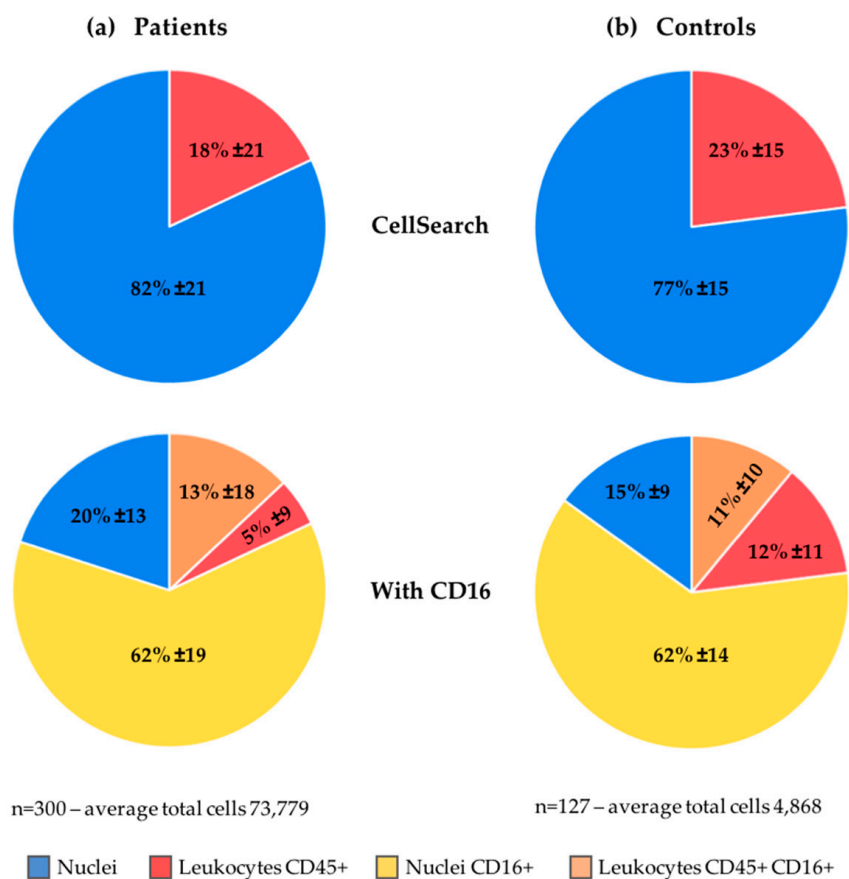


Figure 3. Cell population distribution in (a) patients and (b) controls using the CellSearch definition (upper) and using CD16 expression (lower) for further classification of the cells.

2.4. Increasing Leukocyte Identification by Adding CD16 Immunostaining

Whereas CD45 is basically present on all human leukocytes, it is most strongly expressed on lymphocytes and monocytes and expressed at a relatively low density in granulocytes. CD16 identifies the Fc III receptor present on granulocytes and natural killer cells. We added anti-CD16 labeled to PerCP to the CellSearch staining cocktail to improve the identification of granulocytes especially, using 300 blood samples from 192 NSCLC patients and 127 blood samples from 20 healthy volunteers. Thumbnail images of the nucleated cells identified by ACCEPT were fed into the trained deep learning network as described above. Figure 4 shows typical thumbnail images of nucleated events identified by ACCEPT using the presence of the different markers for classification. Classified CTC with expression of CD45 or CD16 can either be; (1) CTC nonspecifically binding to CD16 or CD45 antibodies; or (2) leukocytes non-specifically binding to the cytokeratin antibodies. The distribution of these cell populations in the patient and healthy volunteer samples is shown in Figure 3. From the 82% ± 21% unidentified nucleated cells in 192 NSCLC patients, 62% ± 19% are identified as granulocytes through CD16-expression. This decreases the number of unidentified cells to 20% ± 13% of the total nucleated cells present in the cartridge. The percentage of CD45-expressing leukocytes that also express CD16 is 13% ± 18%. These are most likely natural killer cells that express CD45 in higher densities compared to granulocytes. In 127 samples from 20 healthy volunteers, 77% ± 15% of the cells are unidentified which was decreased to 15% ± 9% when CD16 was used. The actual number of nucleated cells identified

by either CD45 or CD16 for patients and healthy volunteers is shown in Figure 2. The number of all classifications of nucleated cells was significantly larger in patients compared to healthy volunteers ($p < 0.001$). Figures 2 and 4 show there is still a large number of nucleated cells, especially in the patients, that remain unidentified.

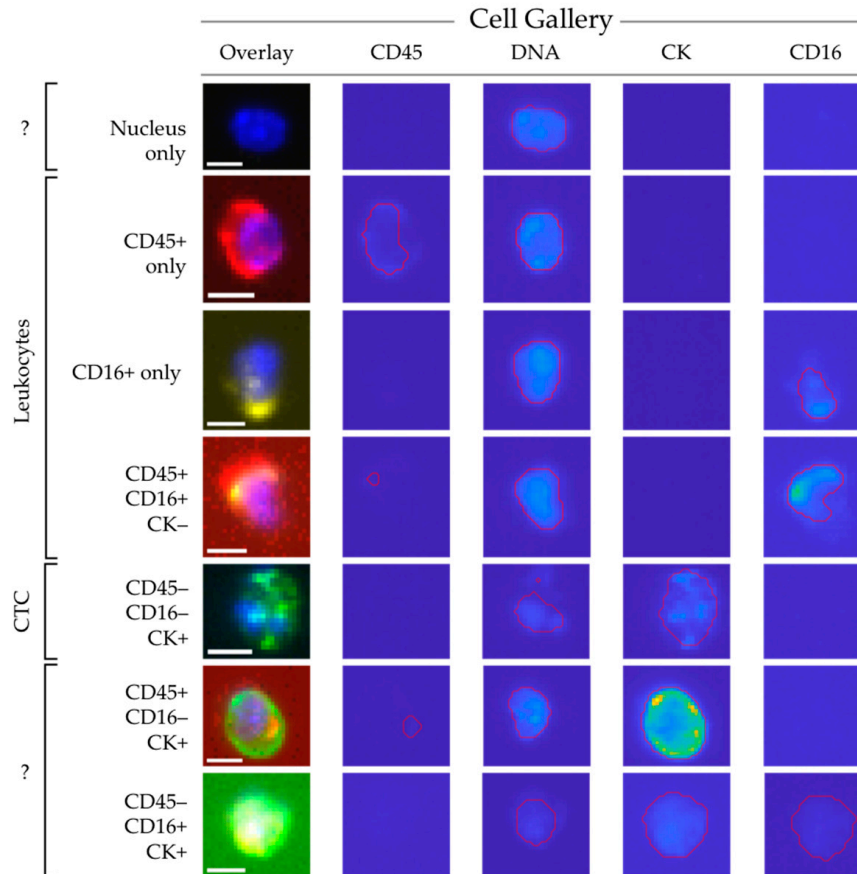


Figure 4. ACCEPT gallery showing cells from all classifications using the presence or absence of several markers. The scale bar is 10 pixels, representing 6.4 μm .

2.5. LED as a Light Source to Improve Excitation Efficiency of CD45-APC and CD16-PerCP

For five NSCLC patients and five healthy volunteers, cartridges were scanned first with the CellTracks—equipped with a mercury arc lamp as a light source—followed by scanning on a microscope equipped with an LED light source. The objectives and filter cubes were kept the same in both systems. The mean percentage of CD45+/CD16- leukocytes found by the CellTracks was $8\% \pm 7\%$ and this increased to $16\% \pm 9\%$ when a LED light source was used, as shown in Figure 5. More pronounced was the increase for CD45+/CD16+ leukocytes, which increased from $1\% \pm 0\%$ to $59\% \pm 25\%$ with the LED light source. Whereas CD16+/CD45- cells comprise a large population of the cells when a mercury arc lamp is used, these cells also appear to be CD45+ when the LED light source was used for the analysis; CD16+/CD45- leukocytes decreased from $58 \pm 17\%$ to $19\% \pm 15\%$. In total, the number of unidentified cells was reduced from $33\% \pm 14\%$ to $6\% \pm 5\%$ when an LED light source was used instead of a mercury arc lamp.

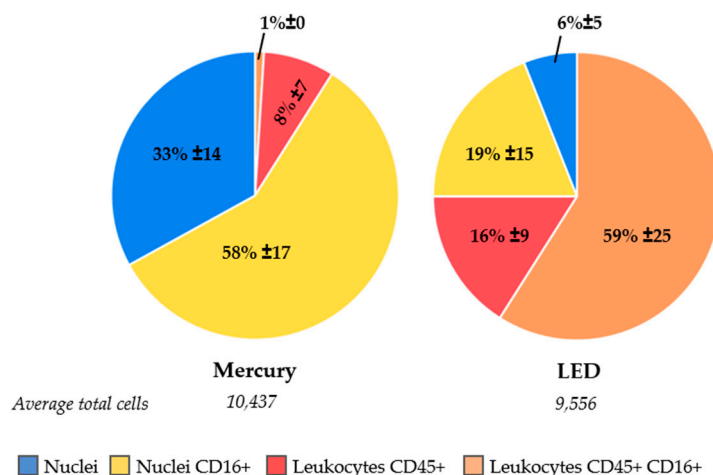


Figure 5. Improved identification of cells by comparison of cell population distributions analyzed with the mercury arc light source in CellSearch (left), followed by analysis with a LED light source on a separate microscope (right).

2.6. Identification of Unstained Nuclei by Adding Wheat Germ Agglutinin Immunostaining

In ten whole blood samples from five NSCLC patients and five healthy volunteers, the lectin wga conjugated to Alexa-Fluor488 was added to the CellSearch staining cocktail for the identification of cellular plasma membranes. In Figure S3, an ACCEPT gallery shows the staining of wga present on all cell classifications. The absence of wga on a nucleus, classifies this event as a “bare nucleus”. The presence of wga on already identified cells was 44% ± 23% for CD16+ granulocytes, 41% ± 20% for CD45+ leukocytes and 88% ± 15% for CD45+/CD16+ leukocytes, see Figure 6. One healthy volunteer sample was spiked with the cancer cell line MCF-7, and 100% of those spiked cells were stained with wga. In contrast, 52% ± 23% of the unidentified cell population was stained with wga, suggesting these are intact cells, whereas the remaining 48% ± 23% of these nuclei appear to be without a cell membrane. This suggests that these are simply bare nuclei, which cannot be identified through immunofluorescence staining.

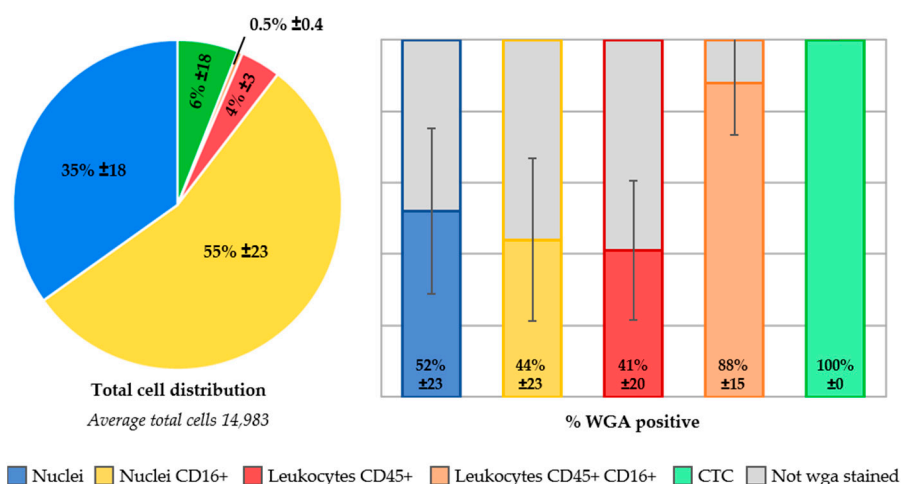


Figure 6. Distribution of cell populations in samples with wheat germ agglutinin (wga)-AlexaFluor488 staining (n = 10). The stacked bar represents the five cell populations present in these cartridges. One healthy volunteer sample was spiked with MCF-7 cancer cells, which are represented in the “CTC” category (6%). Of each population, the percentage of cells that stained positive for wga is displayed on the right side of the image, whereas the population remaining unstained with wga is visualized in grey.

3. Discussion

In this study, we investigated the cell populations present in 659 cartridges after immunomagnetic EpCAM enrichment by the CellSearch system, gathered from 162 healthy volunteers and 192 NSCLC patients. Stored fluorescent microscope images of the surface of the cartridges were analyzed by image analysis for the presence of different cell populations. The identity and number of cells present in the EpCAM-enriched cell suspension is not known, since the software only detects and presents DAPI+/CK+ events for review that are co-located. Any potential tumor cells or other interesting events remain obscured. We therefore explored the identity of all events present in the cartridge after EpCAM enrichment.

Images from the CellTracks microscope were analyzed with advanced imaging techniques. Most of the image analysis tools that were used are already assembled in the open source imaging program ACCEPT, including a tool to compare operator variability in scoring CTC [19,20]. In this image data set we had to deal with samples of low image quality, unequal distribution of the illumination, and extremely crowded images containing many clustered events, see Figure S2. A deep learning based approach appeared to be a good solution since it could easily be adapted to a certain type of signal to segment by feeding it with enough training data. The deep learning based segmentation algorithm was programmed in Python[®] using the Keras [21] framework and is not yet linked to the ACCEPT toolbox, which is developed in MATLAB[®]. To be able to use the deep learning based segmentation, we exported thumbnails identified by ACCEPT and used them as input for the deep learning segmentation network. After segmentation of the thumbnails by the network, the segmentation results, together with the original thumbnails, were fed back into the ACCEPT toolbox for further analysis. For samples with thousands of nucleated events (up to 300,000 with five fluorescent channels each), this became quite a computationally intensive task. Therefore, for the current analysis we restricted that part of our analysis to only 10% of the nucleated events which were randomly selected. Experience obtained with the deep learning based classification algorithm we developed showed very promising results and the new segmentation method also showed that it is often superior to the more traditional segmentation approach currently used in ACCEPT [22]. Thus, in the future, we would like to incorporate a semantic deep learning segmentation into the ACCEPT toolbox to automatically detect and classify CTC, tumor derived extra cellular vesicles, leukocytes, and other cell populations for which reagents are added.

In total, 300 samples obtained from 192 metastatic lung cancer patients, 232 samples obtained from 142 patients with benign tumors, and 127 samples obtained from 20 healthy volunteers were investigated for the presence of nucleated events. The number of nucleated events was significantly higher in cancer patients as compared to the healthy donors and patients with benign disease, see Figure 2 and Figure S1, and further studies are needed to explore the potential clinical utility of this finding. CTC were detected by traditional CellSearch scoring in a small portion of the patient samples and in these samples the CTC number ranged from 1 to 186 cells. This low number of cells could not account for the large differences in the detected nucleated events.

Knowledge of the composition of these nucleated cells is needed to be able to hypothesize about this difference. Most likely, these nucleated events originate from hematopoietic cells that are captured along with the EpCAM enrichment in CellSearch. However, the fraction of nucleated cells on which the leukocyte marker CD45 was detected was remarkably low both in healthy volunteers and metastatic cancer patients, as shown in Figure 3. Since the excitation of CD45-APC with a mercury arc lamp used in the CellTracks system is not optimal, we first looked into a way to increase the detection of granulocytes especially, which express the CD45 antigen in relatively low antigen density compared to leukocytes.

We evaluated this by adding anti-CD16-PerCP antibody to the immunostaining in the CellSearch system. All 300 samples from metastatic lung cancer patients and 127 samples from healthy volunteers were processed with this extra marker. The leukocytes and granulocytes that were identified indeed increased, as shown in Figure 3. However, a considerable number of nucleated cells remained unidentified. Next, we looked into the role of ferrofluids, which are used at a concentration of 40 µg/mL

in CellSearch and are present on the imaging surface of the cartridge [23,24]. The ferrofluids reduce the intensity of the fluorescence; an effect which is most obvious in the blue region (400–460 nm). Therefore, we looked into an improvement of the excitation of the fluorophores. For this, a microscope equipped with an LED light source to improve the APC fluorophore excitation was used [12]. This resulted in a clear decrease of unidentified nucleated cells, see Figure 5. Also, to determine whether these remaining nuclei were enclosed by a cellular plasma membrane, we added wheat germ agglutinin to the CellSearch cocktail [13,14]. This showed that part of the unidentified nucleated cells did not stain with wga, suggesting they are “bare”-like nuclei, see Figure 6.

Whether or not these bare nuclei were actually present in the blood, or if they represent cells that have been damaged through the immunomagnetic selection or staining process, remains an outstanding question. However, a large fraction of leukocytes identified with the current immunostaining are not also stained with wga. This could suggest that the wga-staining assay is not sufficiently sensitive to stain all membranes or that the imaging method is not sufficiently sensitive to detect dim wga fluorescence. A high concentration of wga added to the CellSearch cocktail caused the aggregation of blood cells, and therefore a lower concentration was chosen for this assay. It might, however, be possible that a higher concentration of wga is necessary to locate all membranes present in the blood sample. Also, all samples were analyzed on the CellTracks, but improved identification will be observed when the LED light source is used. Another option is that the leukocyte membranes are damaged to a point that the agglutinin can no longer bind to the residues on the membrane. This may be caused during the 15 min permeabilization process to open the membranes. The CellSearch permeabilization agent is saponin, a detergent which interacts with membrane cholesterol by selectively removing them to form pores, which enables antibodies to enter [25]. It is notable, that the fraction of the spiked MCF-7 cancer cells that were stained with wga was 100%. Compared to cancer cell line cells, the leukocytes are possibly more fragile and susceptible to damage caused during permeabilization.

The wga+/DNA+/CD45-/CD16-/CK- cells are potentially CTC that are not expressing CK for CellSearch detection. However, our observation that a similar population was detected in the blood of healthy volunteers makes this explanation less likely. Confirmation of whether or not CTC are among the unidentified nucleated cells will likely have to come from DNA or RNA assays. Removing the cells from the cartridge might allow them to be selected based on their absence of fluorescence markers and subsequently sorted with fluorescence-activated cell sorting (FACS). Quantitative PCR after DNA isolation and amplification could reveal the presence of mutations, indicating one of the cells is, in fact, a tumor cell. Digital-droplet PCR has a higher sensitivity and could be applied on the entire cell suspension from a cartridge, although the multiplex analysis of mutations is less feasible in this assay. With these approaches, the number of tumor cells will have to be compared to the number of CTC scored on the CellTracks (if present at all) to determine if more or an equal number of tumor cells are present in the cell suspension. Also, analysis of mRNA after FACS isolation might be feasible for the unidentified cells with a cellular membrane present. Quantification of a selection of proteins can indicate if the cells are of hematopoietic or epithelial lineage. However, the permeabilization of cells in the CellSearch might complicate this approach. Therefore, the CellSearch Epithelial Cell Profile Kit should be used. This kit is designed for RNA analysis and enriches EpCAM+ cells without subsequent permeabilization and staining.

After taking several steps to determine the origin of unidentified cells, the possible identity for the cells that remained unidentified thus far might be:

1. Hematopoietic cells not identified through damage incurred during the procedure or inaccessible antigens through the ferrofluids [23,24];
2. Hematopoietic cells without sufficient expression of CD45 and CD16: This suggests they would need additional CD markers for identification or an improved labeling method that amplifies low signals, which might yield increased detection of very dim stained cells and separate the fluorescent signal of densely packed cells [11,26–28]. Early myeloid cells have recently been

observed to surround the tumor in high numbers, and as these cells do not yet express CD16, this remains a possibility [29];

3. CTC with no or low expression of the CK antigens detected by the C11 and A53.B/A2 clones: Since these clones only recognize a subset of the CK present in a cell, it might be beneficial to include antibody clones that recognize all CK. Previously, we have shown that adding several CK clones to the CellSearch antibody cocktail improved the detection of CTC positive patients by 11% [30]. Also, it might be possible that EpCAM+/CK- CTC are present, remaining undetected because of the downregulation of epithelial markers through epithelial-to-mesenchymal transition [31–33]. In order to detect these cells, antibodies specific to this process could be added to the CellSearch immunostaining [34–36];
4. Cells of other origin: Such as circulating stromal, endothelial, or stem cells [37]. Detection of these cells would also require the addition of other antibodies to the assay.

Therefore, there is still much to be achieved toward the identification of the cells that remain anonymous. It is most important to determine if all tumor cells have been located or if some are still missed.

4. Materials and Methods

4.1. Cancer Patients, Patients with Benign Disease, and Healthy Volunteers

Peripheral blood samples were drawn by venipuncture into 10 mL CellSave Preservative Tubes (Menarini Silicon Biosystems, Huntingdon Valley PA, USA) from healthy volunteers and metastatic NSCLC patients. A total of 300 blood samples from 192 NSCLC patients were obtained after patients with advanced NSCLC provided written informed consent. The medical ethical committee of the University Medical Center Groningen and Board of Directors approved the protocol (project number 288.194/7 April 2014; Groningen, The Netherlands). Blood draw from patients occurred before and/or during their treatment. A total of 127 blood samples from 20 healthy volunteers aged 20–55 were obtained from the donor services of the Experimental Centre for Technical Medicine at the University of Twente, and all donors gave written informed consent before donating blood. All healthy volunteer samples were processed 1 day after blood draw, whereas patient samples were processed after 1–4 days. Digitally stored images from 232 samples from 142 patients with benign disease processed with CellSearch from previous studies were included for ACCEPT analysis [10,38].

4.2. Cell Lines and Spiking

Whole blood samples from healthy volunteers were spiked with cancer cell lines PC3, NCI-H460, and MCF-7. PC3 (1.0×10^4 EpCAM antigens) and NCI-H460 (1.4×10^2 EpCAM antigens) were used for frequency detection of CTC with ACCEPT and MCF-7 (8.8×10^5 EpCAM antigens) was used for testing wheat germ agglutinin (wga) staining of cancer cells. This cell line was chosen for its high EpCAM expression to ensure sufficient capture of cancer cells in the cartridge for analysis. All cell lines were obtained from ATCC (Manassas, VA, USA) and have not been authenticated in the past six months. The cells were grown at 37 °C and 5% CO₂ and cultured in RPMI-1640 (Sigma-Aldrich, St. Louis, MO, USA). The culture media was supplemented with L-Glutamine (Sigma), 10% fetal bovine serum (Gibco, Invitrogen, Carlsbad, CA, USA), and 1% penicillin-streptomycin (Gibco, Invitrogen, Carlsbad, CA, USA).

4.3. Processing Blood with CellSearch

Aliquots of 7.5 mL of blood were processed with the CellSearch[®] system (Menarini Silicon Biosystems, Castel Maggiore (BO), Italy) for CTC detection. CellSearch analysis was performed within 96 h after blood was drawn. Antibodies directed against the epithelial cell adhesion antigen (EpCAM) coupled to ferrofluids were used to enrich CTC from the background of leukocytes. The enriched cells were fluorescently labeled with the CellSearch CTC kit (Menarini Silicon Biosystems)

using the nucleic acid dye 4′6-diaminodino-2-phenylindole (DAPI) for DNA staining, anti-cytokeratin monoclonal antibodies (mAbs) C11 and A53.B/A2 labelled with phycoerythrin (PE), and anti-CD45 mAb (clone HI30) labeled with allophycocyanin (APC) for recognizing leukocytes. Peridinin Chlorophyll A Protein (PerCP) labeled mAb anti-CD16 (clone 3G8, Biolegend, San Diego, CA, USA) directed against granulocytes, and Alexa-Fluor488 conjugated to the lectin wga (Thermo Fisher Scientific, Waltham, MA, USA) were added to the extra marker position in the CellSearch Epithelial Cell kit with an end concentration of 2 µg/mL and 3 µg/mL, respectively. This extra marker channel was used for the measurement of PerCP and DIOC on the CellTracks Analyzer II, which generates images of the complete cartridge for all channels.

4.4. Image Acquisition

Fluorescent microscopy images were acquired using a modified CellTracks Analyzer II, which included a 20× and 40× objective and in total six filter cubes next to the standard instruments. Besides the acquisition of the standard channels DAPI, PE, DIOC, and APC, the images with PerCP staining were acquired with a customized PerCP filter cube (excitation 435/40 nm; dichroic 510 nm; emission 676/29 nm (Semrock, Rochester, NY, USA)) using an integration time of 400 ms. When wga was used in the immunostaining, the integration time of the DIOC filter was also 400 ms. To determine if another light source detects weakly stained leukocytes better than the mercury arc lamp of the CellTracks, ten cartridges were automatically scanned with an inverted Nikon Ti-E microscope with filter cube and objective changer (Nikon, Tokyo, Japan). The light source of this microscope is the Lumencor Sola SE II 365 LED. A 10× (0.45NA) objective (Nikon, Tokyo, Japan), a computer-controlled CCD camera (C114400, Orca Flash 4.0LT, Hamamatsu Photonics, Hamamatsu, Japan), an ASI MS-2500 XY stage and filters optimized for LED illumination were used for the acquisition of fluorescent images of the cartridge. The following filters (all from Semrock) were used: DAPI with excitation 377/50 nm, dichroic 409 nm, emission 409 nm/LP; fluorescein isothiocyanate (FITC) with excitation 482/18 nm, dichroic 495 nm, emission 520/28 nm; PE with excitation 543/22 nm, dichroic 562 nm, emission 593/40 nm; APC with excitation 635/18 nm, dichroic 652 nm, emission 680/42 nm; and PerCP with excitation 435/40 nm, dichroic 510 nm, emission 676/29 nm. The integration times were the same as used in the CellTracks: For DAPI 30 ms, FITC 100 ms, PE 200 ms, APC 600 ms, and for PerCP 400 ms. The scanning and image acquisition was controlled by a software program written in Labview (National Instruments, Austin, TX, USA).

4.5. Image Analysis with ACCEPT

In CellSearch, CTC are annotated when the objects are stained with DAPI and CK, but lack CD45 staining and are larger than 4 µm with morphological features consistent with those of a cell [10,38]. In the ACCEPT (Automated CTC Classification, Enumeration, and PhenoTyping) toolbox, every event in the images is analyzed using an advanced multi-scale segmentation approach and several intensity and shape measurements are extracted for every event found [19,39]. ACCEPT is developed in the EU funded “CANCER-ID” and “CTC-Trap” programs, as an open source toolbox for CTC analysis and can be downloaded from <https://github.com/LeonieZ/ACCEPT>. New tools and features for ACCEPT are being developed and made available in the downloadable version at regular intervals.

For this study, the fluorescent images from all samples were uploaded, and all events present in the images were detected. To enhance the detection of cells, especially in areas where the cells formed large cell clusters, we added a background subtraction to the ACCEPT toolbox. Before images are segmented, a Fourier filtering is applied, and large and very fine scales are filtered by multiplying the filtered signal with two Gaussian kernels before the signal is back-transformed by an inverse Fourier transformation. In this way, the background is removed, which enhances the detection in very crowded areas, as well as areas with inhomogeneous illumination.

Fragmented fluorescent signals make a segmentation with an intensity-based segmentation method difficult, which is now implemented in the downloadable version of the ACCEPT toolbox.

In this study this was especially observed in the fluorescent images obtained from the CD16-PerCP stained cells. Therefore, we decided to explore deep learning approaches to determine if this limitation could be overcome. For this approach, all thumbnails containing nucleated cells in the cartridge were found using ACCEPT, and a random selection of these thumbnails was used for further evaluation. Per sample, we exported 10% of the observed cells, but minimally 1000 (if the total amount was high enough) and maximally 5000 thumbnails. Moreover, we exported 759 thumbnails with five fluorescent channels each where the segmentation method of ACCEPT gave satisfactory results. We used these thumbnails (split into single channels) to train a deep learning network to segment the cells. The implementation was done in Keras and we used the U-net architecture with skip connections [21,40]. The set of 3796 thumbnail images was split into a training set of 3037 images (80%) and a validation set of 759 images (20%). The images were preprocessed in the following way: Each thumbnail exported from the ACCEPT toolbox was transformed into a thumbnail of 80×80 pixels by adding a padding at all four sides with the median intensity of the pixels at the image boundary. Moreover, all images scanned with CellTracks were cast from double to uint8 arrays by first dividing by 4095 (the maximum value in Celltracks) and then multiplying by 255 and converting the results to uint8. Images scanned with the inverted Nikon Ti-E microscope were recorded as uint16, thus having a theoretical maximum value of 65,535. However, during our analysis, we noticed that this range was not used and that the range of intensity values used strongly differs between the fluorescent channels. Therefore, we calculated for each fluorescent channel the intensity value that was higher than 99.9% of the measured intensities over all images in all samples. The images were then divided by the resulting intensity per channel, multiplied by 255 and cast to uint8. By taking the value that is above 99.9% instead of the maximum, we accounted for individual saturated pixels where the incident light caused the camera sensor to respond with the highest possible value. To train the deep learning network, we used a batch size of 16 and the network was trained for ten epochs. The resulting convolutional network was then used to segment the exported thumbnails for each sample. The original thumbnails together with the segmentation derived by deep learning were fed back into the ACCEPT algorithm by using the marker characterization processor for further analysis and visualization of results.

4.6. Statistical Analysis

Statistical analysis was performed using SPSS (version 24, SPPS Inc., Chicago, IL, USA). A p -value below 0.05 was considered significant. The non-parametric Mann-Whitney U test with significance level $\alpha = 0.05$ was used to determine if there was a difference in nucleated cell numbers between cohorts of patients and controls. The total nucleated cell amounts were log-transformed for normality. With multiple regression analysis, we determined the association of several variables with the total nucleated cell count, such as age, treatment, CTC score, and sample age before processing. The confounding effect of each variable was also determined.

5. Conclusions

Advanced image analysis was used to identify the cell populations present after immunomagnetic EpCAM enrichment by the CellSearch system from 659 cartridges from NSCLC patients. A significantly larger number of nucleated cells were found, from which the origin could not be identified, compared to samples from healthy volunteers and patients with benign disease. A large portion of the unidentified nucleated cells could be identified by: Adding CD16 to improve the identification of granulocytes; the use of an LED light source to improve fluorescence detection and; the addition of wga to discriminate between bare nuclei and cells. However, there is still much to gain in the identification of the cells that still remain anonymous. These observations must also be confirmed in patients with other cancers. The question remains: What are these cells? Going forward, it is most important to determine if all tumor cells have been located or if some are still missed and reside in this unidentified cell population. This unidentified CTC subpopulation could be of interesting clinical value and remains

thus far unexploited. Therefore, further studies are needed to determine if EpCAM+/CK- CTC are present among the nucleated cells of which the origin is not known.

Supplementary Materials: The following are available online at <http://www.mdpi.com/2072-6694/10/10/377/s1>, Figure S1: Nucleated cell count in controls with benign disease and NSCLC patients, Figure S2: Illustration of the issues that influence the segmentation quality of the ACCEPT toolbox, Figure S3: ACCEPT gallery of cells stained with the membrane lectin wheat germ agglutinin (wga).

Author Contributions: Conceptualization, S.d.W., L.L.Z. and L.W.M.M.T.; Methodology, S.d.W., L.L.Z. and L.W.M.M.T.; Software, L.L.Z. and G.v.D.; Validation, S.d.W. and L.L.Z.; Formal Analysis, S.d.W., L.L.Z. and G.v.D.; Investigation, S.d.W.; Resources patients, T.J.N.H. and H.J.M.G.; Data Curation, S.d.W.; Writing-Original Draft Preparation, S.d.W., L.L.Z. and L.W.M.M.T.; Writing-Review & Editing, S.d.W., L.L.Z., T.J.N.H., H.J.M.G., G.v.D. and L.W.M.M.T.; Visualization, S.d.W. and L.L.Z.; Funding Acquisition, L.W.M.M.T.

Funding: This research was funded by the EU FP7 “CTC-Trap” grant number 305341, the EU IMI “CANCER-ID” grant number 115749-1 and through a research and collaborative agreement with Janssen Diagnostics.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, and in the decision to publish the results.

References

- Racila, E.; Euhus, D.; Weiss, A.J.; Rao, C.; McConnell, J.; Terstappen, L.W.M.M.; Uhr, J.W. Detection and characterization of carcinoma cells in the blood. *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 4589–4594. [[CrossRef](#)] [[PubMed](#)]
- Cohen, S.J.; Punt, C.J.A.; Iannotti, N.; Saidman, B.H.; Sabbath, K.D.; Gabrail, N.Y.; Picus, J.; Morse, M.; Mitchell, E.; Miller, M.C.; et al. Relationship of Circulating Tumor Cells to Tumor Response, Progression-Free Survival, and Overall Survival in Patients With Metastatic Colorectal Cancer. *J. Clin. Oncol.* **2008**, *26*, 3213–3221. [[CrossRef](#)] [[PubMed](#)]
- Cristofanilli, M.; Budd, G. Circulating tumor cells, disease progression, and survival in metastatic breast cancer. *N. Engl. J. Med.* **2004**, *351*, 781–791. [[CrossRef](#)] [[PubMed](#)]
- De Bono, J.S.; Scher, H.I.; Montgomery, R.B.; Parker, C.; Miller, M.C.; Tissing, H.; Doyle, G.V.; Terstappen, L.W.W.M.; Pienta, K.J.; Raghavan, D. Circulating tumor cells predict survival benefit from treatment in metastatic castration-resistant prostate cancer. *Clin. Cancer Res.* **2008**, *14*, 6302–6309. [[CrossRef](#)] [[PubMed](#)]
- Hiltermann, T.J.N.; Pore, M.M.; Van den Berg, A.; Timens, W.; Boezen, H.M.; Liesker, J.J.W.; Schouwink, J.H.; Wijnands, W.J.A.; Kerner, G.S.M.A.; Kruijt, F.A.E.; et al. Circulating tumor cells in small-cell lung cancer: A predictive and prognostic factor. *Ann. Oncol.* **2012**, *23*, 2937–2942. [[CrossRef](#)] [[PubMed](#)]
- Krebs, M.G.; Sloane, R.; Priest, L.; Lancashire, L.; Hou, J.M.; Greystoke, A.; Ward, T.H.; Ferraldeschi, R.; Hughes, A.; Clack, G.; et al. Evaluation and Prognostic Significance of Circulating Tumor Cells in Patients With Non-Small-Cell Lung Cancer. *J. Clin. Oncol.* **2011**, *29*, 1556–1563. [[CrossRef](#)] [[PubMed](#)]
- Hiraiwa, K.; Takeuchi, H.; Hasegawa, H.; Saikawa, Y.; Suda, K.; Ando, T.; Kumagai, K.; Irino, T.; Yoshikawa, T.; Matsuda, S.; et al. Clinical significance of circulating tumor cells in blood from patients with gastrointestinal cancers. *Ann. Surg. Oncol.* **2008**, *15*, 3092–3100. [[CrossRef](#)] [[PubMed](#)]
- Rao, C.; Bui, T.; Connelly, M.; Doyle, G.; Karydis, I.; Middleton, M.R.; Clack, G.; Malone, M.; Coumans, F.A.W.; Terstappen, L.W.M.M. Circulating melanoma cells and survival in metastatic melanoma. *Int. J. Oncol.* **2011**, *38*, 755–760. [[CrossRef](#)] [[PubMed](#)]
- Gazzaniga, P.; Gradilone, A.; De Berardinis, E.; Busetto, G.M.; Raimondi, C.; Gandini, O.; Nicolazzo, C.; Petracca, A.; Vincenzi, B.; Farcomeni, A.; et al. Prognostic value of circulating tumor cells in nonmuscle invasive bladder cancer: A CellSearch analysis. *Ann. Oncol.* **2012**, *23*, 2352–2356. [[CrossRef](#)] [[PubMed](#)]
- Allard, W.J.; Matera, J.; Miller, M.C.; Repollet, M.; Connelly, M.C.; Rao, C.; Tibbe, A.G.J.; Uhr, J.W.; Terstappen, L.W.M.M. Tumor Cells Circulate in the Peripheral Blood of All Major Carcinomas but not in Healthy Subjects or Patients with Nonmalignant Diseases. *Clin. Cancer Res.* **2004**, *10*, 6897–6904. [[CrossRef](#)] [[PubMed](#)]
- Van Dalum, G. The Role and Definition of Cancer Cells in Blood. Ph.D. Thesis, University of Twente, Enschede, The Netherlands, 2015.
- Lumencore, I. SOLA Light Engine. Available online: <http://lumencor.com> (accessed on 21 August 2018).

13. Chazotte, B. Labeling Membrane Glycoproteins or Glycolipids with Fluorescent Wheat Germ Agglutinin. *Cold Spring Harb. Protoc.* **2011**, *2011*. [CrossRef] [PubMed]
14. Legant, W.R.; Shao, L.; Grimm, J.B.; Brown, T.A.; Milkie, D.E.; Avants, B.B.; Lavis, L.D.; Betzig, E. High-density three-dimensional localization microscopy across large volumes. *Nat. Methods* **2016**, *13*, 359–365. [CrossRef] [PubMed]
15. Solórzano, C.; Bouquelet, S.; Pereyra, M.A.; Blanco-Favela, F.; Slomianny, M.C.; Chavez, R.; Lascrain, R.; Zenteno, E.; Agundis, C. Isolation and characterization of the potential receptor for wheat germ agglutinin from human neutrophils. *Glycoconj. J.* **2006**, *23*, 591–598. [CrossRef] [PubMed]
16. Sharon, N.; Lis, H. *Lectins*; Springer: New York, NY, USA, 2007; ISBN 978-1-4020-6953-6.
17. Emde, B.; Heinen, A.; Gödecke, A.; Bottermann, K. Wheat germ agglutinin staining as a suitable method for detection and quantification of fibrosis in cardiac tissue after myocardial infarction. *Eur. J. Histochem.* **2014**, *58*, 2448. [CrossRef] [PubMed]
18. Adair, W.L.; Kornfeld, S. Isolation of the receptors for wheat germ agglutinin and the Ricinus communis lectins from human erythrocytes using affinity chromatography. *J. Biol. Chem.* **1974**, *249*, 4696–4704.
19. Zeune, L.L.; Van Dalum, G.; Terstappen, L.W.M.M.; Van Gils, S.A.; Brune, C. Multiscale Segmentation via Bregman Distances and Nonlinear Spectral Analysis. *SIAM J. Imaging Sci.* **2017**, *10*, 111–146. [CrossRef]
20. Zeune, L.L.; De Wit, S.; Berghuis, A.M.S.; IJzerman, M.J.; Terstappen, L.W.M.M.; Brune, C. How to agree on a CTC: Evaluating the Consensus in Circulating Tumor Cell Scoring. *Cytom. Part A* **2018**. [CrossRef] [PubMed]
21. Chollet, F. Keras. 2015. Available online: <https://keras.io> (accessed on 21 August 2018).
22. Zeune, L.L.; De Wit, S.; Van Dalum, G.; Nanou, A.; Andree, K.C.; Swennenhuis, J.F.; Terstappen, L.W.M.M.; Brune, C. Deep Learning to Identify Circulating Tumor Cells by ACCEPT. In Proceedings of the ISMRC, Montpellier, France, 3–5 May 2018.
23. Scholtens, T.M. Automated Classification and Enhanced Characterization of Circulating Tumor Cells by Image Cytometry. Ph.D. Thesis, University of Twente, Enschede, The Netherlands, 2012.
24. Coumans, F.A.W.; Terstappen, L.W.M.M. Detection and Characterization of Circulating Tumor Cells by the CellSearch Approach. *Methods Mol. Biol.* **2015**, *1347*, 263–278. [CrossRef] [PubMed]
25. Jamur, M.C.; Oliver, C. Permeabilization of Cell Membranes. In *Methods in Molecular Biology (Clifton, N.J.)*; Humana Press: New York, NY, USA, 2010; Volume 588, pp. 63–66.
26. Faget, L.; Hnasko, T.S. Tyramide Signal Amplification for Immunofluorescent Enhancement. In *Methods in Molecular Biology (Clifton, N.J.)*; Humana Press: New York, NY, USA, 2015; Volume 1318, pp. 161–172.
27. Zhou, S.; Yuan, L.; Hua, X.; Xu, L.; Liu, S. Signal amplification strategies for DNA and protein detection based on polymeric nanocomposites and polymerization: A review. *Anal. Chim. Acta* **2015**, *877*, 19–32. [CrossRef] [PubMed]
28. Gustafsdottir, S.M.; Schallmeiner, E.; Fredriksson, S.; Gullberg, M.; Söderberg, O.; Jarvius, M.; Jarvius, J.; Howell, M.; Landegren, U. Proximity ligation assays for sensitive and specific protein analyses. *Anal. Biochem.* **2005**, *345*, 2–9. [CrossRef] [PubMed]
29. Calcinotto, A.; Spataro, C.; Zagato, E.; Di Mitri, D.; Gil, V.; Crespo, M.; De Bernardis, G.; Losa, M.; Mirenda, M.; Pasquini, E.; et al. IL-23 secreted by myeloid cells drives castration-resistant prostate cancer. *Nature* **2018**, *559*, 363–369. [CrossRef] [PubMed]
30. De Wit, S.; Van Dalum, G.; Lenferink, A.T.M.; Tibbe, A.G.J.; Hiltermann, T.J.N.; Groen, H.J.M.; Van Rijn, C.J.M.; Terstappen, L.W.M.M. The detection of EpCAM+ and EpCAM- circulating tumor cells. *Sci. Rep.* **2015**, *5*, 12270. [CrossRef] [PubMed]
31. Serrano, M.J.; Ortega, F.G.; Alvarez-Cubero, M.J.; Nadal, R.; Sanchez-Rovira, P.; Salido, M.; Rodríguez, M.; García-Puche, J.L.; Delgado-Rodríguez, M.; Solé, F.; et al. EMT and EGFR in CTCs cyokeratin negative non-metastatic breast cancer. *Oncotarget* **2014**, *5*, 7486–7497. [CrossRef] [PubMed]
32. Willipinski-Stapelfeldt, B.; Riethdorf, S.; Assmann, V.; Woelfle, U.; Rau, T.; Sauter, G.; Heukeshoven, J.; Pantel, K. Changes in cytoskeletal protein composition indicative of an epithelial-mesenchymal transition in human micrometastatic and primary breast carcinoma cells. *Clin. Cancer Res.* **2005**, *11*, 8006–8014. [CrossRef] [PubMed]
33. Krebs, M.G.; Hou, J.M.; Sloane, R.; Lancashire, L.; Priest, L.; Nonaka, D.; Ward, T.H.; Backen, A.; Clack, G.; Hughes, A.; et al. Analysis of circulating tumor cells in patients with non-small cell lung cancer using epithelial marker-dependent and -independent approaches. *J. Thorac. Oncol.* **2012**, *7*, 306–315. [CrossRef] [PubMed]

34. Kallergi, G.; Papadaki, M.A.; Politaki, E.; Mavroudis, D.; Georgoulas, V.; Agelaki, S. Epithelial to mesenchymal transition markers expressed in circulating tumour cells of early and metastatic breast cancer patients. *Breast Cancer Res.* **2011**, *13*, R59. [[CrossRef](#)] [[PubMed](#)]
35. Armstrong, A.J.; Marengo, M.S.; Oltean, S.; Kemeny, G.; Bitting, R.L.; Turnbull, J.D.; Herold, C.I.; Marcom, P.K.; George, D.J.; Garcia-Blanco, M.A. Circulating Tumor Cells from Patients with Advanced Prostate and Breast Cancer Display Both Epithelial and Mesenchymal Markers. *Mol. Cancer Res.* **2011**, *9*, 997–1007. [[CrossRef](#)] [[PubMed](#)]
36. Wu, S.; Liu, S.; Liu, Z.; Huang, J.; Pu, X.; Li, J.; Yang, D.; Deng, H.; Yang, N.; Xu, J. Classification of Circulating Tumor Cells by Epithelial-Mesenchymal Transition Markers. *PLoS ONE* **2015**, *10*, e0123976. [[CrossRef](#)] [[PubMed](#)]
37. Chen, Y.; Li, P.; Huang, P.H.; Xie, Y.; Mai, J.D.; Wang, L.; Nguyen, N.T.; Huang, T.J. Rare cell isolation and analysis in microfluidics. *Lab Chip* **2014**, *14*, 626–645. [[CrossRef](#)] [[PubMed](#)]
38. Van Dalum, G.; Stam, G.J.; Scholten, L.F.A.; Mastboom, W.J.B.; Vermes, I.; Tibbe, A.G.J.; De Groot, M.R.; Terstappen, L.W.M.M. Importance of circulating tumor cells in newly diagnosed colorectal cancer. *Int. J. Oncol.* **2015**, *46*, 1361–1368. [[CrossRef](#)] [[PubMed](#)]
39. Zeune, L.L.; Van Dalum, G.; Decraene, C.; Proudhon, C.; Fehm, T.; Neubauer, H.; Rack, B.; Alunni-Fabroni, M.; Terstappen, L.W.M.M.; Van Gils, S.A.; Brune, C. Quantifying HER-2 expression on circulating tumor cells by ACCEPT. *PLoS ONE* **2017**, *12*, e0186562. [[CrossRef](#)] [[PubMed](#)]
40. Ronneberger, O.; Fischer, P.; Brox, T. *U-Net: Convolutional Networks for Biomedical Image Segmentation*; Springer International Publishing: Cham, Switzerland, 2015; pp. 234–241.



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).