Article

Not peer-reviewed version

Cancer-associated Systemic Abnormality: A Novel Marker Panel for Early-stage Breast Cancer Characterization

Stefan Schreier , Prapaphan Budchart , Suparerk Borwornpinyo , Lakkana Adireklarpwong , Prakasit Chirappapha , Wannapong Triampo , Panuwat Lertsithichai *

Posted Date: 12 February 2024

doi: 10.20944/preprints202402.0681.v1

Keywords: early-stage breast cancer; circulating tumor cells; rare cells, liquid biopsy; systemic cancer; cytomorphology

Article

Cancer-Associated Systemic Abnormality: A Novel Marker Panel for Early-Stage Breast Cancer Characterization

Stefan Schreier [1,2,3], Prapaphan Budchart [3], Suparerk Borwornpinyo [3,4], Lakkana Adireklarpwong [5], Prakasit Chirappapha [5], Wannapong Triampo [2,6] and Panuwat Lertsithichai [5,*]

1 School of Bioinnovation and Bio-based Product Intelligence, Faculty of Science, Mahidol University, Rama VI Rd, Bangkok 10400, Thailand
2 Thailand Center of Excellence in Physics, Ministry of Higher Education, Science, Research and Innovation, 328 Si Ayutthaya Road, Bangkok 10400, Thailand
3 Premise Biosystems Co., Ltd. Bangkok 10540, Thailand
4 Excellent Center for Drug Discovery, Faculty of Science, Mahidol University, Rama VI, Rd, Bangkok, 10400, Thailand
5 Department of Surgery, Faculty of Medicine Ramathibodi Hospital, Mahidol University, Bangkok, 10400, Thailand
6 Department of Physics, Faculty of Science, Mahidol University, Bangkok 10400, Thailand
* Correspondence: panuwat.ler@mahidol.ac.th

Abstract: Background: Circulating rare cells participate in breast cancer evolution as systemic components of the disease and thus, are a source of theranostic information. Exploration of cancer-associated rare cells is in its infancy. We determined the validity of a rare-cell detection platform and demonstrated the usefulness of a certain rare cell subset as a novel approach to characterize the breast cancer system. Methods: Linearity of the Rarmax® platform was established using a spike-in approach. The platform includes red blood cell lysis, leukocyte depletion and high-resolution fluorescence image recording. Rare cell analysis was conducted on 28 samples (before and after surgery) from 14 patients with breast cancer, 20 healthy volunteers and 9 non-cancer control volunteers. In-depth identification of rare cells, including circulating tumor cells, endothelial-like cells, erythroblasts and inflammation-associated cells, was performed using a phenotype and morphology-based classification system. Results: The platform performed linearly over a range of 5-950 spiked cells, with an average recovery of 84.6%. Circulating epithelial and endothelial-like cell subsets have been demonstrated to be associated with or independent of cancer with tumor presence. Furthermore, certain cell patterns may be associated with treatment-related adverse effects. The sensitivity in detecting tumor-presence and cancer-associated abnormality before surgery was 50% and 85.7%, respectively and the specificity was 100% and 96.6%, respectively. Conclusion: The present study supports the idea of a cancer-associated rare cell abnormality to represent tumor entities as well as systemic cancer. The latter is independent of the apparent clinical cancer.

Keywords: early-stage breast cancer; circulating tumor cells; rare cells; liquid biopsy; systemic cancer; cytomorphology

1. Introduction

Early-stage breast cancer patients undergo treatment with curative intent that commonly includes surgery in combination with adjuvant radiotherapy, chemotherapy, targeted therapy as well as endocrine therapy, when indicated [1]. Although treatment outcomes are generally excellent, a small yet significant fraction in the patient group suffers from cancer recurrence in the long-term. Recurrence rates after curative treatment for breast cancer in long term follow up greater 5 years may

 doi:10.20944/preprints202402.0681.v1

vary from 6% to 22% [2,3]. In comparison, the breast cancer incidence in the general population is approximately 0.18% per year [4]. A lack of proven theories for mechanisms behind recurrence in early-stage breast cancer survivors hinders accurate prediction of recurrence as well as the development of effective prevention strategies [5–7]. The systemic or spectrum theory of breast cancer has explanatory value and suggests tumor cell dissemination taking place as early as the onset of the malignancy yet, without causing overt metastasis [42]. The theory is largely corroborated by the frequent findings of disseminated tumor cells (DTC) in the bone marrow of early-stage patients and years before clinical relevant recurrence [47,48], murine model experiments [40] and more recently findings of circulating tumor cells (CTC) in this patient group [6,17]. In this vein, early-stage breast cancers could be defined by loco-regional disease and a chronic systemic cancer component that is either dormant or indolent. This systemic cancer component apparently has the ability to persist independently of the primary tumor and may then be responsible for, or at least promotes, cancer recurrence. The sum of investigations about systemic breast cancer relate to micro-metastatic indolent cancer i.e. bone marrow DTC [47], CTC [17], immunology and systemic inflammation [8–11] and have helped to increase understanding about breast cancer systemic disease. Further advancement in the field may be achieved by deeper investigations into the cancer-specific circulating rare cell population [12,13] what we would like to refer to as cancer-associated chronic systemic abnormality (CASA). Similar to CTC, CASA denotes blood circulating entities that however, are investigated in comprehension and include systemic representatives of localizable tumor growth, such as tumor-derived CTC, and tumor-associated circulating endothelial cells (CEC) [13–16,20] but also of non-localizable cancer that may include cancer inflammation-related cell types [22], CTC of unknown origin [16,19] and circulating erythroblasts (CEB) [21]. In awareness of the systemic and localizable nature of circulating rare cells, the idea prevailed to assess cancer burden in its entirety based on comprehensive circulating rare cells read-out [18] which would be potentially useful in monitoring response to interventions (in particular post-surgery interventions), to survey the chronic systemic cancer disease during therapy and in follow up as well as to predict recurrence at the earliest time point possible. However, the field of circulating rare cell analysis is in need of intensified marker validation prior to the realization of before mentioned diagnostic potentials, as investigations into rare cells other than CTC and CEC as well as into their association with cancer or tumor growth are sparse [12,20–22]. The primary aim of the present study was to introduce a novel tool for the characterization of early-stage breast cancer based on circulating rare cells. The present findings led to the formation of a specific CASA rare cell panel and supported the notion of a general as well as specific rare cell abnormality in each patient sample. A study design of testing patients before and after surgery was chosen to evaluate the tumor-association of each rare cell marker. This study denotes a prove of concept as the basis to investigate correlation of CASA profiles with various breast cancer states for the benefit of improvements in staging and follow up care. We showcase and advocate the combination of fluorescence marker and cytomorphological analysis of rare cells and demonstrate its association with various breast cancer pathologies and subtypes. A liquid biopsy platform to support fine-stratification of rare cell morphology was employed and included enrichment by the negative selection principle and staining of viable cells in solution all for the purpose of recording the at best blood-native morphological status of rare cells.

2. Materials and Methods

2.1. Study Sample

The study was approved by the Institution's Research Ethics Committee (details at the end of the text). A**ll methods were performed in accordance with the relevant guidelines and regulations. Histopathology grading followed Bloom-Richardson system.** Informed consent was obtained from all study participants prior to blood draw. Blood samples were obtained from 20 healthy individuals, denoted as the healthy cohort, 9 non-cancer but ill individuals, denoted as non-cancer cohort, and 14 predominantly early-stage breast cancer patients, denoted as the cancer cohort. Characteristics of cancer patients are presented in Table 1. A total of 28 blood samples from 14 breast cancer patients

 doi:10.20944/preprints202402.0681.v1

were drawn 2 weeks before and 2 weeks after surgery. All breast cancer patients were therapy-naive at the time of surgery. The non-cancer control cohort comprised a heterogeneous group of 9 individuals with some clinical condition or vague illness without evidence of cancer, including one woman with "unwell feeling" and early night-time fatigue; a type II diabetes patient; a man with an unhealthy lifestyle showing slight rare cell abnormality; a 41-year old man recovering from an acute infection of unknown diagnosis; a 77-year old man with several comorbidities including type II diabetes, kidney failure requiring hemodialysis, and hypertension; a 73-year old woman with symptoms of osteoporosis, a 38-year old woman with chronic pain in the lower spine and showing rare cell abnormality; a 67-year old obese non-diabetic woman with gout-like symptoms, and a 38-year old man recovering from mild acute respiratory illness. The healthy control cohort comprised 10 volunteers for present study and 10 from a previously study [22]. Healthiness was defined as being free from known medical conditions, from any illnesses within the past 2 months, and not taking any medication.

Table 1. Clinical characteristics of 14 breast cancer patients.

Patient code	Age* (years)	Comorbidities/ Vaccination	Histological characteristic	Lymphatic/Node involvement	IHC markers	Type of surgery	Systemic immune-inflammation index (SII)**
BR001	51	No comorbidity/ Covid-19 vaccination 1 month before enrollment	Invasive lobular carcinoma, 26 mm, Grade 2, Free margin (Closest margin 1 mm)	Lymphovascular invasion	ER 95% PR 95% Her-2 0% Ki67 30%	Breast-conserving surgery (BCS) + Sentinel lymph node biopsy (SLNB)	High (3,279)
BR002	74	Diabetes type II, Peripheral vascular disease	Invasive ductal carcinoma, 36 mm, Grade 2, Free margin (Closest margin 12 mm)	Lymphovascular invasion Axillary node 3 of 11	ER < 1% PR < 1% Her-2 1+ Ki67 15%	Mastectomy + Axillary lymph node dissection (ALND)	High (581)
BR003	80	No comorbidity/Covid-19 Vaccination 1 month before enrollment	<u>Left:</u> Invasive ductal carcinoma, 45 mm, Free margin (Closest margin 3 mm)	<u>Left:</u> Lymphovascular invasion	<u>Left:</u> ER 99% PR 20% Her-2 0% Ki67 40%	<u>Left:</u> Mastectomy with SLNB	High (913)
			<u>Right:</u> Invasive ductal carcinoma, 24 mm, Free margin (Closest margin 8 mm)	<u>Right:</u> Lymphovascular invasion Axillary node 3 of 5	<u>Right:</u> ER 99% PR 20% Her-2 0 Ki67 40%	<u>Right:</u> Mastectomy with ALND	
BR005	34	No comorbidity	Invasive ductal carcinoma, 20 mm, Grade III, Free margin (Closest margin 2 mm)	Lymphovascular invasion	ER 0% PR 0% Her-2 2+ Ki67 90%	BCS + SLNB	High (1,488)
BR006	55	No comorbidity	Invasive ductal carcinoma, 28 mm, Grade III, Free margin (Closest margin 2 mm)	Lymphovascular invasion Axilliary node 4 of 48	ER 0% PR 0% Her-2 3+ Ki67 50%	Mastectomy + ALND	Low (469)
BR007	53	No comorbidity	Invasive ductal carcinoma, 25 mm, Grade III, Free margin (Closest margin 5 mm)	Axilliary node 11 of 27	Triple negative Ki67-60%	Mastectomy + ALND	Low (209)

 doi:10.20944/preprints202402.0681.v1

BR008	58	No comorbidity / Covid-19 vaccination 1 month before enrollment	Invasive ductal carcinoma, 30 mm, Grade III, Free margin (Closest margin 5 mm)	No involvement	Triple negative Ki67 50%	BCS + SLNB	High (515)
BR009	59	No comorbidity/ Covid-19 vaccination 3 months before enrollment	Invasive breast carcinoma, 2 mm, Grade II, Free margin (Closest margin 8 mm)	No involvement	ER 100% PR 80% Her-2 0 Ki67 10%	Mastectomy + SLNB	Low (246)
BR012	62	Peptic ulcer	Invasive papillary carcinoma, 19 mm, Grade III, Free margin (Closest margin 33 mm)	Axillary node 1 of 3	ER < 1% PR < 1% Her-2 3+ Ki67 30%	Mastectomy + SLNB	Low (456)
BR013	69	Diabetes, Peptic ulcer/ Influenza vaccination 1 month before enrollment	Invasive ductal carcinoma, 13 mm, Grade II, Free margin (Closest margin 5 mm)	No involvement	ER 95% PR 20% Her-2 0 Ki67 25%	Mastectomy + SLNB	Low (259)
BR014	61	No comorbidity	Invasive ductal carcinoma, 19 mm, Grade II, Free margin	No involvement	ER < 1% PR < 1% Her-2 3+ Ki67 30%	Mastectomy + SLNB	Low (224)
BR016	40	Peptic ulcer/ Covid-19 infection 3 months before enrollment	DCIS (left) and Invasive ductal carcinoma (right) Grade II, Free margin (Closest margin 7 mm)	No involvement	ER 80% PR < 1% Her-2 0 Ki67 10%	Mastectomy + ALND with reconstruction	Low (263)
BR017	66	No comorbidity	Invasive ductal carcinoma, 33 mm, Grade III, Free margin (Closest margin 6 mm)	No involvement	ER 100% PR < 0% Her-2 2+ Ki67 25%	Mastectomy + SLNB	Low (448)
BR018	51	No comorbidity	Invasive ductal carcinoma, 12 mm, Grade III, Free margin (Closest margin 1 mm)	No involvement	Triple negative Ki67 60%	BCS + SLNB	Low (454)

*Age ≥ 60 years can be generally categorized as post-menopause. IHC: immunohistochemical. **Systemic immune-inflammation index (SII) obtained before surgery is an independent prognosis factor for breast cancer patients. SII is calculated using the formula SII = P × N/L, where P, N, and L represent the absolute platelet count (10^9/L), neutrophil count (10^9L), and lymphocyte count (10^9L), respectively. A high SII (cut-off value at 514) is associated with poor disease-free and overall survival [23].

2.2. Cell Preparation and Spiking

Low passage colon cancer cell line HCT116 and breast cancer cell line MCF-7 were cultured in fresh flasks containing RPMI 1640 supplemented with 10% fetal calf serum and containing low glucose Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal bovine serum (FBS), 2 mM glutamine, 0.01 mg/ml insulin and 1% penicillin/streptomycin mix, respectively and were incubated at 37°C in an atmosphere of 5% CO_2. Cell suspensions following trypsinization were accepted for use in case of single cell suspensions and viability greater 90%. The antibody staining capacity by anti-EpCam_FITC (ebioscience) for positive identification and anti-CD45PE(ebioscience) for counter staining was tested prior to spiking. Spike-in counts were assessed depending on the targeted amount. Exactness of spiking concentration for the lowest count of 5 cells was achieved by micro-manipulation as described previously [24]. In brief, model tumor cells were stained using CFSE dye (Cell Division Tracker Kit, Biolegend) diluted to roughly 2 cells per 10uL in a well of a flat bottom

 doi:10.20944/preprints202402.0681.v1

96-well plate and aspirated under microscope vision at 20x magnification. Exactness of higher spiking concentrations was achieved by respective dilution series and by imitating spike-in sample preparation in at least 3 repeats, then used as counting controls. The model tumor cells were then spiked into 5mL to 7mL whole blood, followed by the normal rare cell enrichment and analysis procedure. Experiments were conducted at different days per each data point, thus incorporating variability in blood specimens as well as tumor cell conditions. Experiments were conducted at least in duplicates. For linearity and sensitivity experiments, the spiking levels were targeted to measure 5, 100, 500 and 1000 tumor cells.

2.3. Rare Cell Enrichment

For systemic abnormality analysis, the Rarmax® cell-based liquid biopsy platform (Premise Biosystems Co., Ltd., Version 2.2) was used as previously described following the principle of negative selection [22]. In brief, whole blood nucleated cells were pre-enriched by removing bulk red blood cells and desired rare cells were enriched by removing bulk white blood cells. The analysis was conducted by automated fluorescence microscopy. In detail, peripheral blood was taken by venous puncture collecting 10 mL in a green-top BD Vacutainer blood collection tube containing sodium heparin. The blood sample was kept at room temperature (RT) in the dark and processed at most 3 hours after phlebotomy. Standard chemical lysis buffer (154 mM NH_4Cl, 10 mM $NaHCO_3$, 2 mM EDTA) treatment was applied to remove red blood cells (RBC) from 7 mL whole blood. The cell suspension was incubated twice at RT for a maximum of 3-5 min following centrifugation at 300xg for another 5 minutes each. The final cell pellet was resuspended in 0.5 mL PBS, supplemented with 0.5% bovine serum albumin and allowed resting for 15minutes at room temperature. The cell numbers of nucleated cells subsequent to RBC lysis were determined by hemocytometer (Neubauer) and subjected to enrichment. In brief, peripheral blood rare cell isolation was carried out by automated CD45 positive cell depletion assay (Walderbach II) following the manufacturer's description (SanoLibio GmbH, Germany). Subsequent to enrichment, the sample was split into halves containing 20-30µL each and immediately stained for subsequent analysis by fluorescence microscopy adding anti-CD45PE (ebioscience), anti-EpCam_FITC (ebioscience) in one panel, and anti-CD44Superbright645 (invitrogen), anti-CD24 PE(ebioscience), anti-CD71FITC (ebioscience) and CD45PercPCy5.5 (ebioscience) in the second panel, for each using 0.5µL undiluted dye solution and incubating at RT in the dark for 25 minutes. Epithelial cells were identified exclusively using the EpCam antigen clone 1B7 as the most common and widely accepted antigen in the field for cell membrane staining of epithelial cells without the need of cell perforation and fixation. A washing step followed by diluting the cell suspension with 1 mL PBS, supplemented with 0.5% bovine serum albumin prior to pelleting by centrifugation at 200xg for 3 minutes and resuspending in 120µL using PBS. Nucleus staining followed using 0.5 µL Hoechst 33342 DNA staining (ThermoFischer).

2.4. Image Acquisition and Analysis

The enriched and stained cell suspensions (panel 1 and 2) were aliquoted into 2 flat bottom wells per panel of a 384 well ViewPlate-black plate (Perkin Elmer) and monolayered by centrifugation at 80xg for 2min which was followed by high resolution image acquisition at 40x magnification using the Operetta high content imaging system (PerkinElmer). A standard acquisition protocol was used for all samples allowing intra-patient and inter-patient sample comparability. Images were recorded in a bright field channel, and channels for UV, green, yellow, orange, and red fluorescence emission according to the fluorescent color choices at optimized and fixed exposure times. Columbus analysis software served as a screening and image analysis tool. Sample analysis followed 2 reader assessment with the first person conducting an image screening step according to a standardized protocol and the second person conducting rare cell identification. Image records per sample comprised 117 fields (color z-stack) per well. Rare cell marker positive cells were identified by a cell-like round formation, specific morphologies as detailed in the result section, positive Hoechst DNA staining, positive green (CD326 or CD71) and/or orange (CD44) fluorescence emission in the absence of the typical ring formation or membrane staining as a consequence of positive CD45PE or CD45PerCy5.5 staining.

 doi:10.20944/preprints202402.0681.v1

The hematopoietic status was excluded in case of CD45 signal at background noise level and PE-autofluorescence as identified by dim positive fluorescence signal inside the cellular event and the absence of membrane staining. Sample read-out followed the list of cell types as shown in results (Table 2) starting the analysis with circulating epithelial events (CEE). The high baseline concentration of CEE in healthy donors sufficed reading of 2 x 20 fields out of the 117 fields then extrapolating the concentration to 3.5mL per panel and finally normalizing to our standard of 5mL. All other cells were read through the entire collection of fields. For the assessment of positivity of epithelial events as lead markers, the maximum green fluorescence emission needed to exceed 2SD levels above averaged background noise (surrounding maximum pixels) strictly visible within boundaries of the bright field appearance. Consequently, dim fluorescence emission ranged in signal to background ratio between 1.05 and 1.5, low fluorescence between 1.5 and 1.8 and high greater 1.8. Endothelial-like cells were solely recognized by significance in morphology. Image analysis of cells followed previous descriptions [22]. In brief, each CIC was attributed with a value of the morphological index. A value greater 35 was found to be attributable to malignant disease formerly termed $CD44_{i35}$ hence, the classification of findings with the CD44/24 phenotype as normal (nCIC) and tumor-associated CIC (tCIC). Image analysis of CD71 followed descriptions a previously made [21].

2.5. Statistical Analysis

For linearity, sensitivity, and accuracy assessment of the Rarmax rare cell detection platform NCSS2023 Statistical Software Linear regression and correlation analysis tool was used. Diagnostic performance of the rare cell panel in diagnosing breast cancer was measured using sensitivity, specificity and predictive values, for abnormality cut-offs of each rare cell type, as defined by the limit of detection equation LOD = mean (limit of blank) + (1.68-2)xSD. The receiver operating characteristic (ROC) curve analysis was performed using the Stata v. 14 statistical software (Stata Corp, College Station, Texas, USA), and the area under the curve (AUC) was calculated along with the 95% Confidence Interval (CI). The Tables 4 and S1 contain raw data as source for all calculations of AUC and diagnostic performances as presented in Tables 5–7, respectively. Conclusions of difference in concentration reduction between the cancer-naive and surgical treated cohort were based on the one-tailed paired t-test and was calculated using Excel Microscoft. A p-value < 0.05 was considered statistically significant. The association between each rare cell type and surgical intervention was defined in terms of the response rate, which was defined as the reduction in cell concentration, in percent, after surgery as compared with before surgery. The response was set to 100% if the post-surgery concentration was lower than the cut-off level for that rare cell type, and set as no response if the concentration after surgery was higher than or equal to that before surgery.

3. Results

3.1. Reliability

The spike-in experiment outcome was the number of spiked HCT116 and MCF-7 cells in range of 5 till 950 plotted against the recovered number of tumor cells in the healthy donor samples (n=9). The ratio between the number of observed and the number of expected tumor cells produced a regression slope of 0.89 [95% confidence interval (CI), 0.85–0.94], an intercept of -5.3 (95% CI, -26.8-16.1) (accumulated across 9 data points), and a correlation coefficient (R^2) of 0.997. The coefficient of variation (CV) did not increase as the number of cells spiked decreased, and was measured as the standard deviation of recoveries across all spike levels with a value of 7.61%. The average recovery rate of spiked cells measured 84.6% across all spiking levels in range of 78% to 92.6%. The accurate spiking of very low numbers by micro-manipulation done in triplicates allowed the determination of the lower limit of detection. However, at the spiking level of 5 to 6 tumor cells, all three data points showed a consistent loss of 1 cell into the process implying a near-zero standard deviation for the low-spike. Non-spiked samples were negative resulting in a mean value of zero for the limit of blank. Therefore, the average CV between the 9 spiking points was used that measured 7.61% and would

 doi:10.20944/preprints202402.0681.v1

translate as a SD of 0.38 cells per 5mL at the 5 cell spike level. The LOD was then calculated as 2xSD resulting in a LOD value of 0.76 cells per 5mL or 1 cell per 6.6 mL. In practical terms, at a given restriction of initial WBC count of $8x10^7$ WBC per run due to technical limitations (corresponds to 6.5mL to 20mL whole blood) and the recovery of 85% on average, the reliable detection of at least 1 tumor cell may require a minimal concentration of 1 cell per 5.6mL of whole blood. At given performance, the detection of desired cells would be well supported for the retrieval of meaningful diagnostic information in afflicted individuals with an expected concentration range from 1 to several hundred cells per 10mL. A coarse measure of sustained high sensitivity and platform functionality over the course of test runs was represented by the detection of normoblasts that occur within a detectable concentration range (>1 cell per 5mL) under healthy physiology with 100% certainty [25]. Therefore, repeated failure to detect normoblasts would predict a platform error. All healthy individuals (n=10) as measured on occasions throughout the course of clinical sample testing were positive for normoblasts in range of 1 to 21 cells per 5mL. Platform related sampling error was measured and expected not to exceed 7%. In practice, 1 error was recorded during clinical study conduct until final drafting comprising 102 sample runs and was ascribed to an operator error while using the same platform version and operator within the same setting.

3.2. CASA Panel

Cellular systemic abnormality refers to circulating rare cell abnormalities detected at elevated concentration above baseline (healthy) range, or to de-novo emergence of rare cells, or to cyto-morphological changes in these cells. As cellular systemic abnormality can be related to both physiological and pathological processes [12,21], cancer-associated cellular abnormality (CASA; see Table 2) will require further characterization in terms of its relationship with clinical cancer status as well as its presence in healthy individuals. Certain patterns or panels of rare cells in CASA may, for example, be related to the presence of clinical cancer, or its absence after cancer treatment. Certain panel of rare cells in CASA initially detected may disappear and some may persist after curative surgery. For the purposes of this article, we refer to rare cells in the CASA panel detected in cancer patients whose cancer is present as "cancer-present CASA", in those with absent clinical cancer after curative surgery as "cancer-absent CASA", and in non-cancer subjects as "cancer-naïve CASA".

Table 2. Cancer-associated systemic abnormality (CASA) panel.

Cell type	Phenotype	Description	Diagnostic Implication	References
Circulating Epithelial Cell Events (CEE and CTC)	EpCam+/CD45-/Hoechst+	EpCam+ heterogeneous cell type composition including CEE and CTC	Multi-pathology association; Cancer association	[14]
Circulating Erythroblasts (CEB)	EpCam-/CD71+/CD45-/Hoechst+	CD71+ bone marrow-derived rare cells indicating bone marrow damage (BMD) represented by dyserythropoiesis	Multi-pathology association; Cancer association	[21]
Circulating Endothelial-like cells (CEC)	(CD31+)/CD45-/Hoechst+	Vascular- and bone marrow-derived rare cells representative of vascular dysfunction, repair and neo-angiogenesis	Multi-pathology association; Cancer association	[13,15]
Circulating Inflammatory cells (CIC)	EpCam-/CD71-CD44+/CD24-/CD45-/Hoechst+	Leukocyte-like cells with unknown origin associated with systemic inflammation	Multi-pathology association; Cancer association	[22]

As a consequence of the negative selection method of enrichment, a heterogeneous composition of rare cell types was observed that would come with phenotypic overlaps between different rare cell types. The resulting ambiguous cell identification based on fluorescence signal alone would hamper our goal to distinguish between tumor-derived and systemic rare cell types. Phenotypic overlap was, in particular, frequent at the dim fluorescence emission level. For example, positive EpCam signals should be specific for circulating epithelial cells, but these signals may arise from co-expression in endothelial cells [16], or auto-fluorescence in inflammatory or activated cells [26,27]. Therefore, the

 doi:10.20944/preprints202402.0681.v1

inadequacy of rare cell classification based on phenotyping alone has led us to look deeper into cellular morphology.

It was found useful to classify epithelial cells as identified by the EpCam+/CD45-/Hoechst+ phenotype into circulating epithelial events (CEE) and CTC. CEE and CTC were distinguishable with respect to size, texture under bright field and EpCam signal level. Morphological considerations also allowed us to identify 6 classes of CTC as well as a distinct CEE morphology, as shown in Figure 1 and Table 3. This classification may be helpful in distinguishing cancer-present CASA and cancer-absent CASA, which appear as "tumor" and "systemic", respectively in Table 3 in the row labeled "Cancer association".

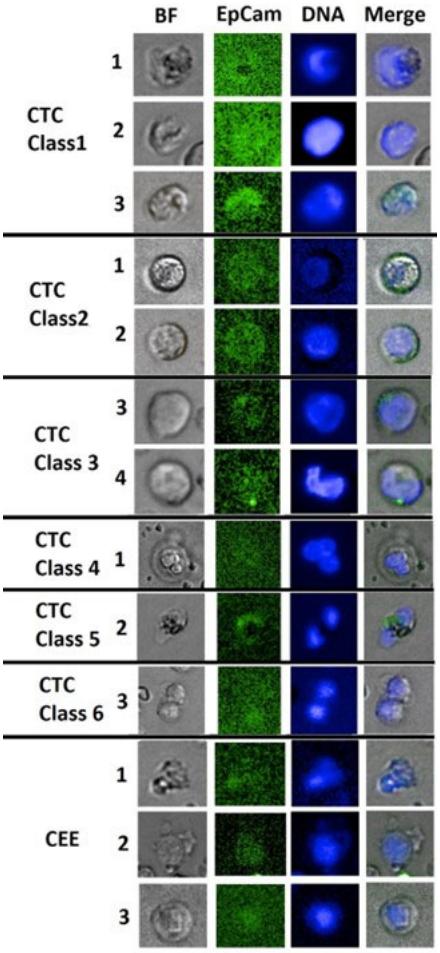

Figure 1. Gallery of circulating cellular epithelial events found in the cancer cohort. Cell images 1-3 denote class 1 cells with dim to low EpCam signal, N/C ratio greater 0.8, major diameters measuring 13.2 μm, 11.8 μm and 10.7 μm and less round to oval cell shapes. Cell images 4-5 denote class 2 cells of round cellular shape with low membrane staining, dim to low Hoechst intensity, N/C ratio from 0.5 to 0.8 and major diameters 12.9 μm, and 11.2 μm. Cell images 6-7 denote class 3 cells with low EpCam signal patches, N/C ratio of 0.5 to 0.8 and major diameters 12.1μm and 10.7 μm. Cell image 8 denotes a class 4 large round CTC with dim EpCam signal, low N/C ratio, nuclear abnormality and a major diameter of 22.3 μm. Cell image 9 denotes a class 5 CTC with bi-nucleation and a low EpCam patch with a major diameter of 15.2 μm. Cell image 10 denotes a class 6 CTC representing mitotic cell pairs or clusters with dim EpCam signal and a major diameter of 17.3 μm. Cell images 11-12 denote CEE of various shapes showing dim to low EpCam signal overlaps in parts with nuclear location and measuring in major diameter 9.8 μm, 16.6 μm, 11.8 μm. BF=bright-field.

 doi:10.20944/preprints202402.0681.v1

Table 3. Morphological identification criteria for circulating epithelial-like cells.

	Class 1	Class 2	Class 3	Class 4	Class 5	Class 6	Non-CTC and CTC in suspicion
Cell feature	Fully nucleated regular large	Round cell (Classic CTC)	Leukocyte-like	Extreme large cells	Multi-nucleation	Cell pairs/ Clusters	CEE
Cell Shape	oval or roundish	roundish to absolute round	Variable	Oval to round	Variable	Oval to round	All shapes, cell budding
BF appearance	Clear outer membrane rim with convex cell body, no inner or second membrane, heterogeneous texture	Strong membrane rim, heterogeneous texture	Strong inner or outer membrane rim heterogeneous or segmented texture	heterogeneous or segmented texture	Heterogeneous	Strong membrane rim heterogeneous texture	various membrane rim, homogeneous or patterned
Cell diameter (major axis)	10.5-18µm	9.5-14µm	8-18µm	>18µm	>8-18µm	>10 µm	6-18µm
EpCam staining intensity	Dim - Low	Low	Dim - High	Dim - Low	Dim - Low	Dim - Low	Dim
EpCam staining distribution	Partly strong, all cell or outer membrane	Clear membrane	Intracellular or partial/ exceeding Nucleus location	Intracellular or Partial	Intracellular or Partial	Intracellular or Partial	Intracellular, often overlaid with nucleus
Nucleus Staining	Low - High	Dim-Low	Dim-Low	Low-High	Dim-Low	Dim-High	Dim -Low
Nucleus texture	Homogeneous, no clear centre,	Homogeneous or rimed	Heterogeneous, polarized, nucleoli	Heterogeneous	Heterogeneous	Heterogeneous	homogeneous
Nucleus shape	cell shape aligned	oval to round	Variable, not aligned	Variable, not aligned	Variable, not aligned	oval to round	Oval, round
N/C-Ratio	>0.8	~0.5	~0.5 – 0.8	~0.5 – 0.8	~0.5 – 0.8	~0.5 – 0.8	~0.5 <1
Cancer Association	Tumor	Systemic	Systemic	Tumor	Systemic	Tumor	Systemic/Tumor

EpCam signal in CTCs was usually dim or low with intracellular or partial membranous distribution and most CTCs would not stand out from a background of WBCs. A careful and time-consuming inspection of these images would be required for their identification.

Class 1 CTC is characterized by a full nucleus with rather homogeneous distribution of chromatin, with a relatively large cellular size (>10.5 µm in major diameter), and EpCam signal within cellular boundaries (Figure 1). The morphology was described in literature and herein found distinct enough to classify as a subtype [28]. An epithelial event with a full nucleus but with weak and declining Hoechst signal from the centre towards the membrane would be considered an apoptotic cell and were classified as CEE. Class 1 CTC events <10.5 µm were classified as Class 3 CTC. The cut-off in diameter denoted a reader estimation and yielded sufficient specificity as will be shown in the following. Larger cells with an irregular convex shape and clear rims under bright-field microscopy are distinct from cellular artifacts, and were also classified as CEE (see Figure 1). The identification of a convex shape helps support the notion of a natural, full-cell nucleus; since an artifactual full nucleus may have been incurred by physical force due to repeated centrifugation prior to imaging analysis. The naturally occurring full-cell nucleus could be explained by a cellular process of nucleus expansion, which would increase cellular stability as a counter-measure to increased shear force in the circulation [28] denoting a morphological adaptation to a new and rough environment upon egress from the tumor. This implies that class 1 CTCs are a part of cancer-present CASA. The evidence for this is the high (> 90%) Class 1 response rate to surgical intervention in 9 out of 14 cancer subjects.

 doi:10.20944/preprints202402.0681.v1

The difference between pre- and post concentration levels was significant (p=0.005). The high response rate means that most of these cells are within normal range after surgery. Further evidence is that the class 1 CTC cut-off at 2 cells per 5 mL has good sensitivity and specificity in distinguishing cancer-present subjects from cancer-absent (surgically removed) and cancer-naïve subjects (64% sensitivity and 98% specificity) [AUC = 0.90; 95% CI: 0.79 to 0.99].

Class 2 CTC has full membrane staining, dim or low EpCam signal, and relatively small oval to round off-centre nucleus. These cells correspond to the classic CTC morphology with similarities to those reported by the Cell-Search system and thus, deemed a suitable for subtyping. Class 2 CTC were frequently found in cancer subjects both before and after surgery, but not in non-cancer and healthy control cohorts. Thus, Class 2 CTCs are both cancer-present and cancer-absent CASA, suggesting cancer-associated processes partly independent of the primary cancer.

Class 3 CTC is identified by a relatively strong, active nuclei and sometimes abnormal shapes (figure 1) along with EpCam patches overlapping with bright-field images, independent of the location of the nucleus, or having clear membrane staining. Therefore, Class 3 CTC denoted a class that comprehended all other morphological variations and might benefit from further subtyping in future investigations. Class 3 CTC can be found in cancer subjects both before and after surgery, but can also be found in cancer-naïve subjects with chronic illnesses. However, the maximum concentration of Class 3 CTC in cancer-naïve subjects was 7 cells per 5mL.

Classes 2 and 3 CTC showed variable changes in cell concentration after surgery, yet the reduction was still significant (p=0.035). These changes were low on average 43.8%±37.7 from 14 positive (without cutoff) out of 14 patients, ranging from 1 to 41 cells (median 7.5 cells per 5mL before surgery, and median 6 cells per 5mL after surgery), which seems to imply the usefulness of classes 2 and 3 CTCs as being diagnostic for cancer status regardless of the presence of cancer, i.e they are cancer-absent markers. The optimal cut-off at the cell concentration ≥ 4 cells per 5mL provides a sensitivity of 71% and a specificity of 93% [AUC=0.92; 95% CI: 0.85 to 0.99] when tested against healthy and non-cancer subjects.

Classes 4, 5, and 6 CTC include extremely large, multi-nucleated cells seen in pairs or clusters. These cells were extremely rare in the present study and were detected only before surgery. Although this finding seems to be a part of cancer-present CASA, the limited data precludes any definitive conclusion.

CEEs are represented by a broad spectrum of cellular and nuclear morphology. CEEs are often smaller in size in the healthy subject when compared to cancer cohort, with dim intracellular Epcam expression that overlapped with the location of the nucleus (Figure 1). In the cancer subject, CEEs include cells that could not be grouped into CTC classes as defined previously. Because of this heterogeneity, the cell types comprising CEEs are unknown. Assuming that the EpCam signal is genuine, or caused by auto-fluorescence, non-epithelial cell types with low or co-expressing EpCam such as endothelial-like cells are most likely to be the major cell type in CEEs [16,29].

CEEs were detected in all subjects. The range of cell concentrations in healthy subjects was between 41 to 268 cells per 5mL (average 106.1 cells per 5mL), in non-cancer subjects this was 60 to 657 cells per 5mL (average 295.4 cells per 5mL), and in cancer subjects before surgery, 93 to 474 (average 231.9 cells) cells per 5mL. To explain the extensive overlap of concentrations among different subjects, we assume that the majority of CEEs are the result of tissue micro-lesions or impaired cell removal independent of clinical illness.

The usefulness of CEE as a marker of systemic abnormality in general may be indicated by the higher cell concentration above the normal range. The cut-off for the determination of abnormal status regardless of disease severity can be estimated using the data from healthy subjects. Using the limit of detection (LOD) criterion, defined as the minimum 1.68×SD, in healthy subjects the LOD was 223 cells per 5mL, so the cut-off at 225 cells per 5mL was used in the present study. Applying this cut-off, CEE elevation was detected in 6 of 14 breast cancer subjects before and in 3 of 14 subjects after surgery, respectively at up to twice the cut-off value. These elevated CEEs are likely cancer-present CASA, despite the relatively weak decrease of 55.7%±42.8% in CEE concentration after surgery.

 doi:10.20944/preprints202402.0681.v1

The reason we believe our data may support the presence of cancer-derived cellular material within the CEE population was the detection of strong response to surgery in 4 of 6 cancer subjects with elevated CEE concentration before surgery, decreasing to within the normal range after surgery. A follow up, longitudinal study of an unhealthy, non-cancer individual seems to show stable CEE concentration with time (within one month), supporting the notion of CEE as a marker of chronic abnormality (data not shown). We therefore declare CEE as cancer-present CASA, but only in the setting of cancer subjects.

A second diagnostically most informative group of cells in the CASA panel (Table 2) are the CECs. These rare cells comprise a vast spectrum of cell subtypes with respect to function and origin. Subsets of CEC may be cancer-derived or detected in the presence of cancer, and may represent evidence of dysfunctional tumor microvasculature or neoangiogenesis, that is, representing ongoing tumor growth.

CEC may also be bone marrow-derived as part of a damage and repair process as well as originate from lymphnodes or any vascular lesion in the body. CECs are commonly present in many types of vascular disorders [33], with associated endothelial dysfunction [30], including certain benign or malignant tumors [16], in hypertensive disorders, pulmonary hypertension [31], and various heart diseases [32].

Because of the low specificity of endothelial markers such as CECs, the demonstration of a causal association between CEC and breast cancer may require substantial knowledge of patient co-morbidities and detailed cancer characteristics, and may possibly require knowledge of single cell genetic alterations typically found in association with cancers [15]. Our present approach is to describe a sensible classification of CEC morphologies whose association with cancer can be demonstrated by their presence or absence after surgical treatment.

CEC may be similar in size and morphology to leukocytes, requiring their identification by the CD31+/CD45-/Hoechst+ phenotype. However, in the present study we disregarded these CECs and focus on those which are morphologically distinct from the rest of the cells found in the enriched blood samples, and distinct from other CEC subtypes. Thus, we did not require additional phenotype identification. We simply classify these cells as being either normal endothelial-like cells, including progenitors, or abnormal CEC when they are morphologically distinct (Figure 2).

CEC abnormality is represented by multi-nucleated cells, aggregation of mature endothelial cells [33,45] or clustering of progenitor cells. Further abnormality includes cellular aberration, showing multiple irregular nuclei including chromosomal irregularity (Figure 2). With increasing abnormality, the presence of cancer is increasingly likely [15]. We have therefore classified CEC events as normal (nCEC), abnormal(aCEC) and aberrant (tCEC). CEC assumed to be associated with cancer is represented by aCEC and tCEC.

nCECs were frequently found in control subjects (2 of 10 in healthy and 6 of 9 in non-cancer subjects). Also, 8 of 14 cancer subjects before surgery were nCEC positive in the range 1 to 67 cells per 5mL, and in 9 of 14 subjects after surgery, showing a decrease of 63.8% ± 46.0% with de-novo appearance in 3 of 9 post-surgery positive subjects. It seems that nCEC may be related to surgery-induced vascular damage.

Aberrant CEC or tCEC is of special interest. All tCEC positive subjects before surgery (6 of 14) were negative for this cell type after surgery, but de novo emergence occurred in one subject after surgery. The average reduction was 85.7%±37.8%, a significant association with surgical intervention and would corroborate previous findings [51]. When including the de-novo positive subject, the specificity was 93% for cancer-presence. Because none of the healthy (n=20) or non-cancer subjects (n=9) was positive for this marker, and prior cytopathological knowledge seems to associate tCEC with severe vascular dysfunction, we interpret the status of the only de-novo positive subject as a true positive for cancer and set the abnormality cutoff to 1 cell per 5mL. Consequently, the sensitivity of this marker for breast cancer is 43%, with a PPV of 70% and a NPV of 83%.

Abnormal CECs were detected in 5 of 14 cancer subjects before surgery. After surgery, 4 subjects were positive, with 2 being de-novo. The average reduction was 50%±53.5% and seems to argue against aCEC as being related to cancer. Also, aCEC was observed in 3 non-cancer subjects. Therefore,

 doi:10.20944/preprints202402.0681.v1

the usefulness of aCEC was limited, although these cells may be used in conjunction with tCEC to help predict systemic cancer. The combination aCEC and tCEC in association with cancer status, points to an abnormality cut-off of 1 cell per 5mL.

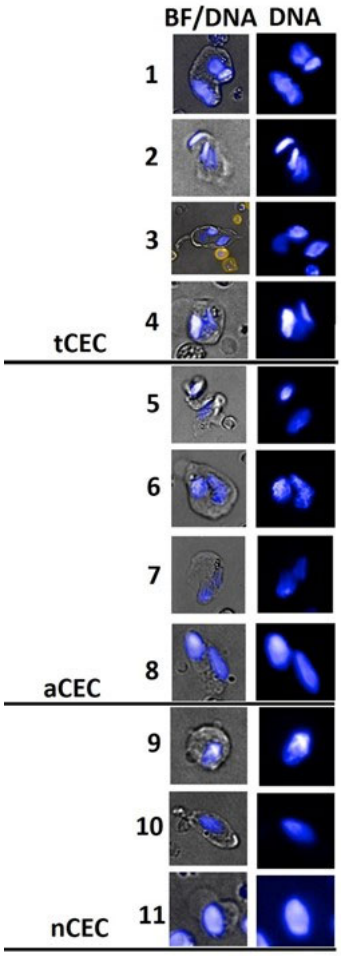

Figure 2. Gallery of circulating endothelial-like cells (CEC) derived from breast cancer patients. Cell images 1-4 demonstrate aberrant CEC (tCEC) with highly irregular nuclear morphology and polyploidy. Cell images 5-8 demonstrate abnormal CEC (aCEC) showing polyploidy but with relatively normal nuclear morphology or aggregation (image 5). Cell images 9-11 demonstrate normal CEC (nCEC) showing one nucleus per cell with oval-shape morphology. These latter include bone-marrow derived progenitor cells, with round to oval shapes.

CEB are classified as having normal or abnormal morphology according to our previous work [21]. These cells are highly sensitive in predicting bone marrow pathology referred to as BMD and may provide insight into cancer behavior and response to treatment. Although the detection of CEBs should always be considered abnormal, these cells may indicate cancer-presence when invasion of tumor cells into the bone marrow has occurred [21]. An elevation above the previous established median cell concentration of 7.5 or 8 cells per 5mL could serve as an indicator of abnormality for purposes of cancer diagnosis [24]. As can be expected in early-stage breast cancer without prior systemic treatment, only 4 of 14 cancer subjects showed mild BMD as indicated by the presence of abnormal CEB, denoted aCEB in Table 3, in the range of 1 to 4 cells per 5mL.

CIC have been shown in previous work to indicate cancer-presence, depending on morphological irregularity [22]. A cut-off greater 5 cells per 5mL was established for normal CIC (nCIC) based on the LOD determination. Concentrations above the cut off would indicate any non-specific abnormality. Indeed, because the average decrease after surgery was only 25.2% ±33.0%, it seems that nCICs are unrelated to cancer-presence.

 doi:10.20944/preprints202402.0681.v1

However, CICs with positive $CD44_{i35}$ status, denoted as aCIC, show a positive predictive value of 87% for cancer status. Systemic abnormality as defined by the combination of nCIC elevation beyond cutoff or the presence of aCIC cells alone was found in 10 of 14 cancer subjects. aCIC alone was found in 6 of 14 subjects. The average reduction of aCICs after surgery was 53.6%±51%, but was highly variable. Three subjects had CIC levels returning to normal, but 4 did not, which included 2 de-novo cases as well. These findings seem to support a systemic cancer component, as part of cancer-absent CASA, independent of primary cancer.

3.3. Diagnostic Performance

We examined the association between cell types in the CASA panel and breast cancer status regardless of surgical status, as well as their association with breast cancer presence, before surgery, and absence after surgery. In addition, we gained insight into possible surgery-related systemic side effects. In Supplementary Table S2, cancer-present CASA panel includes 3 cell categories: CEE, CTC classes 1, 4 and 6 denoted tCTC, and the aberrant CEC cells denoted tCEC. In the same table, cancer-absent CASA panel include 5 categories: CTC classes 2, 3 and 5 denoted sCTC, nCIC, aCIC, aCEB and aCEC, as defined previously. Panel of cells associated with side effects could be defined as those least associated with cancer but showing elevation after surgery. It was found that the combinations of normal and abnormal CEB, CEC and CIC (CEBall, CECall, CICall, respectively, in Supplementary Table S2 were strongly associated with surgical status, but not with cancer, and hence were candidate panels for side effects. The diagnostic performance of markers and combinations therefore specific to tumor-presence was of further interest. Cancer-present subjects were those with the presence of primary tumor before surgery, of whom there were 14. Cancer-absent subjects were those without the primary tumor after surgery, which included the same 14 subjects with breast cancer. Cancer naïve subjects included 20 healthy and 9 non-cancer subjects with some form of illness, as previously described. We used tCTC as the lead marker. We paired tCTC with one other marker, either tCEC, or CEE. We selected 2 cutoff values for tCTC, at 1 and 2 cells per 5mL, for detailed examination. The cutoff for tCEC was set to 1 cell per 5mL, as in most patients only 1 cell per 5 mL was detected. The cutoff for CEE was 225 cells per 5mL, as described previously. We chose the tCTC cutoff at 2 cells per 5mL, as the combined tCTC at this cutoff paired with either tCEC or CEE has a specificity as well as a positive predictive value of 100%, as shown in Table 4. The results may support a reasonable prediction of tumor-presence at given marker combinations.

 doi:10.20944/preprints202402.0681.v1

Table 4. Sensitivity, specificity and predictive values for cancer presence at various tCTC cut-offs.

	Cancer present N = 14	Cancer absent & cancer naïve N = 43	Predictive values
tCTC cut off at 1 cell/5 mL			
tCTC ≥ 1 and (tCEC ≥ 1 or CEE ≥ 225)* : **Positive test**	10	3	PPV: 76.9%
(tCTC = 0 or tCEC = 0) and (tCTC = 0 or CEE < 225)*: **Negative test**	4	40	NPV: 90.9%
Sensitivity & specificity	Sensitivity: 71.4%	Specificity: 93.0%	
tCTC cut off at 2 cells/5 mL			
tCTC ≥ 2 and (tCEC ≥ 1 or CEE ≥ 225)*: **Positive test**	7	0	PPV: 100%
(tCTC < 2 or tCEC = 0) and (tCTC < 2 or CEE < 225)*: **Negative test**	7	43	NPV: 86.0%
Sensitivity & specificity	Sensitivity: 50.0%	Specificity: 100%	

*All are in units of cells/5 mL; PPV: positive predictive value; NPV: negative predictive value.

Table 5. Sensitivity, specificity and predictive values for cancer status, regardless of presence of tumor.

	Cancer, before & after surgery N = 28	Cancer naïve N = 29	Predictive values
sCTC only; cut off at 3 or 4 cells/5 mL (see text)			
sCTC ≥ 3 or 4*: **Positive test**	21	1	PPV: 95.5%
sCTC < 3 or 4*: **Negative test**	7	28	NPV: 80.0%
Sensitivity & specificity	Sensitivity: 75.0%	Specificity: 96.6%	
sCTC cut off at 3 or 4 cells/5 mL combined with aCEC, aCIC or aCEB			
sCTC ≥ 3 or 4 and (aCEC ≥ 1 or aCIC ≥ 1 or aCEB ≥ 1)*: **Positive test**	16	0	PPV: 100%
(sCTC < 3 or 4 or aCEC = 0) and (sCTC < 3 or 4 or aCIC = 0) and (sCTC < 3 or 4 or aCEB = 0)*: **Negative test**	12	29	NPV: 70.1%
Sensitivity & specificity	Sensitivity: 57.1%	Specificity: 100%	

*All are in units of cells/5 mL; PPV: positive predictive value; NPV: negative predictive value.

In a similar manner, the diagnosis of cancer regardless of the presence or absence of the primary tumor can be done using sCTC as the lead marker, paired with any one of aCEC, aCIC or aCEB. These cellular markers may represent some systemic disease related to cancer. The cutoff for sCTC was set to 3 cells per 5 mL if at least 2 sCTC cell types were detected, and to 4 cells per 5mL if only class 3 CTC cells were detected. The cutoff for all other markers aCEC, aCEB and aCIC was set at 1 cell per 5mL, as none of these cells were detected in healthy subjects. sCTC as a stand-alone marker was able to achieve a specificity of 97%, with only 1 non-cancer subject misclassified due to a class 3 CTC concentration higher than the cutoff (see Table 5). However, the combined marker approach achieved 100% specificity and 100% PPV. Another interesting diagnostic approach would be to use the combination of sCTC and tCTC to diagnose primary cancer prior to surgery, as both are useful markers for cancer presence. Using the previous recommended cutoffs, the specificity and NPV for cancer remained high, but now the sensitivity and NPV for cancer are also very high as well, as can be seen in Table 6. Interestingly, the 2 false negative cancer subjects in Table 7 were actually positive for cancer according to the markers tCEC and aCIC.

 doi:10.20944/preprints202402.0681.v1

Table 6. Sensitivity, specificity and predictive values for the presence of cancer.

	Cancer before surgery N = 14	Cancer naïve N = 29	Predictive values
SCTC ≥ 3 or 4 (class3 only) and tCTC ≥ 2* or cancer-presence CASA positive or cancer-absence CASA positive: **Positive test**	12	1	PPV: 92.3%
sCTC < 3 or 4 and tCTC < 2*: **Negative test**	2	28	NPV: 93.3%
Sensitivity & specificity	Sensitivity: 85.7%	Specificity: 96.6%	

*All are in units of cells/5 mL; PPV: positive predictive value; NPV: negative predictive value.

We used all detectable rare cells in the categories denoted CEBall, CECall, CICall as the panel for the side effects of surgery, shown in Supplementary Table S2. The overall average relative increase of 2.8 times suggests that surgery negatively affects homeostasis of the rare cell population, as can be expected due to repair and recovery mechanisms still taking place after 2 weeks. The CEB panel is highly sensitive to any kind of interference to bone marrow homeostasis as has been reported previously. In the present study, 13 of 14 cancer subjects showed CEB elevation after surgery, with an average relative increase of 3.2 times (Supplementary Table S2). The changes in CEC levels, which should represent vascular damage, were highly variable, with only 3 subjects showing an increase, 5 remaining stable, and the rest did not have detectable CECs after surgery. Similarly, changes in CIC levels after surgery were highly variable, with 6 of 14 showing elevation and the rest showed decreased levels. We categorized "response to surgery" as complete, strong, partial and no response. A complete response refers the reduction of the CASA marker concentrations after surgery to below cutoff, or to normal, concentration levels. For the CTC panel, as defined in the previous section, complete response is seen in 5 of 14 subjects after surgery. Responses to surgery other then "complete" were rated according to the response average. Seven out of 14 subjects were rated as strong or partial responders with an average response rate in range of 25% to 70% (strong response: >65%). Lastly, two subjects were rated as non-responders with an average response rate lower than 25%.

Finally, all epithelial events including CTC and CEE were found unrelated to cancer stage. Furthermore, levels of both CIC types in subjects before surgery were found unrelated to the Systemic Immune-Inflammation Index (see Table 1) (nCIC: R^2=0.005 and aCIC: R^2=0.14). However, we observed a correlation between heightened CASA abnormality, either in CEB and/or CEC, and the more aggressive Her-2+ and/or triple negative breast cancer subtypes, with correct classification in 13 out of 14 cancer subjects ($p<0.005$). This finding might support the notion that breast cancer aggressiveness is associated with increased systemic involvement.

4. Discussion

Marker discovery and development is much needed in cell-based liquid biopsy, which could address issues such as tumor heterogeneity [17,34] as well as diagnostic interpretation [35,36]. Traditional approaches to advancement focused mainly on phenotyping and genotyping CTC for cellular identification, perhaps for the advantage of standardization, comparability and the access to the marker. However, standard protein or mRNA markers such as EpCam, cytokeratin, vimentin, CD31 or Her-2, were not sufficiently informative and may have contributed to the current limitations of liquid biopsy. Another, additional, approach for subtyping is in-depth identification and classification of rare cell markers based on cellular morphology. The admittedly subjectivity in reading and interpretation owned to possible overlaps in morphology criteria and fluorescence signal in a minority of cell events can be mitigated by intensive training in person or ultimately by computational pathology. It appears that only a handful of groups have focused on the analysis of CTC based on cytomorphological imaging [37–39,46]. We have also shown feasibility and improvement in the development of cancer-associated markers using an identification strategy that combines phenotyping with morphological analysis, as was applied to erythroblasts and circulating inflammatory cells [21,22]. The present study employs a similar identification approach for circulating epithelial cellular events and combines past and recent knowledge to build up a novel

 doi:10.20944/preprints202402.0681.v1

marker panel herein referred to as CASA for the purpose of breast cancer characterization based on its concurrent systemic abnormalities.

Traditionally, CTCs are assumed to be tumor-derived [14]. With further studies however, it became clear that CTC as well as circulating tumor DNA also exist in the absence of detectable tumors [17,53]. Despite, a significance in difference of CTC (class 2 and 3) between pre- and post surgery measurements (p=0.035), the present study supports the persistence of circulating epithelial cells after tumor resection. Findings were traditionally explained by excessive tumor cell shedding as a consequence of surgical resection and the week-long or months-long persistence of such cells [19]. The theory was considered less helpful in the understanding of what could be denoted as systemic cancer and ignores evidence of nesting and dormancy, CTC half-life, immune response and cell clearance activity. The theory of occult distant and dormant CTC shedding micro-metastases that are concurrent with the primary tumor and in a state of proliferation and apoptotic equilibrium is gaining more popularity [6,17,42]. Such CTC without known origin would be best denoted as systemic CTC and could hypothetically re-nest to become DTC for example in the bone marrow and sufficing the idea of a widespread systemic CTC-DTC cycle. However, given the relative high frequency of epithelial-like cells in the non-cancer control cohort, false positive identification as systemic CTC is likely and reduces the certainty of interpretation. Nevertheless, false positives may not represent mere artefacts but genuine rare cells with the most likely pathological underpinning of tissue inflammation. In consequence, interpretation of CTC by our platform to represent micro- and macro lesion needs caution rather supporting the idea to measure the extend of systemic cancer burden. Description and classification of CTC morphology are common in the literature and in congruence with our finding, yet have never been explicitly associated with cancer presence or absence [39]. Since the class 2 CTC were found highly specific to cancer yet independent of tumor presence, we may attribute this morphological subtype to distant unknown origin and to a representative of systemic cancer. It shall be noted at this point that the study and platform design are insufficient to draw such conclusions for the putative tumor-derived epithelial *Cytokeratin*+ CTC that may seem to fall into either category of class2 or 3. As evidence and in support of the commonly accepted theory of tumor-association only the class1 CTC was found to be strongly correlated with tumor-presence. The class 1 morphology also seems consistent with recent reports of CTC adapting to shear stress in the blood stream by nuclear expansion [28]. A classification of epithelial events (CEE) other than CTC was necessary since the majority of epithelial events were of uncertain nature. The source of this rare cell subset may be a tumor or occult inflammation [44], immature platelets, and additionally may be elevated upon dysfunctional or impaired cell removal by the organs involved. The composition of these tumor-derived cell entities in cancer is most likely dominated by endothelial cell types, as reported in two recent articles [16,29]. However, CEEs similar to CTC class3 were highly heterogeneous in morphology across all subjects in the present study, and a clear identification of epithelial or endothelial cell types is currently impossible. Nonetheless, the results suggested the presence of tumor-derived CEE subpopulations and infers usefulness as co-factor along with class 1 CTC and tCEC to monitor intervention response in particular in subclinical cancers.

The diagnostic performance of our liquid biopsy platform when leveraging the information about tumor-derived and systemic rare cell markers in the CASA system should be of practical interest. For the CTC panel, the cell concentration in subjects before surgery was low, ranging from 2 to 64 cells per 5mL with a median of 9.5 CTC per 5mL, but events were consistently found in all subjects. At this point further technological and methodological improvements are likely to play out in increased functional sensitivity however, we may point out that the biology of early-stage cancer simply has its limits in shedding relevant signal sufficient to be detectable in a volume of 3 mL to 10 mL whole blood. These findings are comparable with those of previous work [37–39]. According to Table 7, the sensitivity for detecting the presence of cancer is quite high at 86%, with 2 out of 14 false negative findings. This false negative identification is likely due to a limitation of the system of morphological classification, suggesting that CTC positive signal has been lost into the CEE class. Importantly, even in the false negative patients, cancer-absent rare cell abnormality was detected. Nevertheless, the performance of the CASA system will require further validation on both cancer

 doi:10.20944/preprints202402.0681.v1

patients and non-cancer controls. There are likely many confounding non-cancer disorders that might present with certain subset of CASA cells, and will lower the diagnostic performance of the CASA system. In particular, some class 3 CTCs are most likely associated with some systemic inflammatory disorders, as these cells are found in relative abundance in non-cancer subjects in the present study. Moreover, our preliminary investigation does not suggest an association between CASA cells and cancer stage, which might be due to the dominance of cancer-absent CASA cells in the circulation.

Consequential to the stratification of markers into tumor-derived or systemic, we identified CEE, tCTC and tCEC as markers to potentially predict cancer-presence based on the high response profile towards surgery. In contrast, sCTC, aCEB, aCEB, aCIC were found insignificantly correlated to tumor presence and thus, warrants investigation for use as surveillance and monitor marker panel in largely subclinical systemic chronic cancer disease during ACT or in follow up. The influence of the surgical intervention is corroborated by our response data of the cancer absent -CASA markers, showing a heterogeneous development in systemic disease over the stretch of the two measurements. A minority of patients showed a significant reduction in abnormality (response >50%) suggesting that this patient subgroup generated a low systemic chronic disease burden in the first place. BMD as indicated by CEB increase was detectable in almost all patients and suggests incomplete recovery still after 2 weeks. There was no association with changes in CIC, implying different causal mechanisms between both markers. Both CIC and CEC responses were highly variable after surgery and uncorrelated. When excluding one subject with de-novo endothelial dysfunction, surgery seemed to have even a beneficial effect in reducing vascular abnormality. CEC elevation or de-novo appearance was observed only in 3 out of 14 cancer subjects. This could be ascribed to recovery processes after surgery similar to observation of CEB elevation. Perhaps worth mentioning is the case of BR13 that showed de-novo elevation of aCEC and tCEC after surgery highlighting the possibility of greater damage in some patients. The adverse reaction is perhaps also indicated by severe inflammation having detected an increase in CIC concentration by a factor of 4.3. Since systemic inflammation and CTCs are known risk factors if not hallmarks for cancer recurrence [43], the patient might be at higher risk of relapse when compared to her peer group. In support, Retsky in 2013 has suggested that surgery-related long-term side effects may be a risk factor for early cancer relapse [5]. Such side effects would provoke sudden shedding of cancer cells from, or growth of, dormant micrometastases due to surgery-induced angiogenesis in these micrometastases, or even surgery-induced activity of single cancer cells [49]. The current belief that metastatic growth must be accompanied by the presence and activity of macrophage and endothelial cells may support this theory [6]. The finding in some studies that mastectomy has lower disease-free survival when compared to breast conserving surgery may be in line with before mentioned theory [41].

We have also looked at the so-called circulating tumor microemboli (CTM) events that are usually considered to be clustered CTCs and are reported to be relatively common, but we were not able to clearly identify such events. Although cellular aggregation was detected frequently in cancer patients, the presence of autofluorescence and leukocytes within such clumps made identification of CTCs impossible [52]. On the other hand, clusters with dim EpCam signal were found in rare cases, which were at given morphology of alleged endothelial origin and consequently, categorized as aCEC. Contrary to previous reports [50], our data may not support specificity to malignancy, as such endothelial clusters have been found in a case with benign adrenal gland tumors (data not shown). Another to be expected herein unlisted cell type is the fibroblast-like cell that was found to be extremely rare and also present in the non-cancer cohort. Therefore, the introduction of an additional subtype for circulating fibroblast-like cells as described in our previous work (Schreier et al. 2020) was meaningless for the purpose of investigating association and behavior in cancer.

Of further note is our decision to design a mixed gender and “open disorder” control cohort thus, focusing on marker specificity and not subject homogenicity given fact that the CASA markers are not gender and breast cancer specific (benign and malignant). The idea was to decrease bias and increase interpretation power. The study was conducted with relative low number of test subjects yet, sufficed statistical significance to reveal a statistical relevant differences between circulating epithelial cells when tested before and after in the same patient.

 doi:10.20944/preprints202402.0681.v1

Major limitations of the present study include a very small control sample size, the subjective nature of current morphological classification, and the purely speculative nature of much of the discussion. However, we believe that the findings in the present study are important for the advancement of liquid biopsy marker use and the discussion about breast cancer minimal residual disease. Further investigations in a larger and more varied sample of breast cancer patients are already underway, as well as follow up studies of patients in the present study. An AI approach to morphological analysis will be developed as well.

5. Conclusions

We evaluated the behavior of a novel rare cell marker panel referred to as cancer-associated systemic abnormality or CASA, based on a group of breast cancer patients before and after surgery and non-cancer controls. The key finding was a strong correlation of in particular fully-nucleated circulating epithelial cells and a weaker correlation of most other rare cells including round circulating epithetial-like cells with tumor-presence supporting the idea of breast cancer as a disease with early onset of systemic disorder. More so, CASA was not substantially effected by surgery and concluding that tumor-derived cell entities might denote a minority in the rare cell population of early-stage breast cancer patients. The findings support the notion of a more inflammation-associated systemic cancer burden that is largely independent of conventional cancer staging. Furthermore, the study showed that the prediction of tumor-presence and absence in early-stage breast cancer patients and survivors, respectively could be facilitated by CASA analysis. We provided novel insights into the nature of surgery-related side effects. Such post-surgery abnormalities remain largely unknown or overlooked. This proof-of-concept study does not support the use of this liquid biopsy platform to screen for cancer in healthy individuals.

Supplementary Materials: The following supporting information can be downloaded at: www.mdpi.com/xxx/s1, Table S1: Test results control cohort; Table S2: test results patients 1-14.

Author Contributions: SS conceived and designed the study, carried out imaging analysis and drafted the manuscript. PrB carried out experimentation and took part in image analysis. PB reviewed the study and took part in drafting the manuscript. SB facilitated the study conduct, reviewed the study and took part in drafting the manuscript. LA, PL and PC provided patient material and data, reviewed the study and took part in drafting the manuscript. WT advised the study, and drafted and reviewed the manuscript. All authors read and approved the final manuscript.

Funding: This study was funded in parts by Premise Biosystems Co. Ltd.

Institutional Review Board Statement: The study is part the project "Advancing cell-based liquid biopsy" and was approved for healthy and non-cancer volunteer recruitment by the Mahidol University Central IRB, Mahidol University (protocol number 2019/197.3007) as well as by Research Ethics Committee of Ramathibodi Hospital (protocol ID 02-65-01) for breast cancer patients. Breast cancer volunteers gave informed consent at the Out-Patients Department of the Department of Surgery, Ramathibodi Hospital.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Conflicts of Interest: The first author and Suparerk Borwornpinyo are shareholders of a company involved in the development and manufacturing of cell separation technology and biomarkers as was employed in this work. **The remaining authors do not have any competing interests to declare.**

References

1. Burstein HJ, Curigliano G, Loibl S et al. Estimating the benefits of therapy for early-stage breast cancer: the St. Gallen International Consensus Guidelines for the primary therapy of early breast cancer 2019. Annals of Oncology 2019; 30(10): 1541-57.
2. Pedersen RN, Esen BÖ, Mellemkjær L, Christiansen P, Ejlertsen B, Lash TL, Nørgaard M, Cronin-Fenton D. The incidence of breast cancer recurrence 10-32 years after primary diagnosis. JNCI: Journal of the National Cancer Institute 2022; 114: 391-9.
3. Elder EE, Kennedy CW, Gluch L, Carmalt HL, Janu NC, Joseph MG, Donellan MJ, Molland JG, Gillett DJ. Patterns of breast cancer relapse. European Journal of Surgical Oncology 2006; 32: 922-7.

4. Ellington TD, Miller JW, Henley SJ, Wilson RJ, Wu M, Richardson LC. Trends in breast cancer incidence, by race, ethnicity, and age among women aged≥ 20 years—United States, 1999–2018. Morbidity and Mortality Weekly Report 2022; 71: 43.
5. Retsky M, Demicheli R, Forget P, De Kock M, Gukas I, A Rogers R, Baum M, Sukhatme V, S Vaidya J. Reduction of breast cancer relapses with perioperative non-steroidal anti-inflammatory drugs: new findings and a review. Current medicinal chemistry 2013; 20: 4163-76.
6. Elkholi IE, Lalonde A, Park M, Côté JF. Breast cancer metastatic dormancy and relapse: An enigma of microenvironment (s). Cancer Research 2022; 16.
7. Liu C, Chen Z, Chen Z, Zhang T, Lu Y. Multiple tumor types may originate from bone marrow-derived cells. Neoplasia 2006; 8: 716-IN3.
8. Coussens LM, Werb Z. Inflammation and cancer. Nature 2002; 420(6917): 860-7.
9. Cole SW. Chronic inflammation and breast cancer recurrence. Journal of clinical oncology: official journal of the American Society of Clinical Oncology 2009; 27: 3418.
10. Pierce BL, Ballard-Barbash R, Bernstein L et al. Elevated biomarkers of inflammation are associated with reduced survival among breast cancer patients. Journal of Clinical Oncology 2009; 27: 3437.
11. Dunn GP, Bruce AT, Ikeda H, Old LJ, Schreiber RD. Cancer immunoediting: from immunosurveillance to tumor escape. Nature immunology 2002; 3: 991-8.
12. Schreier S, Triampo W. The blood circulating rare cell population. What is it and what is it good for? Cells 2020; 9: 790.
13. Hida K, Hida Y, Shindoh M. Understanding tumor endothelial cell abnormalities to develop ideal anti-angiogenic therapies. Cancer science 2008; 99: 459-66.
14. Allard WJ, Matera J, Miller MC et al. Tumor cells circulate in the peripheral blood of all major carcinomas but not in healthy subjects or patients with nonmalignant diseases. Clinical cancer research 2004; 10: 6897-904.
15. Lin AY, Wang DD, Li L, Lin PP. Identification and comprehensive co-detection of necrotic and viable aneuploid cancer cells in peripheral blood. Cancers 2021; 13: 5108.
16. Hu B, Gong Y, Wang Y, Xie J, Cheng J, Huang Q. Comprehensive atlas of circulating rare cells detected by se-ifish and image scanning platform in patients with various diseases. Frontiers in Oncology 2022; 12.
17. Pantel K, Alix-Panabières C. Liquid biopsy and minimal residual disease—latest advances and implications for cure. Nature Reviews Clinical Oncology 2019; 16: 409-24.
18. Schreier S, Triampo W. Systemic cytology. A novel diagnostic approach for assessment of early systemic disease. Medical Hypotheses 2021; 156: 110682.
19. Pachmann K. Current and potential use of MAINTRAC method for cancer diagnosis and prediction of metastasis. Expert Review of Molecular Diagnostics 2015; 15: 597-605.
20. Bhakdi SC, Suriyaphol P, Thaicharoen P et al. Accuracy of tumour-associated circulating endothelial cells as a screening biomarker for clinically significant prostate cancer. Cancers. 2019; 11: 1064.
21. Schreier S, Budchart P, Borwornpinyo S, Arpornwirat W, Triampo W. Circulating erythroblast abnormality associated with systemic pathologies may indicate bone marrow damage. Journal of Circulating Biomarkers 2021;10: 14.
22. Schreier S, Budchart P, Borwornpinyo S et al. New inflammatory indicators for cell-based liquid biopsy: association of the circulating CD44+/CD24− non-hematopoietic rare cell phenotype with breast cancer residual disease. Journal of Cancer Research and Clinical Oncology 2022: 1-2.
23. Li W, Ma G, Deng Y, Chen W, Liu Z, Chen F, Wu Q. Systemic immune-inflammation index is a prognostic factor for breast cancer patients after curative resection. Frontiers in Oncology 2021: 4516.
24. Schreier S, Sawaisorn P, Udomsangpetch R, Triampo W. Advances in rare cell isolation: an optimization and evaluation study. Journal of Translational Medicine 2017;15: 1-6.
25. Schreier S, Borwornpinyo S, Udomsangpetch R, Triampo W. An update of circulating rare cell types in healthy adult peripheral blood: findings of immature erythroid precursors. Annals of Translational Medicine 2018; 6(20).
26. Wizenty J, Ashraf MI, Rohwer N et al. Autofluorescence: A potential pitfall in immunofluorescence-based inflammation grading. Journal of immunological methods 2018; 456: 28-37.
27. Heintzelman DL, Lotan R, Richards-Kortum RR. Characterization of the autofluorescence of polymorphonuclear leukocytes, mononuclear leukocytes and cervical epithelial cancer cells for improved

 doi:10.20944/preprints202402.0681.v1

spectroscopic discrimination of inflammation from dysplasia. Photochemistry and photobiology 2000; 71: 327-32.
28. Xu Z, Li K, Xin Y, Tang K, Yang M, Wang G, Tan Y. Fluid shear stress regulates the survival of circulating tumor cells via nuclear expansion. Journal of Cell Science 2022; 135: jcs259586.
29. El-Heliebi A, Kroneis T, Zöhrer E, Haybaeck J, Fischereder K, Kampel-Kettner K, Zigeuner R, Pock H, Riedl R, Stauber R, Geigl JB. Are morphological criteria sufficient for the identification of circulating tumor cells in renal cancer?. Journal of translational medicine 2013; 11: 1-7.
30. Erdbruegger U, Haubitz M, Woywodt A. Circulating endothelial cells: a novel marker of endothelial damage. Clinica chimica acta 2006; 373: 17-26.
31. Bull TM, Golpon H, Hebbel RP, Solovey A, Cool CD, Tuder RM, Geraci MW, Voelkel NF. Circulating endothelial cells in pulmonary hypertension. Thrombosis and haemostasis 2003; 90: 698-703.
32. Boos CJ, Lip GY, Blann AD. Circulating endothelial cells in cardiovascular disease. Journal of the American College of Cardiology 2006; 48: 1538-47.
33. Tokunaga O, Satoh T, Yamasaki F, Wu L. Multinucleated variant endothelial cells (MVECs) in human aorta: chromosomal aneuploidy and elevated uptake of LDL. InSeminars in thrombosis and hemostasis 1998 (Vol. 24, No. 03, pp. 279-284). Copyright© 1998 by Thieme Medical Publishers, Inc..
34. Kolenčík D, Narayan S, Thiele JA et al. Circulating tumor cell kinetics and morphology from the liquid biopsy predict disease progression in patients with metastatic colorectal cancer following resection. Cancers 2022; 14: 642.
35. Rossi G, Ignatiadis M. Promises and pitfalls of using liquid biopsy for precision medicine. Cancer research 2019; 79: 2798-804.
36. Goeminne JC, Guillaume T, Symann M. Pitfalls in the detection of disseminated non-hematological tumor cells. Annals of oncology 2000; 11: 785-92.
37. Marrinucci D, Bethel K, Bruce RH et al. Case study of the morphologic variation of circulating tumor cells. Human pathology. 2007; 38: 514-9.
38. Marrinucci D, Bethel K, Luttgen M, Bruce RH, Nieva J, Kuhn P. Circulating tumor cells from well-differentiated lung adenocarcinoma retain cytomorphologic features of primary tumor type. Archives of pathology & laboratory medicine 2009; 133: 1468-71.
39. Marrinucci D, Bethel K, Lazar D, Fisher J, Huynh E, Clark P, Bruce R, Nieva J, Kuhn P. Cytomorphology of circulating colorectal tumor cells: a small case series. Journal of oncology 2010; 2010.
40. Hüsemann Y, Geigl JB, Schubert F et al. Systemic spread is an early step in breast cancer. Cancer cell 2008; 13: 58-68.
41. Christiansen P, Mele M, Bodilsen A, Rocco N, Zachariae R. Breast-conserving surgery or mastectomy?: Impact on survival. Annals of Surgery Open 2022; 3: e205.
42. Hellman S, Lecture KM. Natural history of small breast cancers. Journal of clinical oncology. 1994; 12: 2229-34.
43. Van Dalum G, Van der Stam GJ, Tibbe AG et al. Circulating tumor cells before and during follow-up after breast cancer surgery. International journal of oncology 2015; 46: 407-13.
44. Trzpis M, McLaughlin PM, de Leij LM, Harmsen MC. Epithelial cell adhesion molecule: more than a carcinoma marker and adhesion molecule. The American journal of pathology 2007; 171: 386-95.
45. Bethel K, Luttgen MS, Damani S, Kolatkar A, Lamy R, Sabouri-Ghomi M, Topol S, Topol EJ, Kuhn P. Fluid phase biopsy for detection and characterization of circulating endothelial cells in myocardial infarction. Physical biology 2014; 11: 016002.
46. Park S, Ang RR, Duffy SP, Bazov J, Chi KN, Black PC, Ma H. Morphological differences between circulating tumor cells from prostate cancer patients and cultured prostate cancer cells. PloS one 2014; 9: e85264.
47. Slade MJ, Coombes RC. The clinical significance of disseminated tumor cells in breast cancer. Nature clinical practice Oncology 2007; 4: 30-41.
48. Hartkopf AD, Taran FA, Wallwiener M et al. Prognostic relevance of disseminated tumour cells from the bone marrow of early stage breast cancer patients–results from a large single-centre analysis. European journal of cancer 2014; 50: 2550-9.
49. Neeman E, Ben-Eliyahu S. Surgery and stress promote cancer metastasis: new outlooks on perioperative mediating mechanisms and immune involvement. Brain, behavior, and immunity 2013; 30: S32-40.
50. Cima, Igor, et al. "Tumor-derived circulating endothelial cell clusters in colorectal cancer." *Science translational medicine* 8.345 (2016): 345ra89-345ra89.

 doi:10.20944/preprints202402.0681.v1

51. Mehran, Reza, et al. "Tumor endothelial markers define novel subsets of cancer-specific circulating endothelial cells associated with antitumor efficacy." *Cancer research* 74.10 (2014): 2731-2741.
52. Jiang, Xiaocheng, et al. "Microfluidic isolation of platelet-covered circulating tumor cells." *Lab on a Chip* 17.20 (2017): 3498-3503.
53. Rack, Brigitte, et al. "Circulating tumor cells predict survival in early average-to-high risk breast cancer patients." *Journal of the National Cancer Institute* 106.5 (2014): dju066.