# Fine analysis of atypical circulating tumor cells in metastatic breast cancer patients reveals subsets with different prognostic values

Alexia LOPRESTI*[1], Laurys BOUDIN*[1], Pascal FINETTI[1], Séverine GARNIER[1], Anaïs AULAS[1], Maria-Lucia LIBERATOSCIOLI[1], Arnaud GUILLE[1], Quentin DACOSTA[1], Olivier CABAUD[1], Jihane PAKRADOUNI[2], Anthony GONCALVES[3], Daniel BIRNBAUM[1], François BERTUCCI*[1,3], Emilie MAMESSIER*[1], Claire ACQUAVIVA*[1]

* These authors equally contribute to the work.

[1] CRCM, Predictive Oncology Team, INSERM, Institut Paoli-Calmettes, Aix Marseille Univ, CNRS, Marseille, France

[2] Department of Clinical Research and Innovation, Sponsor Unit, Institut Paoli-Calmettes, Marseille, France

[3] Department of Medical Oncology, Institut Paoli-Calmettes, Marseille, France

Correspondence to : Mamessier Emilie, emilie.mamessier@inserm.fr, +33 486 977 261, 13273 Marseille

**Running title: Prognostic value of atypical circulating tumor cells**

**Keywords:** Circulating tumor cells, breast cancer, metastases, stemness, drug resistance, EMT, biomarker, cluster, CTC, *CTM*, epithelioid cells, PFS and OS.

**Conflict of Interest:** AG received non-financial support from Astra Zeneca, Novartis, Pfizer, Roche.

**Funding support:**

This work has been supported by Inserm, Institut Paoli-Calmettes (SIRIC INCa-DGOS-Inserm 6038), and grants from the Ligue Nationale Contre Le Cancer (EL2016.LNCC/DaB, EL2019.LNCC/FB), Association Ruban Rose, and Foundation Groupe EDF. AL was supported by the fellowship DOC4 from the Fondation ARC (n°DOC20180507420). AA was supported by a postdoctoral fellowship from the Fondation ARC (n°PDF20180507565) and the Fondation de France (n°00107936).

**Acknowledgements:**

We would like to thank all the donors who have willingly contributed to the study (PERMED-01 Cohort patients, donors without malignant diseases); the Direction de la Recherche Clinique, for the management of the PERMED-01 cohort; D. Isnardon and M. Rodrigues from the Microscopy and Scientific Imaging platform.

**Abbreviations:**

mBC: metastatic breast cancer

CTCs: Circulating Tumor Cells

iCTCs: isolated Circulating Tumor Cells

gCTC: giant Circulating Tumor Cells

EM status: Epithelial to Mesenchymal status

EMT: Epithelial to Mesenchymal Transition

PFS : Progression free survival

OS : Overall survival

CTM : Circulating tumor microemboli

TN : Triple Negative

TAM : Tumor Associated Macrophage

CAMLs: Cancer-Associated Macrophage-Like cells,

**Abstract**

**Purpose:** Circulating tumor cells (CTCs) have a tremendous potential for diagnosis and treatment of breast cancer patients. Here, we performed a unique analysis of all atypical circulating cells isolated with a filtration-based technology from metastatic breast cancer (mBC) patients.

**Patients and methods:** The PERMED-01 study enrolled patients with mBC, refractory to systemic therapy, and with an accessible lesion to biopsy. We analyzed atypical circulating cells isolated from patients' blood at the time of inclusion using Screencell® Cyto device. For 23 out of 91 analyzed patients, this was completed by advanced immunofluorescence staining of atypical circulating cells. Subsets cut-offs were established using a two-component Gaussian finite Mixture Model, and evaluated for correlation with clinico-pathological data, including progression-free survival (PFS) and overall survival (OS).

**Results:** Three subsets of atypical circulating cells, absent from controls (n=7), were observed in cancer patients (n=91): isolated (*iCTCs),* Clusters (*CTM*), and Giant CTCs (*gCTCs*). CTCs' median number was 8.33 per mL. Co-expression of stem and drug resistance markers was associated with intermediate epithelial to mesenchymal transition phenotype in *CTM* and *gCTCs*, but not in *iCTCs*. Presence of *gCTC* was associated with shorter PFS and OS. Concerning PFS, assigning an immunofluorescence-based Epithelial to Mesenchymal status improved their prognostic value.

**Conclusion:** This study brings to light the diversity of CTCs in mBC patients and their specific molecular profiles regarding epithelial to mesenchymal transition, stemness and drug resistance status. It also highlights the involvement of an atypical circulating cell subset, the *gCTCs*, as a prognostic factor for PFS and OS.

## INTRODUCTION

Metastases are responsible for more than 90% of cancer-associated mortality (1). Therefore, anticipating and understanding their evolution using simple and reliable biomarkers is paramount. Metastases arise from malignant cells issued from a tumor (primary or metastasis) that have acquired an invasive phenotype and, most often, resistance to the first-line treatment. A few tumor cells can enter the blood circulation and, in this circumstance, are specifically called "Circulating Tumor Cells" (CTCs). Some of them, fit enough to survive to the drastic conditions encountered in the blood flow (anoikis, sheer stress, immune defense…), will eventually spread to other organs and then initiate, after a variable period of dormancy, novel malignant foyers (2). CTCs represent a unique access to biological characteristics from distant primary and metastatic lesions endowed with shedding properties (*i.e.* releasing tumor cells). In addition, CTCs' analysis is potentially a powerful approach to study the metastatic process and to determine the tumor response on a real-time basis as therapy proceeds. Finally, because CTCs can be obtained after a simple blood draw of less than 10 milliliters, their sampling is minimally invasive and can be repeated on demand (3).

CTCs enumeration at diagnosis of metastasis is recognized as an independent predictor of overall and progression-free survival (PFS) in metastatic solid tumors such as breast, colon and prostate cancers (4-8). CTCs counts evolution during chemotherapy in metastatic breast cancer patients is clinically predictive for disease progression and survival on a real-time basis as therapy proceeds (8,9). Nonetheless, CTCs prognostic value has a poor added clinical utility in metastatic patients and remains unproven in non-metastatic patients (4,10,11). The current hypothesis is that a mere numerical count of CTCs does not reveal their true nature, especially their role during the metastatic process (12). Indeed, not much is known about critical biological aspects, such as CTC heterogeneity (cytological, phenotypical…), and their relationship with prognosis. However, unbiased and precise cellular characterization is difficult to establish on

rare cells. Major efforts have thus been put into the development of new technologies that are both more sensitive and less biased than the CellSearch™ technology (only FDA-approved test for CTCs identification and enumeration so far) (Janssen Diagnostics, LLC, Raritan, NJ). The Screencell® Cyto system (ScreenCell™, France) is one of the promising technologies allowing to isolate all types of atypical circulating cells, provided that they are bigger and less deformable than leukocytes (13,14). Cytological criteria allow the identification of potential CTCs after the staining of immobilized cells on the filter. Multiple trials studying CTCs from several solid malignancies have been launched using this technology or equivalent (15-18), and the analysis of cells isolated with this device now permits “high sensitivity” detection (19).

Using filtration-based technologies, fragmented data have reported multiple rare and atypical circulating cells (14,20). For example, isolated circulating tumor cells or *iCTCs* are single individualized cells that appear epithelial-like by microscopy (*i.e.*, epithelioid) and display malignant cytological criteria. Despite the fact that additional studies might be required to confirm the epithelial and/or tumor origin of these cells, the naming of CTCs is well accepted, mainly because they have mostly been detected in patients with benign or malign conditions, but not in healthy subjects (13). Clusters of CTCs are also long known under the name of circulating tumor micro-emboli (*CTM*). They share most of the specific malignant cytological criteria used to identify single CTCs and are found in clusters of at least three cells in the blood flow (21). *CTM* are the most likely harbingers of metastases. They have long been reported as an aggravating prognostic factor when identified in the peri-tumoral or intra-tumoral vessels. Their presence in the blood, at distance from the tumor site, was shown to increase by 23- to 50-fold the risk for metastases compared to an equivalent number of *iCTCs* (22). A prevailing view is that many single *iCTCs* die in the circulation, whereas cells from *CTM* are protected from apoptosis, including immune attack or drug mediated-death (23). Giant Circulating Tumor Cells (*gCTCs*), also called Cancer Associated Macrophage-Like cells (CAMLs), are easily distinguished by their

large size. They were first described by Adams *et al*.  but this population remains puzzling and their denomination highly ambiguous (24). Some of these cells have indeed clear epithelioid cytological and phenotypic features, that might categorize these cells as giant circulating epithelioid cells rather than “macrophages”, while others seems to be poly-nucleated. This population has only been observed in cancer patients but not in healthy controls.

In the current study, we provide an exhaustive and simultaneous view of these rare and atypical circulating cell subsets in metastatic breast cancer patients and describe some of their associated biological features. For this, we used a filtration-based technology allowing their fine characterization. We next looked for correlations between each cell subset proportion and prognosis as well as matched metastasis’ clinical characteristics.

## MATERIALS AND METHODS

### PERMED-01 sub-study

The PERMED-01 study was a prospective unicentric clinical trial, promoted by and conducted at the Paoli-Calmettes Institute (Marseille, France), and registered as identifier NCT02342158 at the ClinicalTrials.gov platform. This study was proposed to patients with metastatic cancers, including breast cancers, refractory to at least one line of systemic therapy and with an accessible lesion to biopsy. Its primary objective was to evaluate the number of patients for whom identification of actionable molecular genetic alterations (AGAs) in tumor samples could lead to the delivery of a “targeted therapy”. Among secondary objectives, was the analysis of CTCs subsets isolated from the blood of patients at the time of inclusion. The trial was approved by the French National Agency for Medicine and Health Products Safety, a national ethics committee (CPP Sud-Méditerranée), and our Institutional Review Board (IRB). It was conducted in accordance to the Good Clinical Practice guidelines of the International Conference on Harmonization. All patients gave their informed consent for inclusion, biopsy and blood sampling, and molecular analysis. A total of 91 adult female patients with metastatic breast cancer (mBC) were enrolled after disease progression in PERMED-01 between January 2015 and December 2016 and were sampled. In addition, seven healthy adults lacking any known pathology were recruited as controls. For each subject, a blood sample of 5 mL was collected in a Vacutainer® tube containing EDTA K2 as an anticoagulant. The first milliliters of blood were discarded to avoid endothelial cells contamination during the puncture. All samples were shipped to the laboratory, treated and analyzed within 4 hours.

### Atypical circulating cells enrichment using ScreenCell® CYTO column

Blood samples were treated within 4 hours. To isolate rare circulating cells from whole blood, we used the Screencell® Cyto device. The samples were processed as described elsewhere (Kulemann, Pitman et al. 2015). In brief, 3 mL of peripheral blood were added to a red blood cell lysis/fixative buffer (ScreenCell® FC buffer) and passed through a filter (filters are dotted with 8-µm pores) with a low-pressure vacuum as per the manufacturer's instructions. Filters were dried at room temperature for 5 minutes before being used for further analysis (cytological and/or phenotypical staining).

**Atypical circulating cells staining with May Grünwald Giemsa**

Once dried, the filters were stained successively using three baths: a first bath contained pure May-Grünwald solution (1 minute 30) (Merck Millipore), the second one contained diluted May-Grünwald solution, volume to volume with distilled water (pH 7; 1 minute 30), and finally the third bath contained Giemsa solution, diluted 1:10 in distilled water (6 minutes) (Merck Millipore). After three baths in distilled water, filters were dried at room temperature for 5 minutes. Filters were then read using a Leica™ microsystem light microscope and the NIS NIS-Elements Viewer software (Nikon®).

**Atypical circulating cells subsets cytological criteria**

Rare circulating cells were analyzed for established characteristics that are similar to the malignancy criteria traditionally used for cytological examination. *iCTCs* display a nucleus size equal to or greater than 20 µm, high nuclear-to-cytoplasmic (N/C) ratio (≥0.75), a dense hyperchromatic nucleus -not totally opaque- and an irregular nuclear membrane. The nuclear size and the N/C ratio are the most important criteria. When the four criteria are met, malignancy

is almost certain. If only two (the two most important) or three out of four criteria are met, malignancy is considered as “uncertain” and additional analyses are required (see below: “Uncertain specimens”) (25). The malignancy criteria of *CTM* were mainly based on their global appearance in cluster of at least three cells and the appearance of the nuclei, which should be scattered irregularly through the cluster. Anisokaryosis, *i.e.* nucleus of variable sizes and shapes, is the most important criterion for *CTM*. *gCTCs* are individualized very large cells (50-300 µm), with a voluminous cytoplasm (very small N/C ratio), which can be round or oblong. The nucleus is often multilobular or multiple.

**Immunofluorescent staining of atypical circulating cells for confocal microscopic analysis**

Immunofluorescent staining was done directly on the filter. Briefly, 4% paraformaldehyde was used to fix cells on the filter for 5 minutes (Sigma Aldrich), TBS containing 0.2% Triton X-100 (Sigma Aldrich) was used to permeabilize cells for 5 minutes. After a quick rinse in water, the filter was incubated in blocking buffer containing 3% Bovine Serum Albumin, 1% Donkey serum and 1% Goat serum at room temperature for 30 minutes. Antibodies used are detailed in **Supplementary Table 1**. Primary antibodies were added in blocking buffer and incubated overnight at 4°C. Following three washes with TBS-0.05% Tween20, the secondary antibody mixture was added for 1 hour in the dark at room temperature. After three washes with TBS-0.05% Tween20, the coupled antibody CD45-A488 was incubated for 1 hour in the dark at room temperature. Finally, after three washes with TBS-0.05% Tween20 and water, the filter was counterstained with Sytox Blue Nucleic Acid Stain for 5 minutes at room temperature in the dark, to stain nuclei (Life Technologies, S11348). Finally, the filter was mounted with Kaiser Solution (Sigma Aldrich, 1092420100) and dried at room temperature a few more hours. Immunofluorescence was analyzed with a c-Plan-Apochromat 40x/1.3 oil objective on a LSM880 confocal with spectral detection from Zeiss equipped with a 405 laser diode, an Argon laser and

561 and 633 lasers. Imaged acquisition and spectra unmixing were done using the Zen Black software.

**Statistical analysis**

Correlations between CTCs subsets and clinico-pathological variables were analyzed using logistic regression (Logit link function). Progression-free survival (PFS) was calculated from the date of inclusion in PERMED-01 until the date of first progression or death from any cause. Overall survival (OS) was calculated from the date of inclusion in PERMED-01 until the date of death from any cause. Follow-up was measured from the date of diagnosis to the date of last news for event-free patients. Survivals were calculated using the Kaplan-Meier method and curves were compared with the log-rank test. Uni- and multivariate prognostic analyses were done using Cox regression analysis (Wald test). The variables submitted to univariate analyses included patients' age at inclusion, metastasis-to-diagnosis time, pathological type, grade, and molecular subtype of primary tumor, molecular subtype of metastasis biopsied in PERMED-01, (26) nature and number of metastatic sites at inclusion, number of previous lines of chemotherapy at inclusion, and CTC subsets. Multivariate analyses included the variables significant in univariate analysis ($p \leq 0.05$). The cell subset discretization was based on the distribution of values, and cut-offs were established using a two-component Gaussian finite Mixture Model (GMM) using maximum likelihood estimation on a per-study basis as previously described (27). The likelihood ratio (LR) tests were used to assess the prognostic information of one variable provided beyond that of another variable, assuming a $\chi^2$ distribution. Changes in the LR values ($\Delta$LR-$\chi^2$) measured quantitatively the relative amount of information of one model compared with another. All statistical tests were two-sided at the 5% level of significance. Statistical analysis was done using the survival package (version 2.30) in the R software

(version 2.9.1; http://www.cran.r-project.org/). We followed the reporting REcommendations for tumor MARKer prognostic studies (REMARK criteria).

## RESULTS

### Patients' population

Ninety-one mBC patients enrolled in the PERMED-01 trial and seven controls (*i.*e. healthy subjects without known cancer) were included in the present study. As shown in **Table 1**, the patients' median age at inclusion in PERMED-01 was 55 years (range, 27-79). The median time between diagnosis of metastatic relapse and primary cancer was 3 years (range, 0.5-23). Regarding the primary tumor, the most frequent pathological type was ductal (84%), the most frequent pathological grade was 3 (50%), and the molecular subtypes were mainly HR+/HER2- (56%), then triple-negative (TN) (32%), and HER2+ (12%). The molecular subtype of the PERMED-01 metastatic biopsy, available for 87 out of 91 patients, included more TN cases (46%) than HR+/HER2- (41%) and HER2+ (13%). Nineteen patients (22%) showed discordance between the molecular subtypes observed in the metastatic samples and the primary tumor. The median number of metastatic sites at inclusion in PERMED-01 was 3 (range, 1-8). The most frequent sites were bone (65%), then lymph node (63%) and liver (57%). The median number of lines of systemic chemotherapy received before inclusion was 4 (range, 0-8). With a median follow-up after inclusion of 12 months (range, 1-52), all but one patients showed disease progression and 68 died. The 1-year PFS was 13% (95%CI 8-22), median PFS was 5 months (range, 1-47), 1-year OS was 58% (95%CI 40-70) and median OS was 14 months (range, 1-52).

### Three subsets of atypical circulating cells in mBC patients

We screened all blood samples for the presence of atypical circulating cells, either individual or in cluster, after MGG coloration of a ScreenCell® Cyto filter. We observed several types of atypical circulating cells, which were classified into three subsets, as already described

(14) and based on precise cytological criteria: single/individual epithelioid cells or *iCTCs* **(Figure 1A)**, clusters of cells or *CTM* **(Figure 1B and Supplementary Figure 1A)***,* and giant epithelioid cells or *gCTCs* **(Figure 1C).**

Of note, large nuclei (greater than 20 µm), with irregular nuclear borders, dense and hyperchromatic, but with no clearly visible cytoplasm could often be observed with high frequency in patients' samples and referred to as "uncertain" specimens **(Supplementary Figure 1B)**; these are naked nuclei or residual apoptotic bodies. Complementary analysis is required to understand their identity: whether the pressure applied during the filtration blows apart the cytoplasm from the nucleus of fragile cells, or if the cell is undergoing apoptosis in the bloodstream, thereby displaying a retracted cytoplasm, remains to be determined. Because their origin is uncertain, we have not analyzed these events in the present study. Of note, they were not observed in control samples, and at this stage we cannot rule out a potential role in disease progression. Small (<20 µm) and round naked nuclei or residual apoptotic bodies could also occasionally be observed in patients and control samples **(Supplementary Figure 1C).** These events potentially arise from normal cells.

**Quantification of atypical circulating cells in mBC patients**

All but four patients (96%) had at least one atypical circulating cell (defined as the sum of *iCTCs*, *CTM,* and *gCTCs*) per mL of blood. The median number of CTCs per mL was 8.33 (range, 0-481.6) and the mean number was 27.17 [+/- 71.8]). The distribution of cell subsets in each atypical circulating cells count is shown in **Figure 1D** and described in **Supplementary Table 2.** The number of patients with at least one *iCTC* per mL was 72 (79%), the median number of *iCTCs* per mL was 2 (range, 0-51) and the mean number was 4 [+/- 6.7]. The number of patients with at least one *CTM* per mL was 47 (52%), the median number of *CTM* per mL was

1.33 (range, 0-479.6) and the mean number was 21.90 [+/- 71.1]). The number of patients with at least one *gCTC* per mL was 33 (36%), the median number of *CTM* per mL was 0 (range, 0-10.3) and the mean number was 1.2 [+/- 2.1]). Thus, the distribution of the different cell subsets was variable between patients: *iCTCs* were more frequently present than *CTM*, which were more frequently present than *gCTCs;* when present, the *CTM* and *iCTCs* were in larger concentration per mL of blood than were *gCTCs*.

**The intermediate EMT status of *CTM* and *gCTCs* (but not *iCTCs*) is associated with the co-expression of stemness and drug resistance markers**

To improve their characterization, we looked deeper into the biology of these cell subsets by analyzing specific molecular markers using advanced multicolor confocal analysis. We analyzed 23 samples, which included a total of 2,152 atypical circulating cells. These cells included a majority of *CTM* (1,742), a minority of *gCTCs* (74), and 336 *iCTCs*, in agreement with the distribution observed through the whole series of patients. We set up a 6-color immunofluorescence staining simultaneously targeting leukocytes (CD45), epithelial markers (EPCAM and Pan-cytokeratin: pan-KRT), mesenchymal marker (VIM), stem cell markers (CD133 (28) and/or LGR5 (29)), drug resistance markers (ABCG2 and/or ABCB1/P-GP/MDR1 (efflux pumps)), and SytoxBlue as a DNA labeling dye for nuclei (**Figure 2; Supplementary Figure 2**). We analyzed this combination of antibodies by spectral confocal microscopy. All cells retained for analysis did not show a neat CD45 ring at the cell surface and did not belong to the leukocyte's lineage. The staining observed with the combination of these molecular markers allowed to estimate, for each CTC subset, the epithelial-to-mesenchymal transition status, the stemness attributes (30), and the potential ability to resist to drugs based on the expression of efflux pumps (31). The results are shown in **Figure 3 A-B, and Supplementary Figure 3**.

A strictly epithelial (EPCAM+ and/or KRT+ but VIM-) status was found on very few circulating cells, independently of their subset (2.5% of *iCTCs*, 0.6% of *CTM,* and 0% of *gCTCs*). The balance between epithelial and mesenchymal markers was otherwise highly variable among the subsets. The phenotype of *iCTCs* was predominantly strictly mesenchymal at 64.7%, with only 26.3% of them presenting a dual epithelial/mesenchymal phenotype. The majority of cells composing *CTM* harbored a dual epithelial/mesenchymal phenotype (82.5%), with only 16.4% of them being mesenchymal-only. The *gCTCs* were also mostly double positive for epithelial and mesenchymal markers (63.6%), whereas 27.4% were strictly mesenchymal. Altogether, these results suggest that displaying a purely epithelial phenotype on atypical circulating cells is probably not favored in the blood flow. This could be because it is a better phenotype to detach from the tumor or because the environment in the blood might favor the survival of cells with a dual epithelial/mesenchymal phenotype or a strict mesenchymal phenotype (**Figure 3A**).

The LGR5 stem cell marker was expressed in 52% of *iCTCs*, 89% of *CTM*, and 99% of *gCTCs*. *CTM* and *gCTCs* seemed to be the subsets endowed with the highest stem potential (**Figure 3A**). Finally, we evaluated the expression of two efflux pumps, which can be involved in the mechanism of drug resistance, a specific trait from the PERMED-01 patients' cohort. Efflux pumps were expressed on 49% of *iCTCs*, 69% of *CTM,* and 89% of *gCTCs*, suggesting that the *gCTC* subset was the most likely to resist to systemic treatments (**Figure 3A).**

Interestingly, when we looked at the co-expression of all these markers, we observed that the pattern of EMT, stemness and/or drug resistance markers differed on the various cell subsets (**Figure 3B; Supplementary Figure 3**). The *iCTCs* were predominantly mesenchymal and displayed variable profiles for stem cells and/or drug resistance markers. The *CTM* subset, in its majority (83%), simultaneously expressed epithelial and mesenchymal markers. In this population, 60% of the cells also expressed stem cells and drug resistance markers. The expression pattern associated with purely mesenchymal *CTM* was less evident. The main co-

expression was that 13% of mesenchymal *CTM* co-expressed stem markers. Finally, *gCTCs* predominantly co-expressed stem and drug resistance markers, independently of their EMT status (56% of the cells with mixed EMT status and 33% of cells with mesenchymal status) (**Figure 3B; Supplementary Figure 3**).

Altogether, this analysis shows that *iCTCs* are a very peculiar category of atypical circulating cell with a majority of them being purely mesenchymal with random display of stem and efflux pump markers. Otherwise, *CTM* and *gCTCs* associate, in majority, an intermediate EMT state with the expression of stem markers, very often together with efflux markers. This confers specific attributes to those two subsets, which could contribute to their association with poorer prognosis.

### ***Atypical circulating cells subsets*** **and clinico-pathological correlations**

We then searched for correlations between all atypical circulating cells or each individual subsets and clinico-pathological features of patients and tumors (logit function test). Before analysis, we defined objectively, for each atypical circulating cells subset, a positivity cut-off using a two-component Gaussian finite Mixture Model. Above the cut-off, patients were considered “positive” for this atypical circulating cell subset. The number of positive patients was 56 (62%) for all atypical circulating cells (cut-off: >4.5 cells/mL), 49 (54%) for *iCTCs* (cut-off: >1.8 cells/mL), 47 (52%) for *CTM* (cut-off: >0.9 cells/mL), and 42 (46%) for *gCTCs* (cut-off: >0.3 cells/mL) **(Table 1)**.

The results of correlation analyses are shown in **Figure 4A**. Atypical circulating cells in their totality (“all atypical cells”) were associated with the presence of liver metastases (p=3.13E-02) and with the positivity of each cell subset (positivity is attributed to scores of CTCs above the cut-off defined for this subset). The correlations were different according to the subsets. The

presence of *iCTCs* was more frequently associated with the HR+/HER2- subtype of primary cancer (p=1.29E-02) and of metastatic lesion (p=3.22E-02) than with the TN subtype, and with the presence of liver metastases (p=4.56E-02), the absence of skin metastases (p=2.86E-02), and the positivity of *gCTCs* (p=2.29E-03). The presence of *CTM* was associated with a shorter time between metastatic relapse and diagnosis of primary cancer (p=4.51E-02) and the TN *versus* HR+/HER2- subtype of primary cancer (p=3.39E-02). The presence for *gCTCs* was associated with the existence of peritoneal metastases (p=3.75E-02) and the positivity of *iCTCs* (p=2.29E-03). None of the atypical circulating cells subsets was associated with patients' age, number of metastatic sites and number of previous lines of chemotherapy at inclusion, and pathological type and grade of primary tumor

***The gCTCs status is an independent prognostic factor for PFS and OS***

We then wondered if the presence and/or the different subsets of atypical circulating cells were associated with patients' survival. The results of uni- and multivariate analyses (Wald test) are shown in **Table 2**. In univariate analysis, patients with numbers of atypical circulating cells in their totality above the positivity cut-off ("positive") displayed shorter PFS than the "negative" (under the cut-off) patients (HR=1.37, 95%CI 0.89-2.10), but the difference was not significant (p=0.158). Such association was also not significant for *iCTCs* (HR=1.51, 95%CI 0.99-2.29; p=0.056) and for *CTM* (HR=1.16, 95%CI 0.77-1.76; p=0.478). By contrast, it was significant regarding *gCTCs* with a HR for PFS event of 1.94 (95%CI 1.24-3.01) in the "positive" patients as compared to the "negative" patients (p=2.98E-03). The 1-year PFS was 5% (95%CI 1-18) in the "*gCTC*-positive" group *versus* 20% (95%CI 12-35) in the "*gCTC*-negative" group (p=2.46E-03, log-rank test; **Figure 4B**). The other clinico-pathological features associated with PFS in univariate analysis included the patients' age at inclusion (p=3.34E-02), the pathological type of primary tumor (p=1.76E-02), and the number of previous lines of chemotherapy at inclusion

(p=4.20E-02). In multivariate analysis, the *gCTC* status remained significantly associated with shorter PFS (p=5.09E-03), suggesting independent prognostic value.

Similar results were observed with OS (**Table 2)**. In univariate analysis, the patients' status for all atypical circulating cells, *iCTCs*, and *CTM* was not significantly associated with OS. By contrast, it was significant regarding *gCTCs* with a HR for death of 2.46 (95%CI 1.47-4.12) in the "positive" patients as compared to the "negative" patients (p=5.84E-04). As shown in **Figure 4C**, the 1-year OS was 71% (95%CI 59-85) in the "*gCTC*-negative" group *versus* 42% (95%CI 29-62) in the "*gCTC*-positive" group (p=1.06E-04, log-rank test). The other variables associated with OS in univariate analysis included the pathological type of primary tumor (p=1.90E-02), the presence of brain/meningeal metastases (p=1.15E-03), and the number of previous lines of chemotherapy at inclusion (p=8.50E-03). Again, in multivariate analysis, the *gCTC* status remained significant and associated with shorter OS (p=7.26E-03), suggesting independent prognostic value.

We then explored the prognostic value of additional molecular features of atypical circulating cells in the small subgroup of 23 patients randomly chosen. Before analysis, the molecular status of cell subsets was discretized using a 50% positivity arbitrary cut-off: a patient was considered as positive when more than 50% of cells of this subset were stained for the given molecular marker. In univariate analysis for PFS (**Supplementary Table 3**), only the epithelial-mesenchymal mixed status was associated with shorter PFS (HR=2.87, 95%CI 1.08-7.68; p=3.53E-02, Wald test). Given the association between this molecular status and *gCTCs*, both associated with PFS, we tested their prognostic independence in multivariate analysis (**Supplementary Table 3**): interestingly, and despite the small number of samples, both variables tended towards significance with respective p-values of 0.082 and 0.054, suggesting that both cytological and molecular variables might provide complementary prognostic information. **Supplementary Figure 4A** shows the PFS of 23 patients according to the presence

or not of atypical circulating cells of intermediate EM status and the presence or not of *gCTCs*. The patients without *gCTCs* and without atypical circulating cell of intermediate EM status (*gCTC* 0 + EM 0) displayed 71% 6-months PFS (95%CI 45-100) *versus* 17% (95%CI 3-100) for patients with both statutes positive (*gCTC* 1 + EM 1) and 40% (95%CI 19-85) for patients with "hybrid" statutes (*gCTC* 0 + EM 1 and *gCTC* 1 + EM0) (p=1.12E-02, log-rank test). Such prognostic complementarity between both variables was tested using the likelihood ratio (LR) test: the *gCTC* status added prognostic information to that provided by the epithelio-mesenchymal mixed status ($\Delta$LR-$\chi^2$=4.08, p=4.34E-02; **Supplementary Figure 4B**), and conversely, the epithelio-mesenchymal mixed status tended to add prognostic information to that provided by the *gCTC* status ($\Delta$LR-$\chi^2$=2.89, p=0.089). This result suggested that investigating both the molecular phenotype of atypical circulating cells and their cytological aspect could improve the prognostic value of CTCs.

## DISCUSSION

It is accepted that circulating epithelial cells, defined on Cellsearch™ criteria as EPCAM$^{+}$ CD45$^{-}$ circulating cells, represent a new outcome-associated biomarker independent from known prognostic factors in metastatic breast cancers (4,5,32-34). Recent and more sensitive technologies revealed that multiple cell subsets can be found in the blood of patients with solid tumors (14,22,24,35). Single-use filter-based systems provide a high degree of sensitivity to isolate atypical rare circulating cells of “large” size, while letting most erythrocytes and leukocytes pass through (14). This revealed the diversity of CTCs regarding the maintenance of their epithelial status, their size and presentation. With this high sensitivity other atypical, potentially non-malignant cells, might also be trapped on the filters. Those can be endothelial cells which, although not supposedly found in the blood flow, can however be identified in very rare occasions, notably in cancer patients (35). We barely detected any endothelial cells in the analyzed samples. However, we detected large cells that we named *gCTCs,* not found in blood in healthy situation. They are large cells that were described as CAMLs (cancer-associated macrophage-like cells), reported to come in five morphological variants (amorphous, oblong, spindle-shaped, round, and tadpole-shaped) (24). A recent theory suggests that those cells might be fusion hybrids between tumor-associated macrophages (TAM) and malignant cells (36,37). We found that EPCAM and KRTs were present in the majority of *gCTC* (as shown in **Supplementary Figure 2** and **Figure 3A**), unmistakably identifying them as cells of epithelial origin. In the other cases, the macrophage origin of *gCTCs* in our study is probably invalid because *gCTCs* did not display CD45 at their membrane. It is interesting to note that if *gCTCs* are the result of cell fusion, their potential polyploidy could enhance their chances to adapt to environmental changes, which would participate in the increased tumorigenic potential compared to other subsets of CTCs. Understanding the nature of those cells could help deciphering their role in cancer evolution. Notably, it would be of great interest in future studies

to analyze the transcriptomic profile of these cells. In the meantime, we observed that the majority of these *gCTCs* (>98%) displayed high level of the LGR5 stem cell marker as well as the efflux pump ABCB1 suggesting a strong potential as tumor-initiating and drug-resistant cells respectively. Their numbers were previously shown to increase in chemotherapy-treated but not in hormone-therapy treated patients, suggesting that the release into the circulation of some *gCTCs* might be influenced by the type of treatment (24). Finally, *gCTCs* were correlated to peritoneal metastasis **(Figure 4)**. Peritoneum metastasis in breast cancer probably arises with systemic spreading of tumor cells. The molecular characteristics (stem and drug resistance) of *gCTC*s might result in better survival in the systemic circulation and thus increased numbers of *gCTCs* and/or increased capacity to form metastases. Peritoneal metastasis from breast cancer is a rare and challenging clinical presentation in which a better understanding of *gCTCs*' role could be helpful. Adding to the fact that *gCTCs* remain an intriguing population, we found that their presence was an independent predictor for PFS and OS. This relationship with OS did not depend on a defined threshold but only on their presence or not.

Recently, the center of attention has been drawn on the *CTM* population (9,38-40). Although less frequently observed in the circulation than *iCTCs*, *CTM* have 23- to 50-fold increased metastatic potential (22,38). In line with this, we found that 89% of cells found within cluster were positive for the LGR5 stem cell marker, whereas only 52% of *iCTCs* were faintly LGR5-positive. Interestingly, the detection of stem cell markers was not necessarily observed on the most mesenchymal cells, but rather on cells with a mixed epithelial/mesenchymal phenotype. In a cohort of patients with progressing metastatic breast cancer, 30% of patients had *CTM* using the Cluster-iChip technology and their presence correlated with shorter PFS in a cohort of 30 patients (41). We found a higher prevalence of *CTM* in our cohort of patients (52%), potentially explained by the more advanced metastatic status of our cohort. It is also possible that the filter-based technology we used was better suited to isolate *CTM* that the CTC-iChip

technology (41). We were not able to correlate the presence of *CTM* to shorter patients' PFS, nor OS despite a larger cohort size (n = 91). Our cut-off selection might not be optimal, but aims at being objective, being based on the distribution of data (GMM). Indeed, we have tested other arbitrary cut-offs for *CTM* and *iCTCs*, and found that some, like *CTM* cut-off < or > to 15) have a prognostic value (**Supplementary Figure 5)**. However, these thresholds need larger series for validation in independent samples, which is why we pursued with the GMM cut-off. The lack of correlation between the presence of *CTM* and shorter PFS or OS is also probably inherent to the specificity of this cohort, which is composed of advanced mBC. One can hypothesize that at a later stage, the predictive value of *CTM* is less powerful than at earliest stages. In this case, as previously reported for *CTCs*, a variation between time points might be more informative (42,43).

Outside the fact that *CTM* expressed high level of LGR5 stem cell marker, they also showed the highest level of epithelio-mesenchymal mixed phenotype. This might contribute to their previously reported association with a poor prognosis as it was reported that cells with a dual epithelio-mesenchymal phenotype are endowed with higher metastatic risk (44). The proposed explanation is that the mesenchymal phenotype confers increased invasiveness to cancer cells, and the epithelial phenotype higher proliferation abilities. Our phenotypic analysis of *CTM* provides additional rationale for their high metastatic potential.

Our work presents the largest cohort of mBC patients analyzed so far using the ScreenCell®Cyto system. We show, for the first time with this device, that CTCs can be considered as a robust independent prognosis factor in mBC patients, both regarding progression-free and overall survivals. Of note that patients in our study had received many lines of treatment. Positivity for *gCTCs* might help oncologists identify which patients, after several lines of systemic therapy, might benefit from best supportive care alone. Or, if validated, the identification of *gCTC*s might provide a tool for better prognostic stratification in early clinical trials testing novel therapeutics, which frequently enroll late-stage, already pre-treated patients.

Altogether, this study comforts the validity and utility of this easy, relatively inexpensive and convenient system to study CTCs in breast cancer patients. Finally, it also highlights the involvement of an interesting atypical circulating cell subsets, the *gCTCs*, as a prognostic factor for PFS, which can be improved when combined with complementary immunofluorescent staining and confocal microscopy analysis. A better characterization of molecular make-up of these *gCTCs* might also help identifying potential targets for future anti-metastasis-treatment.

In conclusion, this original study focuses on the identification of subtypes of CTCs, rather than the mere quantification of total CTCs. It allowed the identification of *gCTCs* as potential prognostic biomarker in metastatic breast cancers, which can potentially be refined by immunophenotypic analysis. The correlation with clinical and molecular data is new, and deserves to be confirmed on larger series. The use of immunohistochemistry also indicated that the co-expression of stem and drug resistance markers is associated with intermediate epithelial-mesenchymal cells in *CTM* and *gCTCs,* but not in *iCTCs*. This may explain why *CTM* and *gCTCs* are associated with poor prognosis through stem and drug resistance attributes.

## REFERENCES


1. Gupta GP, Massague J. Cancer metastasis: building a framework. *Cell* 2006;**127**(4):679-95 doi S0092-8674(06)01414-0 [pii]

10.1016/j.cell.2006.11.001.

2. Dasgupta A, Lim AR, Ghajar CM. Circulating and disseminated tumor cells: harbingers or initiators of metastasis? *Mol Oncol* 2017;**11**(1):40-61 doi 10.1002/1878-0261.12022.
3. Marrinucci D, Bethel K, Kolatkar A, Luttgen MS, Malchiodi M, Baehring F*, et al.* Fluid biopsy in patients with metastatic prostate, pancreatic and breast cancers. *Phys Biol* 2012;**9**(1):016003 doi 10.1088/1478-3975/9/1/016003

S1478-3975(12)97631-6 [pii].

4. Cristofanilli M, Budd GT, Ellis MJ, Stopeck A, Matera J, Miller MC*, et al.* Circulating tumor cells, disease progression, and survival in metastatic breast cancer. *N Engl J Med* 2004;**351**(8):781-91 doi 10.1056/NEJMoa040766

351/8/781 [pii].

5. Cristofanilli M, Hayes DF, Budd GT, Ellis MJ, Stopeck A, Reuben JM*, et al.* Circulating tumor cells: a novel prognostic factor for newly diagnosed metastatic breast cancer. *J Clin Oncol* 2005;**23**(7):1420-30 doi 23/7/1420 [pii]

10.1200/JCO.2005.08.140.

6. Cohen SJ, Punt CJ, Iannotti N, Saidman BH, Sabbath KD, Gabrail NY*, 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**(19):3213-21 doi 10.1200/JCO.2007.15.8923

26/19/3213 [pii].

7. de Bono JS, Scher HI, Montgomery RB, Parker C, Miller MC, Tissing H*, et al.* Circulating tumor cells predict survival benefit from treatment in metastatic castration-resistant prostate cancer. *Clin Cancer Res* 2008;**14**(19):6302-9 doi 10.1158/1078-0432.CCR-08-0872

14/19/6302 [pii].

8. Plaks V, Koopman CD, Werb Z. Cancer. Circulating tumor cells. *Science* 2013;**341**(6151):1186-8 doi 10.1126/science.1235226

341/6151/1186 [pii].

9. Paoletti C, Miao J, Dolce EM, Darga EP, Repollet MI, Doyle GV*, et al.* Circulating Tumor Cell Clusters in Patients with Metastatic Breast Cancer: a SWOG S0500 Translational Medicine Study. *Clin Cancer Res* 2019;**25**(20):6089-97 doi 10.1158/1078-0432.CCR-19-0208

1078-0432.CCR-19-0208 [pii].

10. Goodman CR, Seagle BL, Friedl TWP, Rack B, Lato K, Fink V*, et al.* Association of Circulating Tumor Cell Status With Benefit of Radiotherapy and Survival in Early-Stage Breast Cancer. *JAMA Oncol* 2018;**4**(8):e180163 doi 10.1001/jamaoncol.2018.0163

2679564 [pii].

11. Thery L, Meddis A, Cabel L, Proudhon C, Latouche A, Pierga JY*, et al.* Circulating Tumor Cells in Early Breast Cancer. *JNCI Cancer Spectr* 2019;**3**(2):pkz026 doi 10.1093/jncics/pkz026.

12. Werner S, Stenzl A, Pantel K, Todenhofer T. Expression of Epithelial Mesenchymal Transition and Cancer Stem Cell Markers in Circulating Tumor Cells. *Adv Exp Med Biol* 2017;**994**:205-28 doi 10.1007/978-3-319-55947-6_11.
13. Desitter I, Guerrouahen BS, Benali-Furet N, Wechsler J, Janne PA, Kuang Y*, et al.* A new device for rapid isolation by size and characterization of rare circulating tumor cells. *Anticancer Res* 2011;**31**(2):427-41 doi 31/2/427 [pii].
14. Mu Z, Benali-Furet N, Uzan G, Znaty A, Ye Z, Paolillo C*, et al.* Detection and Characterization of Circulating Tumor Associated Cells in Metastatic Breast Cancer. *Int J Mol Sci* 2016;**17**(10) doi E1665 [pii]

ijms17101665 [pii]

10.3390/ijms17101665.
15. Mascalchi M, Falchini M, Maddau C, Salvianti F, Nistri M, Bertelli E*, et al.* Prevalence and number of circulating tumour cells and microemboli at diagnosis of advanced NSCLC. *J Cancer Res Clin Oncol* 2016;**142**(1):195-200 doi 10.1007/s00432-015-2021-3

10.1007/s00432-015-2021-3 [pii].
16. Awe JA, Saranchuk J, Drachenberg D, Mai S. Filtration-based enrichment of circulating tumor cells from all prostate cancer risk groups. *Urol Oncol* 2017;**35**(5):300-9 doi S1078-1439(16)30415-X [pii]

10.1016/j.urolonc.2016.12.008.
17. Nicolazzo C, Raimondi C, Francescangeli F, Ceccarelli S, Trenta P, Magri V*, et al.* EpCAM-expressing circulating tumor cells in colorectal cancer. *Int J Biol Markers* 2017;**32**(4):e415-e20 doi A297E484-D303-40BA-B2BA-D8C5DB17EC1F [pii]

10.5301/ijbm.5000284.
18. Rosenbaum MW, Cauley CE, Kulemann B, Liss AS, Castillo CF, Warshaw AL*, et al.* Cytologic characteristics of circulating epithelioid cells in pancreatic disease. *Cancer* 2017;**125**(5):332-40 doi 10.1002/cncy.21841.
19. Drucker A, Teh EM, Kostyleva R, Rayson D, Douglas S, Pinto DM. Comparative performance of different methods for circulating tumor cell enrichment in metastatic breast cancer patients. *PLoS One* 2020;**15**(8):e0237308 doi 10.1371/journal.pone.0237308.
20. Rosenbaum MW, Cauley CE, Kulemann B, Liss AS, Castillo CF, Warshaw AL*, et al.* Cytologic characteristics of circulating epithelioid cells in pancreatic disease. *Cancer Cytopathol* 2017;**125**(5):332-40 doi 10.1002/cncy.21841.
21. Krebs MG, Hou JM, Sloane R, Lancashire L, Priest L, Nonaka D*, 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**(2):306-15 doi 10.1097/JTO.0b013e31823c5c16.
22. Aceto N, Bardia A, Miyamoto DT, Donaldson MC, Wittner BS, Spencer JA*, et al.* Circulating tumor cell clusters are oligoclonal precursors of breast cancer metastasis. *Cell* 2014;**158**(5):1110-22 doi 10.1016/j.cell.2014.07.013

S0092-8674(14)00927-1 [pii].
23. Au SH, Storey BD, Moore JC, Tang Q, Chen YL, Javaid S*, et al.* Clusters of circulating tumor cells traverse capillary-sized vessels. *Proc Natl Acad Sci U S A* 2016 doi 201524448 [pii]

1524448113 [pii]

10.1073/pnas.1524448113.
24. Adams DL, Martin SS, Alpaugh RK, Charpentier M, Tsai S, Bergan RC, *et al.* Circulating giant macrophages as a potential biomarker of solid tumors. *Proc Natl Acad Sci U S A* 2014;**111**(9):3514-9 doi 10.1073/pnas.1320198111

1320198111 [pii].
25. Wechsler J. Circulating Tumor cells from solid cancers. 2015.
26. McShane LM, Altman DG, Sauerbrei W, Taube SE, Gion M, Clark GM, *et al.* REporting recommendations for tumour MARKer prognostic studies (REMARK). *Br J Cancer* 2005;**93**(4):387-91 doi 10.1038/sj.bjc.6602678.
27. Lehmann BD, Bauer JA, Chen X, Sanders ME, Chakravarthy AB, Shyr Y, *et al.* Identification of human triple-negative breast cancer subtypes and preclinical models for selection of targeted therapies. *J Clin Invest* 2011;**121**(7):2750-67 doi 10.1172/JCI45014.
28. Brugnoli F, Grassilli S, Al-Qassab Y, Capitani S, Bertagnolo V. CD133 in Breast Cancer Cells: More than a Stem Cell Marker. *J Oncol* 2019;**2019**:7512632 doi 10.1155/2019/7512632.
29. Yang L, Tang H, Kong Y, Xie X, Chen J, Song C, *et al.* LGR5 Promotes Breast Cancer Progression and Maintains Stem-Like Cells Through Activation of Wnt/beta-Catenin Signaling. *Stem Cells* 2015;**33**(10):2913-24 doi 10.1002/stem.2083.
30. De Angelis ML, Francescangeli F, Zeuner A. Breast Cancer Stem Cells as Drivers of Tumor Chemoresistance, Dormancy and Relapse: New Challenges and Therapeutic Opportunities. *Cancers (Basel)* 2019;**11**(10) doi E1569 [pii]

10.3390/cancers11101569

cancers11101569 [pii].
31. Pasello M, Giudice AM, Scotlandi K. The ABC subfamily A transporters: multifaceted players with incipient potentialities in cancer. *Semin Cancer Biol* 2019 doi S1044-579X(19)30250-0 [pii]

10.1016/j.semcancer.2019.10.004.
32. Bidard FC, Peeters DJ, Fehm T, Nole F, Gisbert-Criado R, Mavroudis D, *et al.* Clinical validity of circulating tumour cells in patients with metastatic breast cancer: a pooled analysis of individual patient data. *Lancet Oncol* 2014;**15**(4):406-14 doi 10.1016/S1470-2045(14)70069-5

S1470-2045(14)70069-5 [pii].
33. Yan WT, Cui X, Chen Q, Li YF, Cui YH, Wang Y, *et al.* Circulating tumor cell status monitors the treatment responses in breast cancer patients: a meta-analysis. *Sci Rep* 2017;**7**:43464 doi 10.1038/srep43464.
34. Cristofanilli M, Pierga JY, Reuben J, Rademaker A, Davis AA, Peeters DJ, *et al.* The clinical use of circulating tumor cells (CTCs) enumeration for staging of metastatic breast cancer (MBC): International expert consensus paper. *Crit Rev Oncol Hematol* 2019;**134**:39-45 doi 10.1016/j.critrevonc.2018.12.004.
35. Williamson SC, Metcalf RL, Trapani F, Mohan S, Antonello J, Abbott B, *et al.* Vasculogenic mimicry in small cell lung cancer. *Nat Commun* 2016;**7**:13322 doi 10.1038/ncomms13322.
36. Zhang Y, Zhou N, Yu X, Zhang X, Li S, Lei Z, *et al.* Tumacrophage: macrophages transformed into tumor stem-like cells by virulent genetic material from tumor cells. *Oncotarget* 2017;**8**(47):82326-43 doi 10.18632/oncotarget.19320

19320 [pii].

37. Gast CE, Silk AD, Zarour L, Riegler L, Burkhart JG, Gustafson KT*, et al.* Cell fusion potentiates tumor heterogeneity and reveals circulating hybrid cells that correlate with stage and survival. *Sci Adv* 2018;**4**(9):eaat7828 doi 10.1126/sciadv.aat7828

aat7828 [pii].

38. Wang C, Mu Z, Chervoneva I, Austin L, Ye Z, Rossi G*, et al.* Longitudinally collected CTCs and CTC-clusters and clinical outcomes of metastatic breast cancer. *Breast Cancer Res Treat* 2017;**161**(1):83-94 doi 10.1007/s10549-016-4026-2

10.1007/s10549-016-4026-2 [pii].

39. Larsson AM, Jansson S, Bendahl PO, Levin Tykjaer Jorgensen C, Loman N, Graffman C*, et al.* Longitudinal enumeration and cluster evaluation of circulating tumor cells improve prognostication for patients with newly diagnosed metastatic breast cancer in a prospective observational trial. *Breast Cancer Res* 2018;**20**(1):48 doi 10.1186/s13058-018-0976-0

10.1186/s13058-018-0976-0 [pii].

40. Costa C, Muinelo-Romay L, Cebey-Lopez V, Pereira-Veiga T, Martinez-Pena I, Abreu M*, et al.* Analysis of a Real-World Cohort of Metastatic Breast Cancer Patients Shows Circulating Tumor Cell Clusters (CTC-clusters) as Predictors of Patient Outcomes. *Cancers (Basel)* 2020;**12**(5) doi 10.3390/cancers12051111.

41. Sarioglu AF, Aceto N, Kojic N, Donaldson MC, Zeinali M, Hamza B*, et al.* A microfluidic device for label-free, physical capture of circulating tumor cell clusters. *Nat Methods* 2015;**12**(7):685-91 doi 10.1038/nmeth.3404

nmeth.3404 [pii].

42. Jauch SF, Riethdorf S, Sprick MR, Schutz F, Schonfisch B, Brucker SY*, et al.* Sustained prognostic impact of circulating tumor cell status and kinetics upon further progression of metastatic breast cancer. *Breast Cancer Res Treat* 2019;**173**(1):155-65 doi 10.1007/s10549-018-4972-y.

43. Jansson S, Bendahl PO, Larsson AM, Aaltonen KE, Ryden L. Prognostic impact of circulating tumor cell apoptosis and clusters in serial blood samples from patients with metastatic breast cancer in a prospective observational cohort. *BMC Cancer* 2016;**16**:433 doi 10.1186/s12885-016-2406-y

10.1186/s12885-016-2406-y [pii].

44. Zheng X, Carstens JL, Kim J, Scheible M, Kaye J, Sugimoto H*, et al.* Epithelial-to-mesenchymal transition is dispensable for metastasis but induces chemoresistance in pancreatic cancer. *Nature* 2015;**527**(7579):525-30 doi 10.1038/nature16064

nature16064 [pii].

## TABLES

**Table 1: Patient and tumor characteristics**

Clinical data and CTCs status (negative vs positive). The positivity cut-off was defined using a two-component Gaussian finite Mixture Model (GMM)(cf text). Above the cut-off, patients were considered “positive” for this atypical circulating cell subset.

**Table 2: Univariate and multivariate analysis for PFS and OS**

## FIGURE LEGENDS

**FIGURE 1. Pictures of atypical circulating cells observed in the blood of patients with mBC and subsets distribution for 91 patients of the cohort.**

**A-C.** Cytological staining with May-Grünwald solution of cells isolated from blood samples on a filter. Small black dots are filter pores (for clarity some examples are marked with a white asterix). Scale represents 50µm. Cells of interest are marked with an arrow. The figure shows representative examples of: **A.** Single epithelioid cells or *iCTCs*, **B.** Cluster of epithelioid cells or *CTM*, **C.** Giant epithelioid cells or *gCTCs*. **D.** Graphic representation of the result of the screening of 91 metastatic cancer patients and 7 cancer-free donors as controls, indicating the number of *iCTC*, *CTM* and *gCTC* per ml in each case.

**FIGURE 2. Examples of the immunofluorescence staining of *CTM* isolated on ScreenCell® filters.**

The blood of a mBC patient was filtered and cells on the filter were immunostained. Blood cells are detected with the expression of CD45 antigen (see SD2 empty arrowheads), whereas CTCs are detected *via* the expression of epithelial markers (EPCAM and pan-KRT) and mesenchymal marker (VIM). Stem cell marker (Lgr5) and drug resistance marker (PgP) expression are also assessed. Representative image of a *CTM* with intermediate EMT status and no expression of stemness or drug resistance markers. Scale bar represents 10µm.

**FIGURE 3. Correlations between CTC subsets and molecular markers.**

**A.** Data for the phenotypic markers expressed by atypical circulating cell subsets. **B.** Alluvial plot representation of the correlations between CTC subsets and molecular markers. The graph shows for the three substets of atypical cells (*CTM, gCTCs*, and *iCTCs*) the combined expression of EMT, stemness and efflux pump markers. The height of the blocks represents the size of the population: in the left column, the pink block represents the amount of *CTM*, the green one of *gCTCs*, and blue one of *iCTCs*; in the middle column the white block represents cells with epithelial status, light grey with mixed EMT status, and grey with mesenchymal status; in the right hand side column the light pink block the cells negative for both stemness and efflux pump markers, the light orange block cells expressing the tested efflux pump, the orange block represents the amount of cells expressing stem makers tested, and the yellow block cells expressing both stemness and efflux pump markers. The thickness of a stream represents the amount of cells contained in blocks interconnected by the stream.

**FIGURE 4: Correlations between CTCs and CTC subsets and clinicopathological variables and correlation between *gCTCs*' presence and progression-free survival (PFS) and overall survival (OS).**

**A.** Forest plot representation of the correlation between the total number of atypical cells detected, or individual atypical cell subsets (*iCTCs*, *CTM*, *gCTCs*), and clinical data for all 91 patients. Odds ratio are indicated with confidence intervals (horizontal lines). Data statistically significant are in black, data statistically insignificant (crossing 0, vertical line) in grey. TN: triple-negative, HR+: Hormone Receptive-positive, HER2-/+: HER2 negative/positive. The bottom part of the forest plot shows the correlation between the total number of atypical cells and the different subsets. **B-C.** Kaplan-Meier PFS (**B**) and OS (**C**) curves of mBC patients according to the presence of *gCTCs* (data from table 2).

**SUPPLEMENTARY DATA:**

**Supplementary Table 1: Antibodies used for immunofluorescence**

**Supplementary Table 2: Distribution of CTC subsets in the mBC population**

**Supplementary Table 3: Univariate and multivariate analysis for PFS according to molecular and cytological profiles**

**Supplementary Figure 1: Pictures of atypical circulating cells observed in the blood of patients with mBC.**

Cytological staining with May-Grünwald solution of cells isolated from blood samples on a filter. Small black dots are filter pores (for clarity some examples are marked with a white asterix). Scale represents 50µm. Cells of interest are marked with an arrow. The figure shows representative examples of: **A.** Clusters of epithelioid cells of *CTM*, **B.** "Uncertain" specimens, including cells with no clearly visible cytoplasm and a cell with N/C ratio <0.75, **C.** Round naked nuclei or residual apoptotic bodies. Scale represents 50µm.

**Supplementary Figure 2: Examples of the immunofluorescence staining of atypical cells isolated on ScreenCell® filters.**

The blood of mBC patients were filtered and cells on the filter were immunostained. Blood cells are detected with the expression of CD45 antigen (empty arrowheads), whereas CTCs are detected *via* the expression of epithelial markers (EPCAM and pan-KRT) and mesenchymal marker (VIM). Stem cell marker (Lgr5) and drug resistance marker (PgP) expression are also

assessed. Representative images of **A.** *iCTCs,* **B.** *CTM* and **C.** *gCTCs*, are shown and indicated by white arrowheads. Scale bar represents 10µm.

**Supplementary Figure 3: Alluvial plot representation of the combined expression of EMT, stem and drug resistance markers on CTCs subsets individually (*CTM, gCTCs*, and *iCTCs*).**

The same data as in figure 3 are represented for each cell subset separately. Alluvial plot representation of the correlations between CTC subsets and molecular markers. The graph shows for the three subsets of atypical cells separately (for *CTM* in **A.***,* for *gCTCs* in **B.,** and for *iCTCs* in **C.**) the combined expression of EMT, stemness and efflux pump markers. The height of the blocks represents the size of the population. The thickness of a stream represents the amount of cells contained in blocks interconnected by the stream. EMT status: Ep=epithelial, Mixed=epithelial + mesenchymal, Mes=mesenchymal; Other markers: Efflux=Efflux pump, Stem=stemness marker, Mixed= Stemness + Efflux pump markers.

**Supplementary Figure 4: Kaplan-Meier curves of OS and PFS of patients based on the presence of *gCTCs* and intermediate EM status.**

**A.** Survivals were calculated using the Kaplan-Meier method and were compared with the log-rank test to evaluate the prognostic value of combined *gCTC* and intermediate EM statuses. Presence of *gCTC* is marked as 1 vs absence as 0. Presence of an intermediate EM status in all CTCs' subsets is represented as 1 value, absence as 0. **B.** Prognostic complementarity: the values are given for prognostic information of each variable colored in grey (Mixed EM and *gCTC*) on its own (LR-$\chi 2$) and when added to the other variable colored in blue ($\Delta$LR-$\chi 2$). * indicates $p \leq 0.05$ and indicates trend for significance with $p \leq 0.10$.

**Supplementary Figure 5: Kaplan-Meier curve of OS for patients based on the global population of CTCs (A) or on individual subsets (B: *iCTCs*, C: *CTM*, and D: *gCTCs*).**

For *gCTCs* their presence or absence was considered, and cut-off values were applied for all atypical circulating cells, *CTM* and *iCTCs* (based on arbitrary set cut-offs).

**Table 1. Patient and tumor characteristics**

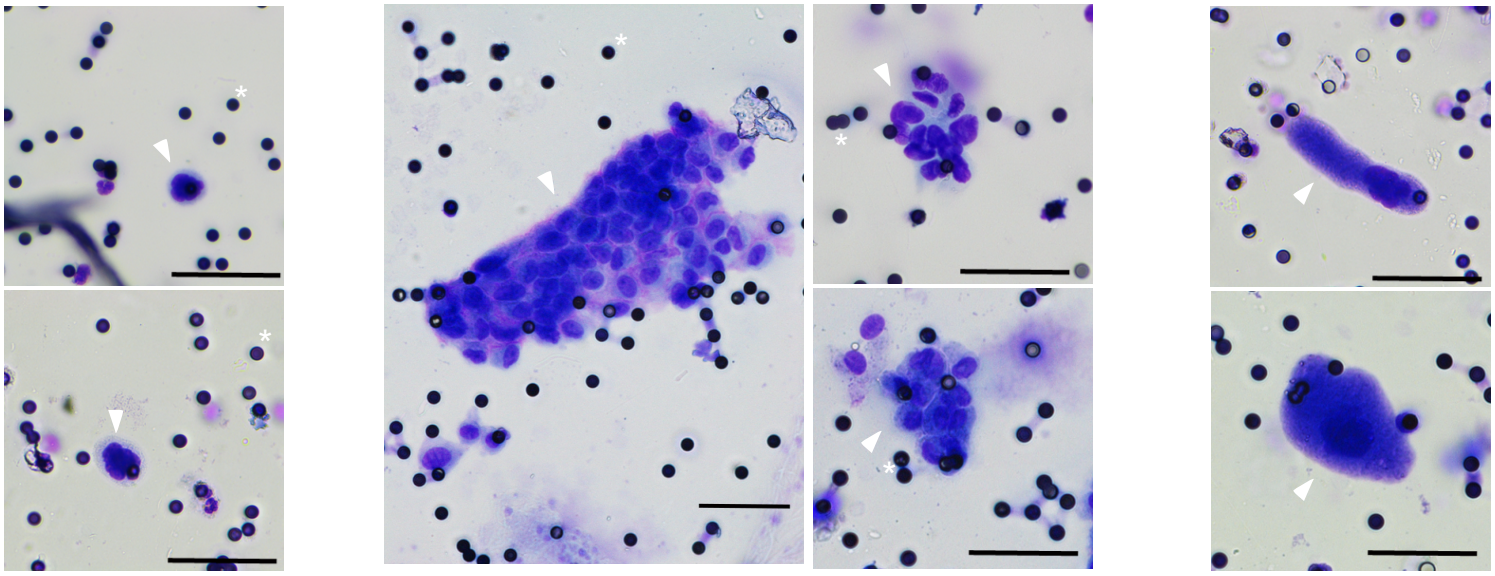
| Characteristics | | N (%) |
|---|---|---|
| Age at inclusion, years | | 55 (27-79) |
| Metastasis to diagnosis interval, years | | 3 (0.59-23) |
| Pathological type of primary tumor | ductal | 68 (84%) |
| | lobular | 6 (7%) |
| | other | 7 (9%) |
| Pathological grade of primary | 1 | 5 (7%) |
| | 2 | 32 (43%) |
| | 3 | 37 (50%) |
| Molecular subtype of primary | HR+/HER2- | 50 (56%) |
| | HER2+ | 11 (12%) |
| | TN | 29 (32%) |
| Molecular subtype of metastasis | HR+/HER2- | 36 (41%) |
| | HER2+ | 11 (13%) |
| | TN | 40 (46%) |
| Bone metastasis | no | 32 (35%) |
| | yes | 59 (65%) |
| Liver metastasis | no | 39 (43%) |
| | yes | 52 (57%) |
| Lung-pleural metastasis | no | 45 (49%) |
| | yes | 46 (51%) |
| Brain-meningeal metastasis | no | 81 (89%) |
| | yes | 10 (11%) |
| Lymph node metastasis | no | 34 (37%) |
| | yes | 57 (63%) |
| Skin metastasis | no | 73 (80%) |
| | yes | 18 (20%) |
| Peritoneum metastasis | no | 81 (89%) |
| | yes | 10 (11%) |
| Other metastatic site | no | 66 (73%) |
| | yes | 25 (27%) |
| Number of metastasic sites at inclusion, N | | 3 (1-8) |
| Number of previous chemotherapy lines at inclusion, N | | 4 (0-8) |
| Atypical circulating cells (all subsets) | negative | 35 (38%) |
| | positive | 56 (62%) |
| *gCTC* | negative | 49 (54%) |
| | positive | 42 (46%) |
| *iCTC* | negative | 42 (46%) |
| | positive | 49 (54%) |
| *CTM* | negative | 44 (48%) |
| | positive | 47 (52%) |
| Follow-up median, months (range) | | 12 (1-52) |
| PFS events, N (%) | | 90 (99%) |
| median PFS time, month (min-max) | | 5 (1-47) |
| 1-year PFS, % [95% CI] | | 13% [8-22] |
| OS events, N (%) | | 68 (75%) |
| median OS time, month (min-max) | | 14 (1-52) |
| 1-year OS, % [95% CI] | | 58% [49-70] |

# Figure 1

**A.** Single epitheliod cells or *iCTCs*

**B.** Cluster of epitheliod cells or *CTM*

**C.** Giant epitheliod cells or *gCTCs*



**D.**

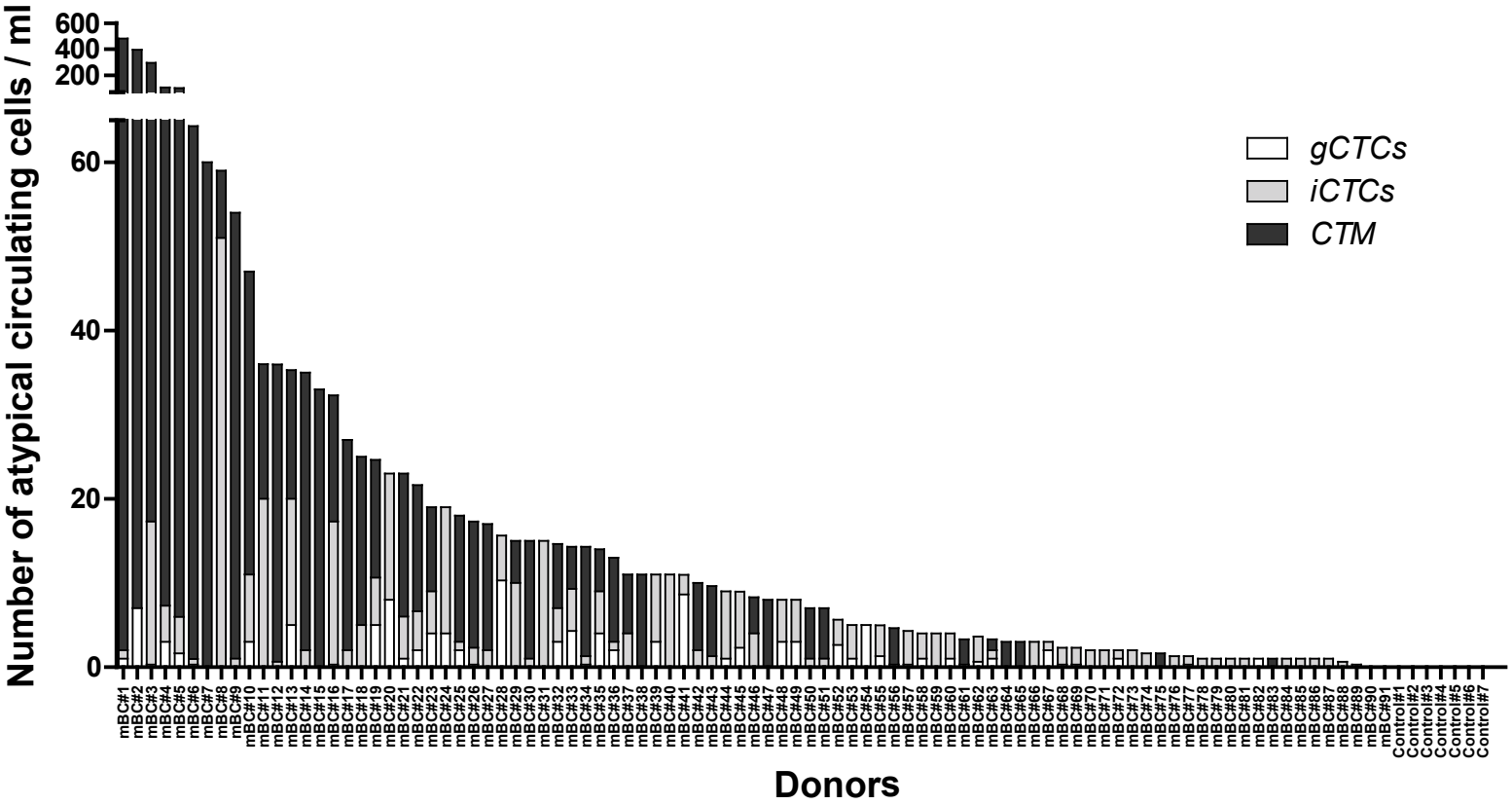
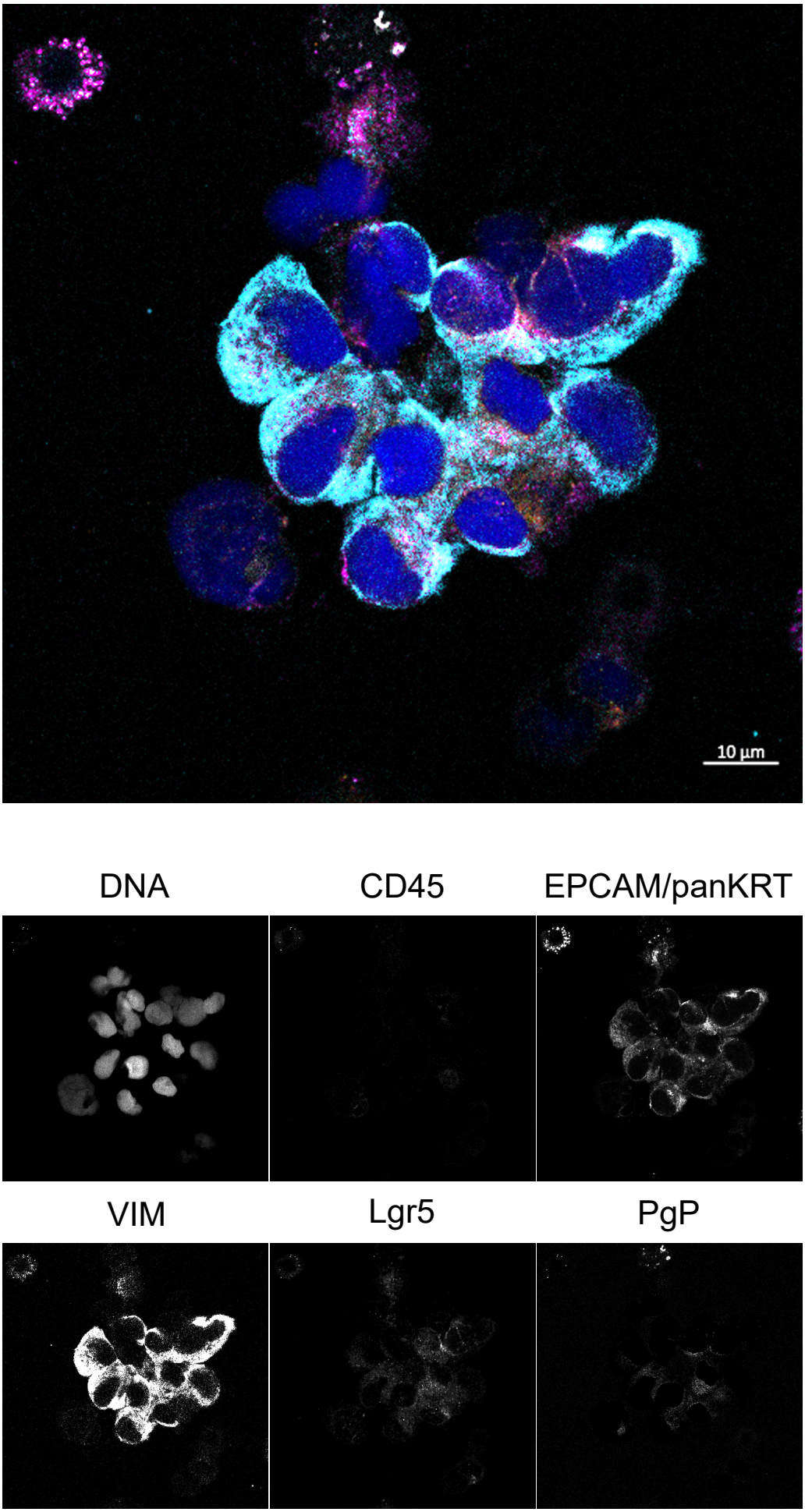


Figure 2

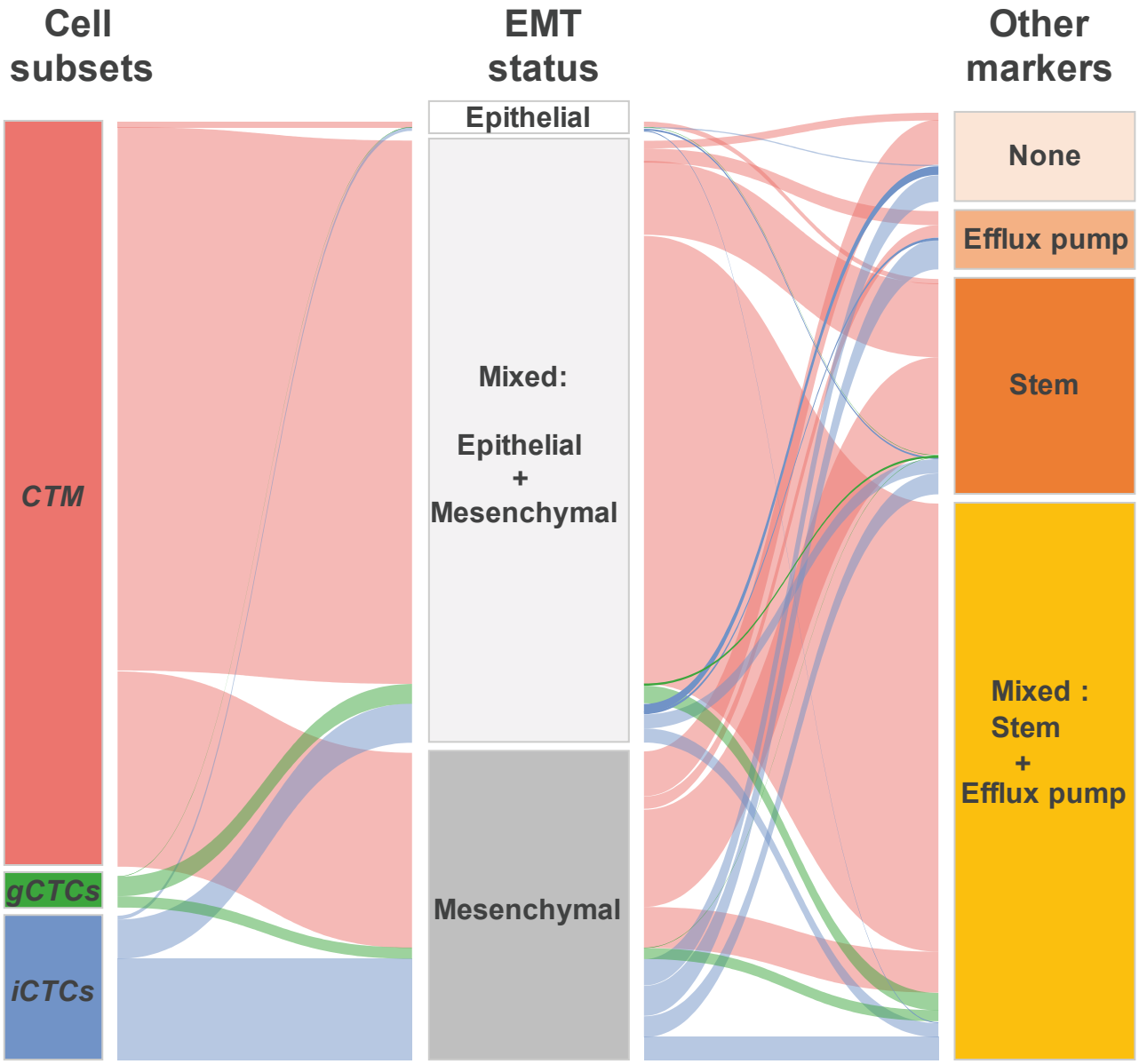


# Figure 3

## A

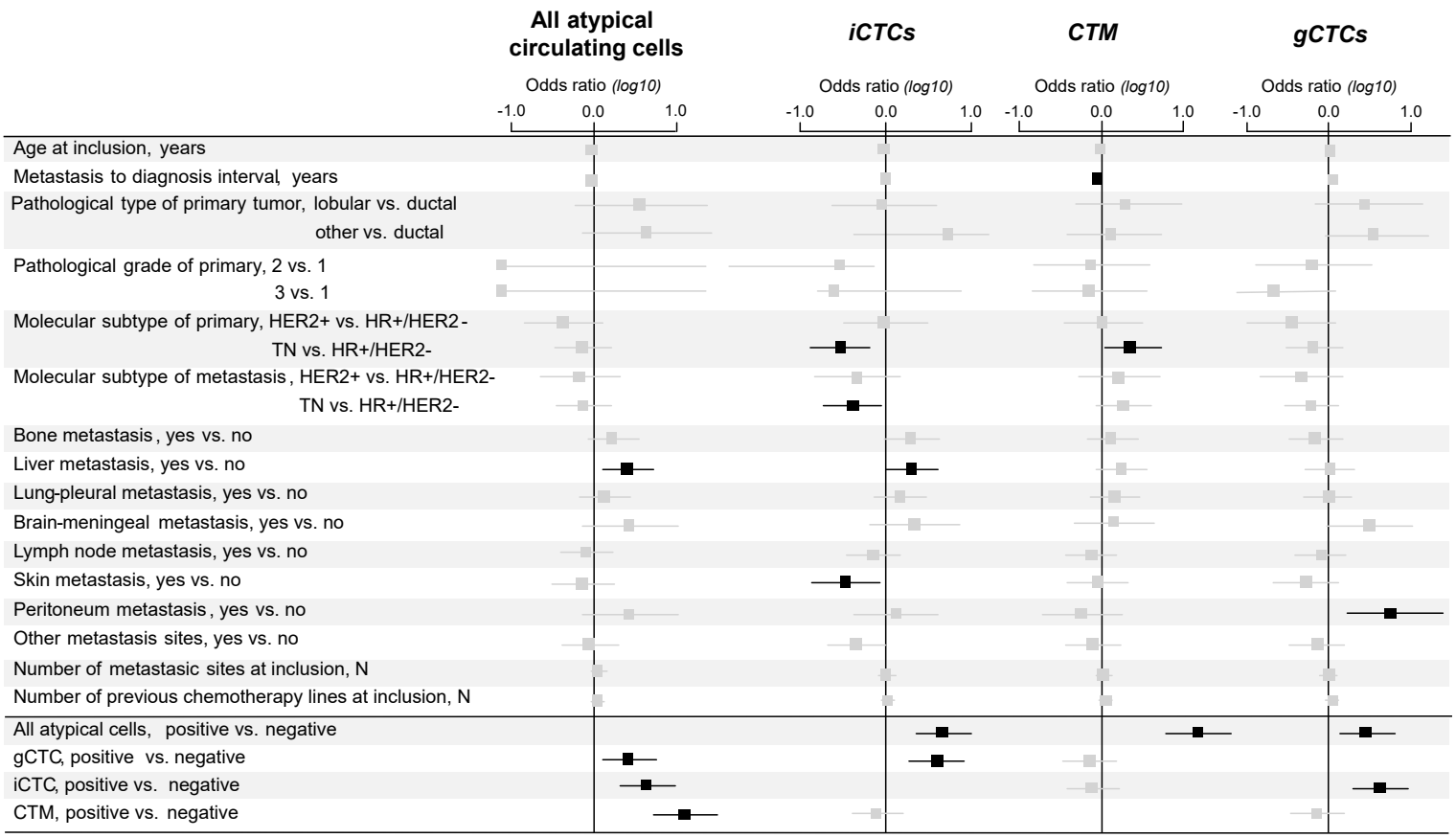
| Subsets | Total nb of cells analyzed | Nucleus size, mean (SD) | Percentage of cells positive for: | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | Leukocyte | EMT status | | | | Stem and drug resistance markers | |
| | | | CD45 (leukocyte marker) (%) | Epithelial markers (%) | Epithelial + mesenchymal markers (%) | Mesenchymal markers (%) | None of the EMT markers (%) | Stem cell markers (%) | Drug Resistance markers (%) |
| ***iCTC*** | 336 | 20.8 (+/-3.8) | 0 | 2.6 | 27.5 | 69.9 | 0 | 52.1 | 49 |
| ***CTM*** | 1742 (164 cl) | 16.3 (+/-3.6) | 0 | 0.6 | 82.5 | 16.4 | 0.5 | 89.4 | 69.4 |
| ***gCTC*** | 74 | 29.8 (+/-8.3) | 35 (diffuse) | 0 | 63.6 | 27.4 | 9 | 98.9 | 89.3 |

## B

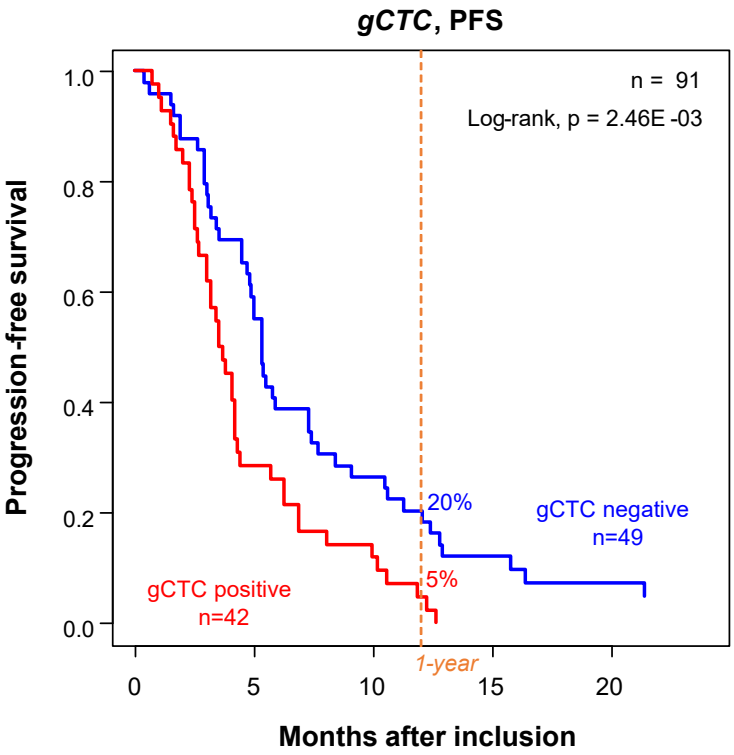


# Figure 4

**A.**


All atypical circulating cells
iCTCs
CTM
gCTCs
Odds ratio (log10)
-1.0
0.0
1.0
Age at inclusion, years
Metastasis to diagnosis interval, years
Pathological type of primary tumor, lobular vs. ductal
other vs. ductal
Pathological grade of primary, 2 vs. 1
3 vs. 1
Molecular subtype of primary, HER2+ vs. HR+/HER2-
TN vs. HR+/HER2-
Molecular subtype of metastasis, HER2+ vs. HR+/HER2-
TN vs. HR+/HER2-
Bone metastasis, yes vs. no
Liver metastasis, yes vs. no
Lung-pleural metastasis, yes vs. no
Brain-meningeal metastasis, yes vs. no
Lymph node metastasis, yes vs. no
Skin metastasis, yes vs. no
Peritoneum metastasis, yes vs. no
Other metastasis sites, yes vs. no
Number of metastasic sites at inclusion, N
Number of previous chemotherapy lines at inclusion, N
All atypical cells, positive vs. negative
gCTC, positive vs. negative
iCTC, positive vs. negative
CTM, positive vs. negative

**B.**


gCTC, PFS
n = 91
Log-rank, p = 2.46E -03
Progression-free survival
1.0
0.8
0.6
0.4
0.2
0.0
20%
5%
gCTC negative
n=49
gCTC positive
n=42
1-year
0
5
10
15
20
Months after inclusion

**C.**

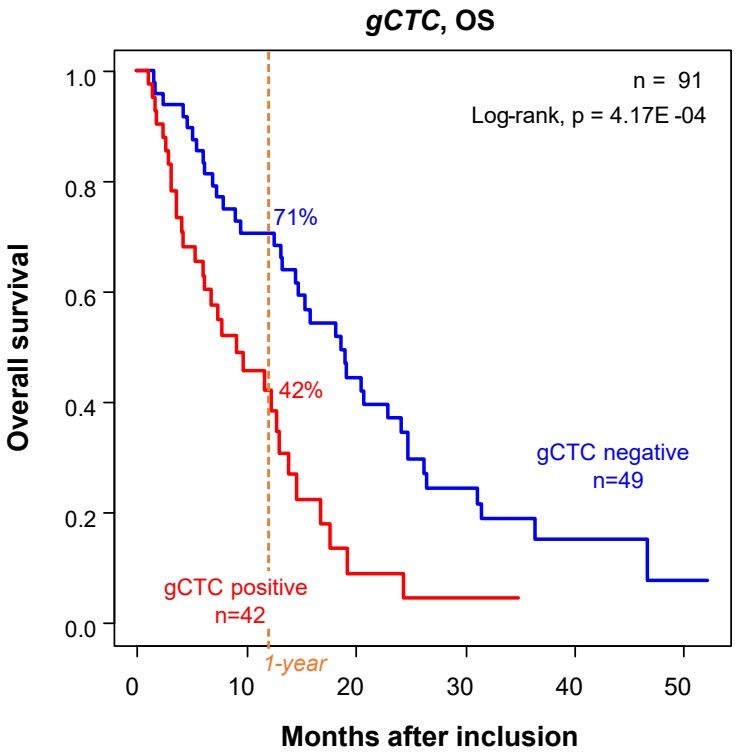

gCTC, OS
n = 91
Log-rank, p = 4.17E -04
Overall survival
1.0
0.8
0.6
0.4
0.2
0.0
71%
42%
gCTC negative
n=49
gCTC positive
n=42
1-year
0
10
20
30
40
50
Months after inclusion

**Table 2: Univariate and multivariate analysis for PFS and OS**

| | | PFS | | | | | | OS | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Univariate | | | Multivariate | | | Univariate | | | Multivariate | | |
| | | N | HR [95%CI] | p-value | N | HR [95%CI] | p-value | N | HR [95%CI] | p-value | N | HR [95%CI] | p-value |
| Age at inclusion, years | | 91 | 1.02 [1.00-1.04] | **3.34E-02** | 81 | 1.02 [1.00-1.04] | 0.055 | 91 | 1.01 [0.99-1.04] | 0.286 | | | |
| Metastasis to diagnosis interval, years | | 91 | 1.02 [0.98-1.06] | 0.34 | | | | 91 | 1.03 [0.99-1.08] | 0.173 | | | |
| Pathological type of primary tumor | lobular vs. ductal | 81 | 1.34 [0.57-3.12] | **1.76E-02** | 81 | 1.04 [0.43-2.51] | 0.934 | 81 | 1.75 [0.62-4.92] | **1.90E-02** | 81 | 1.65 [0.58-4.70] | 0.351 |
| | other vs. ductal | | 3.29 [1.44-7.52] | | 81 | 2.24 [0.94-5.34] | 0.069 | | 3.05 [1.36-6.84] | | 81 | 3.3 [1.43-7.59] | **5.07E-03** |
| Pathological grade of primary | 2 vs. 1 | 74 | 0.74 [0.28-1.97] | 0.453 | | | | 74 | 0.64 [0.19-2.18] | 0.533 | | | |
| | 3 vs. 1 | | 0.59 [0.23-1.55] | | | | | | 0.53 [0.16-1.80] | | | | |
| Molecular subtype of primary | HER2pos | 90 | 1.47 [0.74-2.91] | 0.151 | | | | 90 | 1.50 [0.67-3.39] | 0.262 | | | |
| | TN | | 1.99 [0.96-4.12] | | | | | | 1.98 [0.85-4.64] | | | | |
| Molecular subtype of metastasis | HER2pos | 87 | 0.87 [0.43-1.76] | 0.866 | | | | 87 | 1.26 [0.58-2.73] | 0.694 | | | |
| | TN | | 1.05 [0.67-1.67] | | | | | | 1.25 [0.73-2.14] | | | | |
| Bone metastasis | yes vs. no | 91 | 0.89 [0.57-1.38] | 0.591 | | | | 91 | 1.04 [0.62-1.74] | 0.892 | | | |
| Liver metastasis | yes vs. no | 91 | 0.81 [0.52-1.25] | 0.335 | | | | 91 | 1.01 [0.62-1.66] | 0.958 | | | |
| Lung-pleural metastasis | yes vs. no | 91 | 1.09 [0.72-1.66] | 0.679 | | | | 91 | 1.06 [0.65-1.71] | 0.817 | | | |
| Brain-meningeal metastasis | yes vs. no | 91 | 1.42 [0.73-2.75] | 0.304 | | | | 91 | 3.15 [1.58-6.28] | **1.15E-03** | 81 | 2.9 [1.40-6.00] | **4.04E-03** |
| Lymph node metastasis | yes vs. no | 91 | 1.06 [0.68-1.65] | 0.784 | | | | 91 | 1.14 [0.68-1.91] | 0.607 | | | |
| Skin metastasis | yes vs. no | 91 | 1.02 [0.60-1.74] | 0.942 | | | | 91 | 1.43 [0.81-2.51] | 0.219 | | | |
| Peritoneum metastasis | yes vs. no | 91 | 1.20 [0.62-2.32] | 0.597 | | | | 91 | 1.50 [0.71-3.16] | 0.287 | | | |
| Other metastatic site | yes vs. no | 91 | 0.95 [0.59-1.53] | 0.845 | | | | 91 | 1.34 [0.80-2.23] | 0.269 | | | |
| Number of metastasic sites at inclusion, N | | 91 | 1.00 [0.86-1.15] | 0.948 | | | | 91 | 1.12 [0.96-1.30] | 0.158 | | | |
| Number of previous chemotherapy lines at inclusion, N | | 91 | 1.11 [1.00-1.24] | **4.20E-02** | 81 | 1.08 [0.97-1.20] | 0.158 | 91 | 1.18 [1.04-1.33] | **8.50E-03** | 81 | 1.17 [1.04-1.32] | **1.10E-02** |
| all atypical circulating cells | positive vs. negative | 91 | 1.37 [0.89-2.10] | 0.158 | | | | 91 | 1.50 [0.91-2.49] | 0.115 | | | |
| iCTC | positive vs. negative | 91 | 1.51 [0.99-2.29] | 5.60E-02 | | | | 91 | 1.51 [0.93-2.45] | 9.30E-02 | | | |
| CTM | positive vs. negative | 91 | 1.16 [0.77-1.76] | 0.478 | | | | 91 | 1.35 [0.83-2.20] | 0.221 | | | |
| gCTC | positive vs. negative | 91 | 1.94 [1.25-3.01] | **2.98E-03** | 81 | 2.03 [1.24-3.33] | **5.09E-03** | 91 | 2.46 [1.47-4.12] | **5.84E-04** | 81 | 2.28 [1.25-4.18] | **7.26E-03** |