

## ARTICLE OPEN



# Tumor-agnostic ctDNA levels by mFAST-SeqS in first-line HR-positive, HER2 negative metastatic breast cancer patients as a biomarker for survival

Noortje Verschoor<sup>1</sup>✉, Vanja de Weerd<sup>1</sup>, Mai N. Van<sup>1</sup>, Jaco Kraan<sup>1</sup>, Marcel Smid<sup>1</sup>, Joan B. Heijns<sup>2</sup>, Jan C. Drooger<sup>3</sup>, Johanna M. Zuetenhorst<sup>4</sup>, Annemieke van der Padt-Pruijsten<sup>5</sup>, Agnes Jager<sup>1</sup>, Stefan Sleijfer<sup>1</sup>, John W. M. Martens<sup>1</sup> and Saskia M. Wilting<sup>1</sup>

This prospective cohort study reports aneuploidy score by mFast-SeqS as a strong prognostic marker in MBC patients. mFAST-SeqS is an affordable and easily implementable method for the assessment of total ctDNA levels and, as such, provides an alternative prognostic tool. One mixed cohort (cohort A,  $n = 45$ ) starting any type of treatment in any line of therapy and one larger cohort (cohort B,  $n = 129$ ) consisting of patients starting aromatase inhibitors (AI) as first-line therapy were used. mFAST-SeqS was performed using plasma of blood in which CTCs (CellSearch) were enumerated. The resulting aneuploidy score was correlated with categorized CTC count and associated with outcome. The aneuploidy score was significantly correlated with CTC count, but discordance was observed in 31.6% when applying cut-offs of 5. In both cohorts, aneuploidy score was a significant prognostic marker for both PFS and OS. In the Cox regression models, the HR for aneuploidy score for PFS was 2.52 (95% CI: 1.56–4.07), and the HR for OS was 2.37 (95% CI: 1.36–4.14). Results presented here warrant further investigations into the clinical utility of this marker in MBC patients.

*npj Breast Cancer* (2023)9:61; <https://doi.org/10.1038/s41523-023-00563-w>

## INTRODUCTION

Circulating tumor cells (CTCs) have been extensively researched as noninvasive markers of tumor load in metastatic breast cancer<sup>1</sup>. CTCs can be quantified by the frequently used CellSearch method (Menarini Silicon Biosystems Inc.). The prevalence of CTCs is approximately 60% in metastatic breast cancer (MBC) patients. In about 50% of patients, CTCs are above the validated cut-off value of 5 in 7.5 ml blood, thereby indicating aggressive disease<sup>1–3</sup>. For breast cancer, there is the level I evidence of CTC count above this cut-off value for being prognostic for progression-free survival (PFS) and overall survival (OS), independent of receptor subgroup and therapy line<sup>4,5</sup>. However, the clinical utility of CTC enumeration has not been proven yet to this date<sup>6,7</sup>. The enumeration of CTCs requires specialized equipment and trained staff, which could hamper clinical implementation.

Alternatively, circulating tumor DNA (ctDNA) can be used to estimate tumor load in the blood. Previous work showed the prognostic value of the levels of ctDNA in MBC patients, but a high level of evidence has not yet been achieved<sup>8</sup>. ctDNA level is often expressed as the variant allele frequency (VAF) of the dominant tumor-derived mutation, determined by next-generation sequencing (NGS) panels. Such panels report on mutations and copy number variations in a number of known cancer-associated genes and thus provide valuable, yet restricted information about the interrogated tumor genome. In contrast to tumor-informed approaches like targeted NGS, approaches such as whole genome sequencing or whole exome sequencing are used to obtain more extensive information on tumor load, as well as on possible targets for treatment and predictive biomarkers. However, most

assays are rather expensive and need considerable bioinformatical support to interpret their results<sup>9,10</sup>. Furthermore, applying these methods requires a relatively high cfDNA input, mostly found only in patients with late-stage disease<sup>11</sup>. Moreover, most of the above-mentioned methods have a long turn-around time, prohibiting their use as biomarkers of tumor load in daily clinical practice.

In contrast, the modified fast aneuploidy screening test-sequencing system (mFAST-SeqS) represents a fast and fairly affordable assay to estimate the fraction of ctDNA by assessing chromosomal aneuploidy instead of mutations. Since in the majority of malignancies, aneuploidy is a hallmark, this method is readily applicable in multiple tumor types and is, therefore, tumor-agnostic<sup>12</sup>. Because it is an amplicon-based NGS method, it only requires standard equipment present at virtually any molecular and/or diagnostics laboratory. The method was initially designed to pre-screen plasma samples for a high ctDNA content to aid further analysis with methods that require high input, for example, complete or shallow whole-genome sequencing. In that setting, it has been previously shown that an aneuploidy score of five or higher derived by mFAST-SeqS had a sensitivity of 100% and a specificity of 80% for the presence of copy number aberrations revealed by shallow whole-genome sequencing<sup>13</sup>. Moreover, aneuploidy score of five or higher is strongly associated with a mutation allele frequency of 5–10% or higher, assessed by targeted sequencing<sup>14</sup>.

In a proof-of-concept study by Suppan et al., it has subsequently been shown that the aneuploidy score by itself had prognostic value in MBC<sup>15</sup>. Thus, this blood-based aneuploidy score might be a relevant clinical biomarker. Notwithstanding the promising results, this pilot study was performed in a small group of 29 MBC

<sup>1</sup>Department of Medical Oncology, Erasmus MC Cancer Institute, Rotterdam, The Netherlands. <sup>2</sup>Department of Medical Oncology, Amphia, Breda, The Netherlands. <sup>3</sup>Department of Medical Oncology, Breast Cancer Center South Holland South, Ikazia Hospital, Rotterdam, The Netherlands. <sup>4</sup>Department of Medical Oncology, Franciscus Gasthuis & Vlietland, Rotterdam/ Schiedam, the Netherlands. <sup>5</sup>Department of Internal Medicine, Breast Cancer Center South Holland, Maasstad Hospital, Rotterdam, The Netherlands.

✉email: n.verschoor@erasmusmc.nl

patients of different receptor subtypes and in different lines of treatment. The prognostic value for aneuploidy score is yet to be validated for MBC patients. Additionally, it is currently not yet fully elucidated to what extent CTC count and levels of ctDNA correlate.

In the present work, we investigate the correlation between CTC count and the ctDNA-load, estimated by the aneuploidy score derived by mFAST-SeqS. Secondly, we aim to assess the clinical validity of the aneuploidy score by analyzing the prognostic value for survival in two separate MBC cohorts.

## RESULTS

### Description of samples

A total of 92 and 157 patients were included in cohorts A and B, respectively. From cohort A, blood draws that were not taken at the start of a new line of therapy ( $n = 31$ ), and non-unique patient samples ( $n = 14$ ) were excluded. From cohort B, ctDNA isolation could not be performed in twelve samples because there was no sufficient amount of plasma available. As a result, correlation analysis between aneuploidy and CTC count was conducted in 45 and 145 patients, respectively. For assessment of prognostic value, patients from cohort B that did not start a first-line aromatase inhibitor ( $n = 10$ ) and patients that started an aromatase inhibitor because of poor clinical performance ( $n = 6$ ) were excluded, upon which 129 patients remained (flowchart in supplementary figure 1). The baseline characteristics of both cohorts were different, mainly due to the inclusion criteria (Table 1). The median follow-up in cohort A was 11 months (95% CI: 1–37 months), in which 40 patients experienced progressive disease and 30 patients died. In cohort B, the median follow-up time was 27 months (95% CI: 2–73 months), in which 81 patients had progressive disease, and 58 patients died. The most important difference between the two cohorts was that patients in cohort A were unselected for the receptor subtype. Furthermore, 37 patients (82.2%) were treated with one or multiple lines of therapy before starting a new line of therapy, as opposed to cohort B, which consisted only of patients that were sampled before the start of first-line aromatase inhibitors. The more advanced stage of disease of patients in cohort A is reflected in a greater proportion of patients with a high CTC count ( $p = 0.02$ , chi-square) and with a high aneuploidy score, although this was not significantly different ( $p = 0.13$ , chi-square).

### Comparison between biomarkers

In all samples combined ( $n = 190$ ), there was a significant correlation between aneuploidy score and CTC count (Spearman's  $\rho = 0.49$ ,  $p < 0.01$ ). Additionally, there was a significant trend for a higher median aneuploidy score with increasing CTC count (Jonkcheere–Terpstra test  $p < 0.001$ , Fig. 1). Seventy-three patients (41.9%) had low scores on aneuploidy score and CTC count (both  $< 5$ ), whereas 46 patients (26.4%) had high scores ( $\geq 5$ ) on both markers (correlation plots shown in supplementary figure 2). There was a significant association between a high aneuploidy score and a high CTC count (chi-square test,  $p < 0.001$ ). However, discordancy was also observed in 55 patients, of which 34 (19.5%) only had a high CTC count, and 21 (12.1%) only had a high aneuploidy score. The aneuploidy score was not different between patients with and without the visceral disease (Mann–Whitney test,  $p = 0.54$ ), nor between patients with liver metastases and with metastases at other sites (Mann–Whitney test,  $p = 0.08$ ). Additionally, the baseline CA15.3 measurement was available for 128 patients, of whom 74% showed levels above the clinically used cut-off of 30. There was no correlation between CA15.3 and aneuploidy score (Spearman Rho 0.06,  $p = 0.50$ ). When considering samples with an aneuploidy score of 5 or higher, alterations in chromosomes 1q and 8q (copy number gain) as well as chromosomes 8p, 10q, 13q, and 17p (copy number loss, Supplementary Fig. 3) were most frequently observed. Lastly,

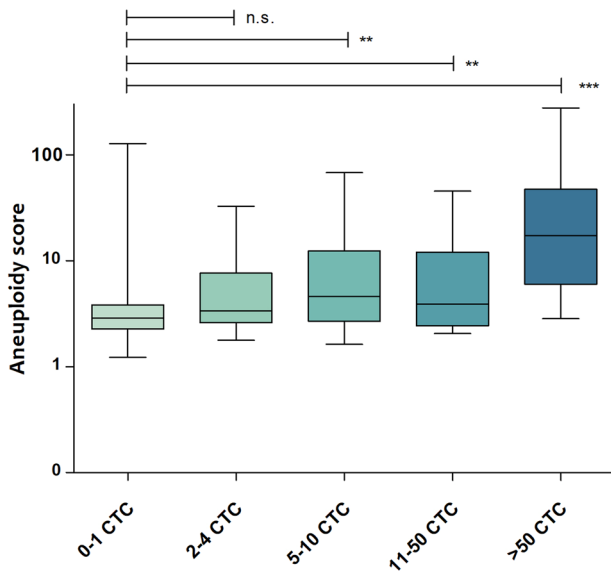
**Table 1.** Baseline characteristics of cohort A ( $n = 45$ ) and cohort B ( $n = 129$ ).

Characteristic	Cohort A ( $n = 45$ )	Cohort B ( $n = 129$ )
<b>Age</b>		
Mean (range) (yr)	56 (33–79)	68 (40–89)
<b>WHO PS at registration—no. (%)</b>		
0	18 (40.0)	41 (31.8)
1	17 (37.8)	74 (57.4)
2–3	5 (11.1)	14 (10.9)
Unknown	5 (11.1)	0
<b>Subtype (%)</b>		
HR+	30 (66.7)	129 (100)
HER2+	9 (20.0)	0
TNBC	6 (13.3)	0
<b>Disease-free interval (months)</b>		
Mean $\pm$ SD	49 $\pm$ 61	182 $\pm$ 220
<b>Visceral disease—no. (%)</b>		
	31 (68.9)	63 (48.8)
<b>Previous treatment lines (%)</b>		
0	8 (17.8)	129 (100)
1–2	20 (44.4)	0
>2	17 (37.8)	0
<b>Treatment after blood draw</b>		
Endocrine therapy	15 (33.3)	129 (100)
Addition of CDK4/6i	6 (13.3)	36 (27.9)
Chemotherapy	25 (55.6)	0
Targeted therapy	5 (11.1)	0
<b>Baseline CA15.3</b>		
Median (IQR)	69 (232)	77 (279)
<30 (%)	12 (36.4)	21 (22.1)
$\geq 30$ (%)	21 (63.6)	74 (77.9)
Missing	12	34
<b>CTC count</b>		
Median (IQR)	9 (187)	3 (20)
<5 (%)	18 (40.0)	76 (58.9)
$\geq 5$ (%)	27 (60.0)	53 (41.1)
<b>Aneuploidy score</b>		
Median (IQR)	4.2 (24.3)	3.6 (5.1)
<5 (%)	24 (53.3)	83 (64.3)
$\geq 5$ (%)	21 (46.7)	46 (35.7)
<b>Survival status</b>		
Alive	15 (33.3)	71 (55.0)
Breast cancer-related death	29 (64.4)	42 (32.6)
Not breast cancer-related death	0	5 (3.9)
Unknown	1 (2.2)	11 (8.5)

supplementary figure 4 demonstrates a significant correlation (Spearman's  $\rho = 0.66$ ,  $p < 0.001$ ) between VAF and aneuploidy score in patients with a detectable mutation by the OncoPrint™ Breast cfDNA panel ( $n = 24$ ).

### Association of aneuploidy score with survival measures

Next, the potential prognostic value of an elevated aneuploidy score was evaluated. The Kaplan–Meier curves in Fig. 2a–d show that a high aneuploidy score was significantly associated with PFS and with OS in cohort A as well as in cohort B (all  $p < 0.05$ , log-rank



**Fig. 1 Association between aneuploidy score and CTC count.** Boxplots showing the association between classes of circulating tumor cell counts and median aneuploidy score (Jonckheere–Terpstra test for trend  $p < 0.001$ ). The boxes indicate the 25th to 75th percentile, the middle line indicates the median, and the whiskers represent minimal and maximal values. The lines above the graph show interclass comparisons by two-sided Mann–Whitney  $U$  tests: n.s. = non significant, \*\* =  $p < 0.001$ , \*\*\* =  $p < 0.0001$ .

test). When combining both aneuploidy and CTC count in cohort B, patients that had both an elevated aneuploidy score and an elevated CTC count had the shortest PFS (6.0 months, 95% CI: 5.3–6.8 months) and OS (31.2 months, 95% CI: 21.8–40.7 months). Patients with a high aneuploidy score and a low CTC count had comparable OS to the patients with both markers elevated, whereas patients with elevated CTC count and a low aneuploidy score did not (median 25.9 and 60.9 months, respectively, log-rank test for trend  $p = 0.04$ , Fig. 2e). For PFS, this effect was not observed, although patients with both markers elevated clearly had worse outcomes on endocrine therapy (log-rank test for trend  $p < 0.0001$ , Fig. 2f). Multivariable Cox regression with PFS as outcome measure demonstrated that a high aneuploidy score was a significant prognostic marker for PFS with a hazard ratio of 2.52 (95% CI: 1.56–4.07), independent of visceral disease, while the CTC count was not significantly associated. For OS, the hazard ratio in the multivariable model for a high aneuploidy score was 2.37 (95% CI: 1.36–4.14), next to significant associations for WHO performance 2 status and the presence of visceral disease (Table 2).

## DISCUSSION

In the present work, we established the independent prognostic value of the aneuploidy score for PFS and OS in MBC patients. First, we validated the findings of Suppan et al. by showing that aneuploidy score correlated with outcome. Next, we further emphasized these findings in a larger, homogeneous cohort of patients that were treated with first-line aromatase inhibitors. Finally, multivariable analysis showed that the aneuploidy score was prognostic, independent of other common clinical determinants.

As expected, the aneuploidy score correlated with CTC count, as both represent circulating tumor-specific markers of cancer cell turnover. The copy number profiles that we found on the chromosome arm level support the tumor-specificity of our mFast-SeqS results since these are in line with what was reported before in breast cancer<sup>16</sup>. It was remarkable that CTC count did

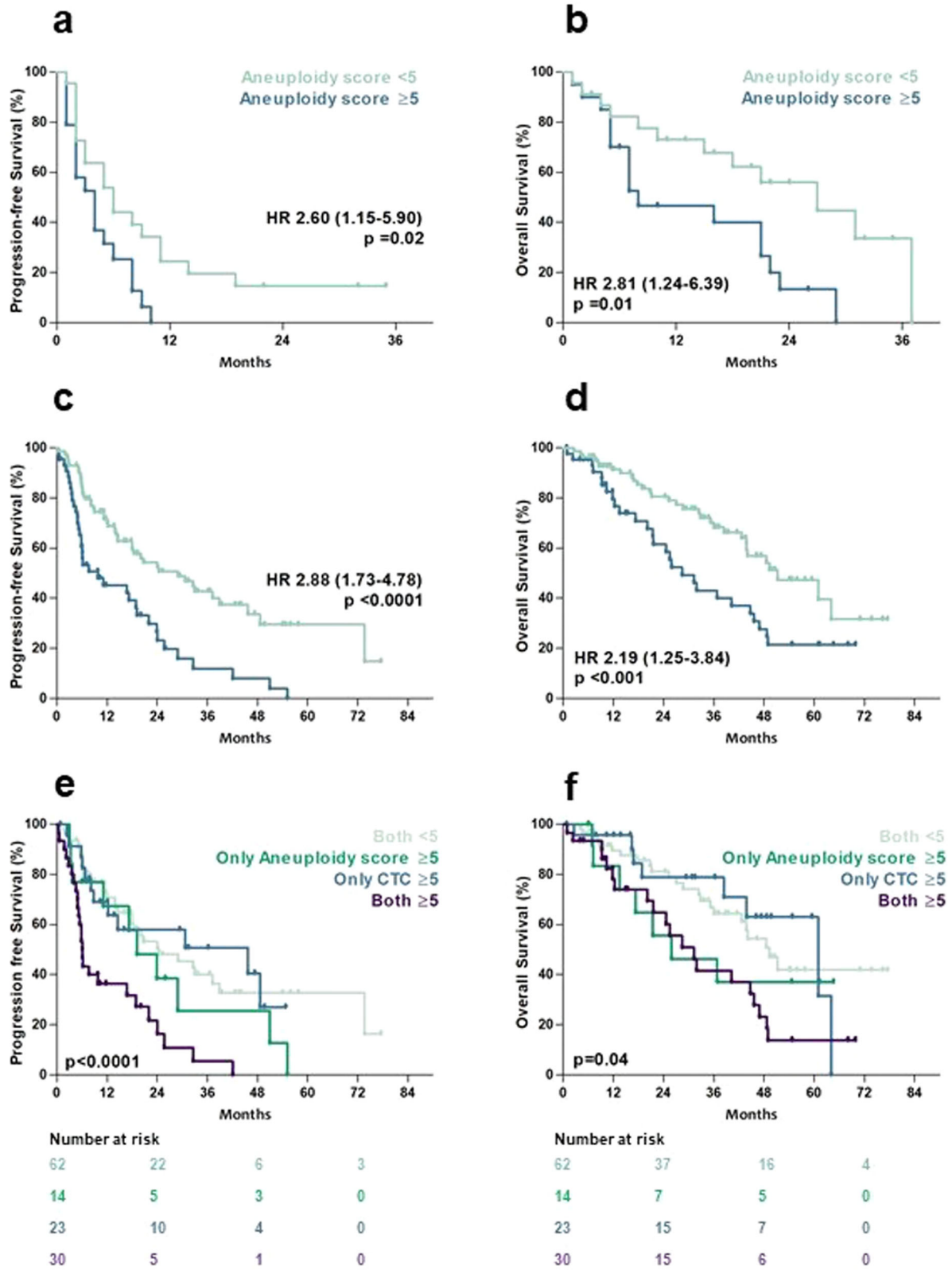
not prove to be a significant prognostic factor in our cohort, which is contradictory to previous literature<sup>4</sup>. In the analysis by Bidard et al., only 14% of patients started with endocrine treatment in different lines of treatment. Patients starting with aromatase inhibitors as first-line treatment have a relatively indolent disease, as also demonstrated by the lower median CTC count in our cohort. Thus, this specific patient group might have been underrepresented in the previously presented pooled analysis, and a different cut-off might be more suitable.

Important limitations of the mFast-SeqS include the limited sensitivity compared to the assessment of mutations and the low resolution. Furthermore, since this method is aimed at detecting aneuploidy, it cannot be used to detect targetable alterations, for example, mutations in PIK3CA. It is already known from the literature that measuring copy number alterations in circulating DNA harbors a higher detection limit than mutations or structural rearrangements because the analysis is based on increased or decreased numbers of reads of a specific segment, whereas the DNA sequence itself has not changed from healthy DNA<sup>14,17</sup>. In addition, due to the LINE-1-based amplicons, specific (focal) copy number alterations cannot always be measured. For example, it is not possible to assess HER2 amplification status by this method due to the fact that there are no LINE-1 elements on the HER2 amplicon. However, one might hypothesize that for establishing a prognosis in metastatic disease, a very sensitive assay is not always required. The benefits of the aneuploidy score, being the easy implementation and interpretation, the low costs, and its usefulness in multiple tumor types, might outweigh these disadvantages in this context.

A limitation of the current analysis is that the cut-off of five that we used for the aneuploidy score was not established for association with outcome, although this threshold was used in other cohort studies for this purpose as well<sup>18,19</sup>. Our aim was to validate the prognostic value that was previously reported for MBC patients. However, it requires a larger, independent cohort to certify the most appropriate threshold. Further analysis with more patient data will help with finding the optimal, reliable cut-off before assessing the possible clinical utility of the aneuploidy score.

The major strength of our analysis includes the two independent cohorts that we used to validate the previous findings, one of which consisted of a well-annotated cohort of patients all about to start the same first-line endocrine therapy. We showed that baseline CA15.3, which is currently used as a biomarker in the clinic, does not correlate to the aneuploidy score nor to the outcome. However, this marker is used in a more longitudinal fashion in clinical practice, and we did not collect serial samples. That the aneuploidy score could serve as an early response marker of treatment was shown by Suppan et al., but our cohorts are not suited to validate this finding.

Collectively, our study demonstrates that mFAST-SeqS is a promising method to assess tumor load without the need for prior knowledge of the mutational landscape. It is widely applicable because it requires no specialist equipment or computationally intensive bioinformatical pipelines. Furthermore, it is affordable. Lastly, since the DNA of virtually all solid tumors harbors aneuploidy to some extent and because low input of only 1 ng cfDNA is required, this method is applicable in multiple metastatic settings. Further research should first determine the most appropriate cut-off and the added value of serial sampling, and next, whether there is clinical utility in treatment stratification of MBC patients by aneuploidy score. In an era with increasing treatment possibilities and increasing health care costs, the applicability of the aneuploidy score might lie in identifying those patients in need of escalating therapy.



**Fig. 2 Kaplan–Meier estimates of probabilities of survival.** Progression-free survival (panel a) and overall survival (panel b) in cohort A and of progression-free survival (panel c) and overall survival (panel d) in cohort B. Panel e and f show Kaplan–Meier estimates of probabilities of progression-free survival and overall survival for aneuploidy score combined with CTC count in cohort B, with an associated number at risk tables below. All p-values were derived from two-sided log-rank tests.



**Table 2.** A. Univariate and multivariate Cox regression analysis for PFS in cohort B. B. Univariate and multivariate Cox regression analysis for OS in cohort B.

Characteristic	Univariate		Multivariate	
	HR (95% CI)	p-Value	HR (95% CI)	p-Value
<b>A</b>				
Age	0.99 (0.97–1.01)	0.18		
<i>WHO PS (ref 0)</i>				
1	1.26 (0.78–2.05)	0.35		
2	0.75 (0.28–1.97)	0.56		
Disease free interval	1.00 (1.00–1.00)	0.68		
Visceral disease (Y/N)	1.49 (0.96–2.32)	0.07	1.75 (1.12–2.75)	0.02
CDK4/6i (Y/N)	0.71 (0.42–1.21)	0.21		
CA15.3 ≥ 30 (Y/N)	1.01 (0.53–1.91)	0.98		
Aneuploidy score ≥ 5 (Y/N)	2.45 (1.57–3.82)	<0.001	2.52 (1.56–4.07)	<0.001
CTC-count ≥ 5 (Y/N)	1.69 (1.08–2.63)	0.02	1.33 (0.83–2.11)	0.24
<b>B</b>				
Age	1.02 (0.98–1.05)	0.09	1.02 (0.99–1.05)	0.15
<i>WHO PS (ref 0)</i>				
1	1.95 (1.02–3.74)	0.04	1.54 (0.78–3.04)	0.22
2	3.95 (1.53–10.15)	0.004	3.43 (1.17–10.0)	0.03
Disease free interval	1.00 (1.00–1.00)	0.46		
Visceral disease (Y/N)	2.08 (1.23–3.53)	0.007	2.74 (1.56–4.79)	<0.001
CDK4/6i (Y/N)	0.96 (0.52–1.75)	0.88		
CA15.3 ≥ 30 (Y/N)	0.92 (0.40–2.12)	0.85		
Aneuploidy score ≥ 5 (Y/N)	2.03 (1.21–3.40)	0.007	2.37 (1.36–4.14)	0.002
CTC count ≥ 5 (Y/N)	1.41 (0.84–2.36)	0.20		

P-values indicate a significant difference in hazard ratio compared to the reference cohort.

## METHODS

### Patient inclusion

Between May 2015 and September 2021, patients with MBC were consecutively included in one of the two following prospective cohort studies, which were approved by the Medical Research Ethics Committee of the Erasmus MC Cancer Institute (Rotterdam, the Netherlands): in cohort A (Liquid Biobank, Erasmus MC ID MEC 17–238) patients in any line of treatment and with any clinical subtype were included when there was a clinical need for CTC enumeration. In cohort B (Caremore-AI, Erasmus MC ID MEC 14–588), patients with hormone-receptor, HER2 negative MBC that was about to start treatment with first-line aromatase inhibitors (AI) ± CDK4/6 inhibitors were prospectively included. Exclusion criteria were adjuvant chemotherapy within 6 months, other anticancer therapy within two weeks, hormonal antitumor treatment within 1 week prior to the start of AI, and medical conditions prohibiting adequate follow-up. All patients provided written informed consent for the data presented here.

### Blood samples and assays

Blood was drawn in CellSave preservative tubes. CTCs were enumerated in 7.5 mL peripheral blood with the CellSearch® Circulating Epithelial Cell Kit within 96 h, according to the manufacturer's instructions. The remaining blood was centrifuged at 1700×g for 10 min, and plasma was transferred to a new tube. Then a second centrifugation step was performed at 12,000×g for 10 min at 4 °C. Plasma was stored at –80 °C until further processing. For cfDNA isolation, 0.3–1.6 mL plasma was isolated using the Maxwell® (MX) RSC LV ccfDNA Plasma Custom Kit (Promega, Madison, WI, USA) for cohort A and the QIAamp Circulating Nucleic Acid Kit (Qiagen) for cohort B and quantified by the Quant-iT dsDNA High-sensitivity Assay (Invitrogen, Life Technologies) according to the manufacturer's instructions. The Qubit Fluorometer (Invitrogen) was

used as a readout. cfDNA was then stored at –20 °C. Aneuploidy score was obtained by mFAST-SeqS essentially as previously described<sup>19</sup>. Briefly, 1 ng of cfDNA was amplified using a Phusion high fidelity polymerase (New England Biolabs) by a single primary primer pair for specific amplification of LINE-1 sequences throughout the genome (2' at 98 °C, followed by 8 cycles of 10" at 98 °C, 2' at 57 °C and 2' at 72 °C). To increase the complexity of the resulting sequencing libraries, a random spacer was introduced to the forward primer as described by Fadrosch et al.<sup>20</sup> (see Supplementary Table 1 for primer sequences). By a second PCR step (2' at 98 °C, followed by 18 cycles of 10" at 98 °C, 15" at 65 °C and 15" at 72 °C), indexes and adapters were added using Illumina index primer sequences coupled to the common sequence in the LINE-1 primers (Supplementary Table 1). The resulting libraries were pooled equimolarly and sequenced on a MiSeq system generating 150 bp single-end reads to reach at least 90,000 reads per sample. Read counts per chromosome arm were normalized to the total library size, and subsequently, a Z-score per chromosome arm was calculated relative to healthy female controls. The short arms of chromosomes 13, 14, 15, 21, and 22, as well as chromosome Y, were excluded due to the insufficient presence of LINE-1 elements. Finally, the resulting Z-scores were squared and summed and compared to genome-wide squared and summed values of the controls, yielding an overall aneuploidy score, which indicates the number of standard deviations the sample deviating from the healthy controls.

### Statistical considerations

The correlation between CTC count and aneuploidy score was calculated by Spearman's correlation coefficient and the Jonckheere–Terpstra test for trends across ordered groups. Additionally, to account for the fact that the CTC count can be zero, the proportion of patients with a high aneuploidy score (≥5) was compared between groups with a low CTC count (CTC < 5) and

a high CTC count (CTC  $\geq 5$ ) with a chi-square test. In a small subgroup of patients, the VAF of the driver mutation by the Oncomine™ Breast cfDNA Assay was previously described<sup>21</sup>. The correlation between VAF and aneuploidy score was calculated by Spearman. Next, for both cohorts, survival differences were explored and visualized with the Kaplan-Meier method and compared univariate using the log-rank test for PFS and OS. Cohort A was too small to correct for known prognostic factors. Therefore, multivariable analysis was performed with Cox proportional-hazards regression in cohort B only, thereby eliminating the possible effect of hormone and HER2 receptor subtype and line of treatment. The following factors were considered: age, disease-free interval (both as continuous variables), presence of visceral metastases, the concurrent start of CDK4/6 inhibitors, baseline CA15.3, CTC-count, and aneuploidy score (all as dichotomous variables, CTC count and aneuploidy score were scored low or high according to the previously established cut-off of five). Factors with a *p*-value of  $<0.1$  at univariable analysis were included in the multivariable model using backward stepwise selection. This lenient significance level was used to prevent the false rejection of borderline significant variables. *P*-values  $<0.05$  in the multivariable model were considered significant. Survival was defined from the date of blood draw to radiological or clinical progression (for PFS) or death (for OS). If this did not occur, patients were censored at the date of last contact. All statistical analyses were performed using IBM SPSS Statistics v25 (IBM Corp., Armonk, NY, USA), and for all tests, two-sided *p*-values of  $<0.05$  were considered statistically significant. Figures were generated with GraphPad Prism 5 and Biorender. Results were reported according to the REMARK guidelines<sup>22</sup>.

### Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this article.

### DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request.

### CODE AVAILABILITY

Code for mFast-Seqs processing can be obtained from the corresponding author upon reasonable request.

Received: 17 February 2023; Accepted: 26 June 2023;

Published online: 14 July 2023

### REFERENCES

- Cristofanilli, M. et al. Circulating tumor cells, disease progression, and survival in metastatic breast cancer. *N. Engl. J. Med.* **351**, 781–791 (2004).
- Hayes, D. F. et al. Circulating tumor cells at each follow-up time point during therapy of metastatic breast cancer patients predict progression-free and overall survival. *Clin. Cancer Res.* **12**, 4218–4224 (2006).
- Pierga, J. Y. et al. High independent prognostic and predictive value of circulating tumor cells compared with serum tumor markers in a large prospective trial in first-line chemotherapy for metastatic breast cancer patients. *Ann. Oncol.* **23**, 618–624 (2012).
- Bidard, F.-C. et al. Clinical validity of circulating tumour cells in patients with metastatic breast cancer: a pooled analysis of individual patient data. *Lancet Oncol.* **15**, 406–414 (2014).
- Febbo, P. G. et al. NCCN task force report: evaluating the clinical utility of tumor markers in oncology. *J. Natl Compr. Cancer Netw.* **9**, S-1–S-32 (2011).
- Smerage, J. B. et al. Circulating tumor cells and response to chemotherapy in metastatic breast cancer: SWOG S0500. *J. Clin. Oncol.* **32**, 3483–3489 (2014).
- Bidard, F. C. et al. Efficacy of circulating tumor cell count-driven vs clinician-driven first-line therapy choice in hormone receptor-positive, ERBB2-negative metastatic breast cancer: the STIC CTC randomized clinical trial. *JAMA Oncol.* **7**, 34–41 (2021).

- Dawson, S.-J. et al. Analysis of circulating tumor DNA to monitor metastatic breast cancer. *N. Engl. J. Med.* **368**, 1199–1209 (2013).
- Murtaza, M. et al. Non-invasive analysis of acquired resistance to cancer therapy by sequencing of plasma DNA. *Nature* **497**, 108–112 (2013).
- Rebecca, J. L. et al. Detection of chromosomal alterations in the circulation of cancer patients with whole-genome sequencing. *Sci. Transl. Med.* **4**, 162ra154–162ra154 (2012).
- Bos, M. K. et al. Whole exome sequencing of cell-free DNA—a systematic review and Bayesian individual patient data meta-analysis. *Cancer Treat. Rev.* **83**, 101951 (2020).
- Bakhoun, S. F. & Cantley, L. C. The multifaceted role of chromosomal instability in cancer and its microenvironment. *Cell* **174**, 1347–1360 (2018).
- Heitzer, E. et al. Tumor-associated copy number changes in the circulation of patients with prostate cancer identified through whole-genome sequencing. *Genome Med.* **5**, 30 (2013).
- Belic, J. et al. mFast-SeqS as a monitoring and pre-screening tool for tumor-specific aneuploidy in plasma DNA. *Adv. Exp. Med. Biol.* **924**, 147–155 (2016).
- Suppan, C. et al. Untargeted assessment of tumor fractions in plasma for monitoring and prognostication from metastatic breast cancer patients undergoing systemic treatment. *Cancers (Basel)* **11**, 1171 (2019).
- Nik-Zainal, S. et al. Landscape of somatic mutations in 560 breast cancer whole-genome sequences. *Nature* **534**, 47–54 (2016).
- Adalsteinsson, V. A. et al. Scalable whole-exome sequencing of cell-free DNA reveals high concordance with metastatic tumors. *Nat. Commun.* **8**, 1324 (2017).
- Angus, L. et al. Detection of aneuploidy in cerebrospinal fluid from patients with breast cancer can improve diagnosis of leptomeningeal metastases. *Clin. Cancer Res.* **27**, 2798–2806 (2021).
- Mendelaar, P. A. J. et al. Genome-wide aneuploidy detected by mFast-SeqS in circulating cell-free DNA is associated with poor response to pembrolizumab in patients with advanced urothelial cancer. *Mol. Oncol.* **16**, 2086–2097 (2022).
- Fadrosh, D. W. et al. An improved dual-indexing approach for multiplexed 16S rRNA gene sequencing on the Illumina MiSeq platform. *Microbiome* **2**, 6 (2014).
- Kruger, D. T. et al. High ctDNA molecule numbers relate with poor outcome in advanced ER+, HER2- postmenopausal breast cancer patients treated with everolimus and exemestane. *Mol. Oncol.* **14**, 490–503 (2020).
- Sauerbrei, W., Taube, S. E., McShane, L. M., Cavenagh, M. M. & Altman, D. G. Reporting recommendations for tumor marker prognostic studies (REMARK): an abridged explanation and elaboration. *J. Natl Cancer Inst.* **110**, 803–811 (2018).

### ACKNOWLEDGEMENTS

We express our gratitude to all women who contributed their specimens to the Liquid Biobank and Caremore-AI trials. Furthermore, we would like to thank all our collaborators that included women in these trials. This trial was funded by EU-FP7, Project number 601760-2, and by the Dutch Cancer Society (KWF), grant number 12481. The current study was presented as a poster presentation (P2-03-14) at the San Antonio Breast Cancer Symposium, December 6–10, 2022, in San Antonio, Texas.

### AUTHOR CONTRIBUTIONS

S.M.W. and S.S. conceived the study. N.V., S.M.W., and M.S. contributed to the study design and the methodology. N.V., J.B.H., J.C.D., J.M.Z., A.v.d.P.-P., A.J., V.d.W., M.N.V., and J.K. contributed to data collection. N.V., M.S., and S.M.W. contributed to the data analysis and interpretation. N.V. contributed to the manuscript preparation. A.J., S.S., J.W.M.M., and S.M.W. contributed to the paper review. All authors read and approved the final paper.

### COMPETING INTERESTS

The authors declare no competing interests.

### ADDITIONAL INFORMATION

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41523-023-00563-w>.

**Correspondence** and requests for materials should be addressed to Noortje Verschoor.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023