Supporting information

An Automated High-Throughput Counting Method for Screening Circulating Tumor Cells in Peripheral Blood

Mengxia Zhao[§^], Perry G. Schiro[§^], Jason S. Kuo[^], Karen M. Koehler[+], Daniel E. Sabath[+], Viorica Popov[‡], Qinghua Feng[‡], Daniel T. Chiu*[^]

[^] *Department of Chemistry, University of Washington, Seattle, WA, 98195.*

[+] *Departments of Laboratory Medicine and Medicine, University of Washington, Seattle, WA, 98195.*

‡ *Department of Pathology, University of Washington, Seattle, WA, 98195*

[§] *These authors contributed equally to this project*

[*] *To whom correspondence should be addressed.*

Daniel T. Chiu

Fax: (206) 685-8665

E-mail: chiu@chem.washington.edu

Table of Contents

Statistical consideration of the 3 different thresholds

Simple threshold

The mean background level of our method is 1.2 counts/mL, ranging from 0 to 4 counts/mL. Similar to the other published works[39], we may set up a threshold at 5 counts/mL as a cut off value to distinguish the CTC events from the background. After the normalization to 7.5 mL, this threshold would be 38 counts/7.5 mL.

***Threshold set using mean plus two standard deviations*.**

The mean background level of our method is 1.2 counts/mL with a standard deviation of 1.6 counts/mL. As a result, the corresponding threshold would be 4.4 counts/mL. For 7.5 mL of the blood, the threshold would be 33 counts/7.5 mL

Threshold set using Z-test.

The false positive data (Figure 3a) has an average value of 0.6 cells per sample (0.5 mL), and a standard deviation of 0.8 cells per sample (0.5 mL). Because there was only one measurement per clinical sample, we assume Poisson distributed error for the clinical samples. For the Z test,

$$Z = \frac{S - B}{\sqrt{\sigma_S^2 + \sigma_B^2}}$$

Where S is the counts in the clinical sample and B = 0.60 is the average counts in the blank and $\sigma_B^2 = 0.8^2$ is the variance in the blank. For the clinical sample σ_S^2 can be approximated by S, so we get

$$Z = \frac{S - 0.6}{\sqrt{S + 0.64}}$$

For S = 4.2, Z = 1.63, which is a 90% confidence interval. For 7.5 mL, we have 4.2 × 15 = 63. Any sample with S > 4 will have a confidence level > 90% for a two-tail test, and > 95% for a one-tail test. Since a sample cannot have a negative number of counts, we should use the one-tail test. As a result, this threshold (63 counts/7.5mL) would be used for a one-sided test with a confidence level higher than 95%.

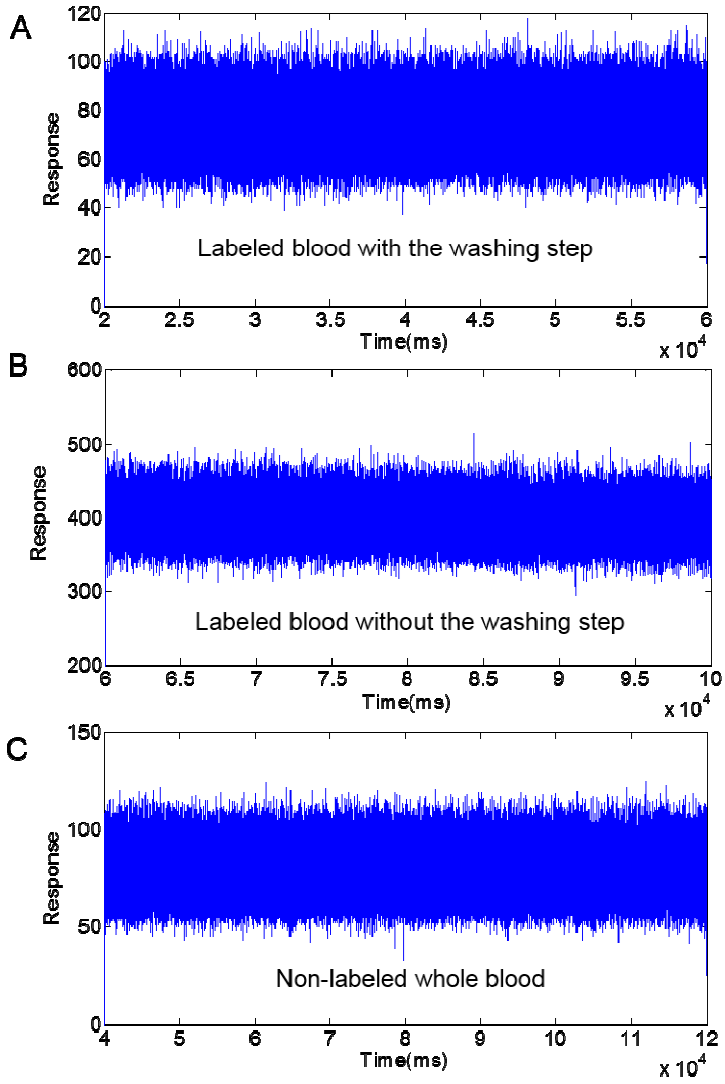
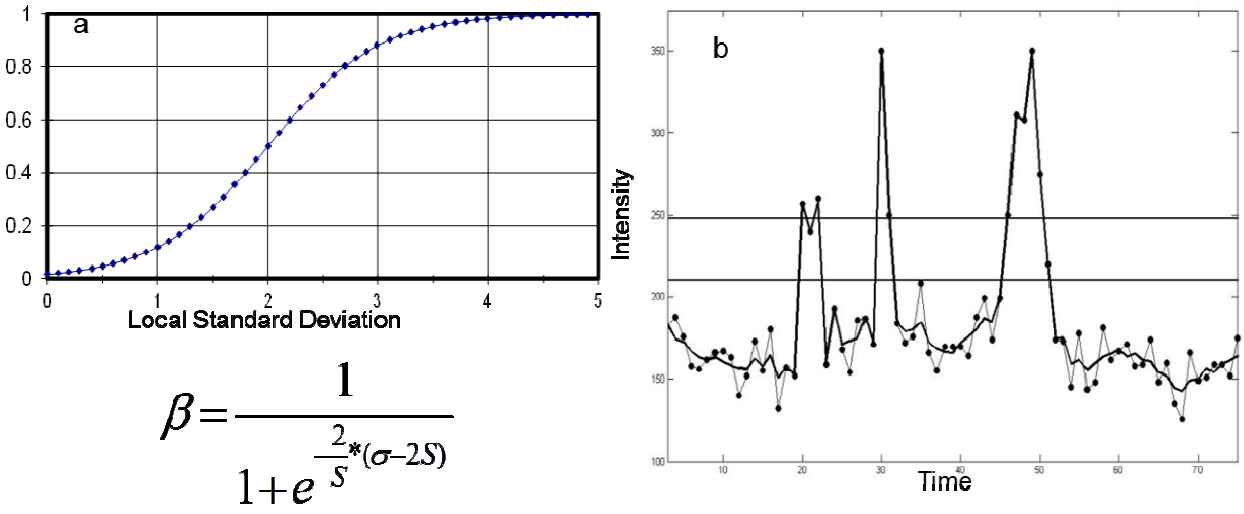


Figure 1S. Fluorescence background levels in PE (phycoerythrin) detection channel, when we flow, A) healthy blood sample labeled with anti-EpCAM-PE with the same washing step used in this paper, B) healthy blood sample labeled with anti-EpCAM-PE without any washing step, and C) healthy blood sample without any labeling and washing steps. The average level of background in figure 1SA, 1SB and 1SC is 79.4, 397.4 and 74.6, respectively.



$$\beta = \frac{1}{1 + e^{\frac{2}{S}*(\sigma - 2S)}}$$

Figure 2S . Sigmoidal burst filtering. (a) A sigmoidal function based on the deviation of the intensity for a time bin from the local median intensity. (b) An example APD trace showing the uncorrected (dots and thin line) and corrected (dark line) data. The corrected data has significantly reduced noise, allowing for a lower threshold value in identifying CTCs without changing the intensity of the identified CTCs.

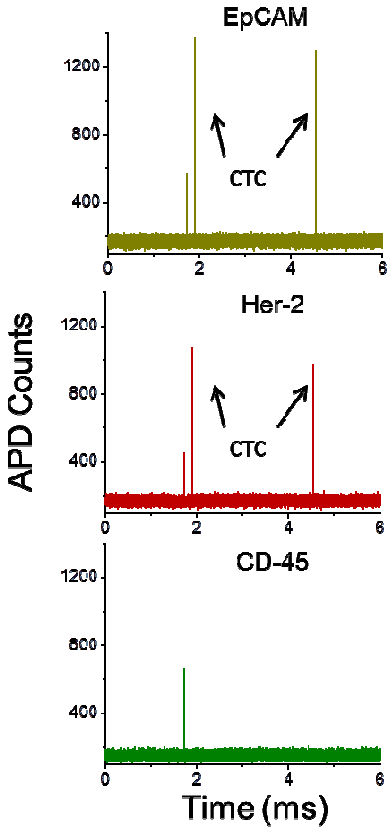


Figure 3S Example of APD data for the detection scheme using EpCAM, Her2 and CD45 as biomarkers. There were 3 detectable events in 5 ms, and the first event at 1.7 ms was not a CTC event, because it had a strong CD45 signal together with the EpCAM and Her2 responses. The last two events are defined as CTCs, because both EpCAM and Her2 expression are high without any significant CD45 signals . This experiment shows that the Her2 positive CTCs could be counted using the current labeling and detection scheme.

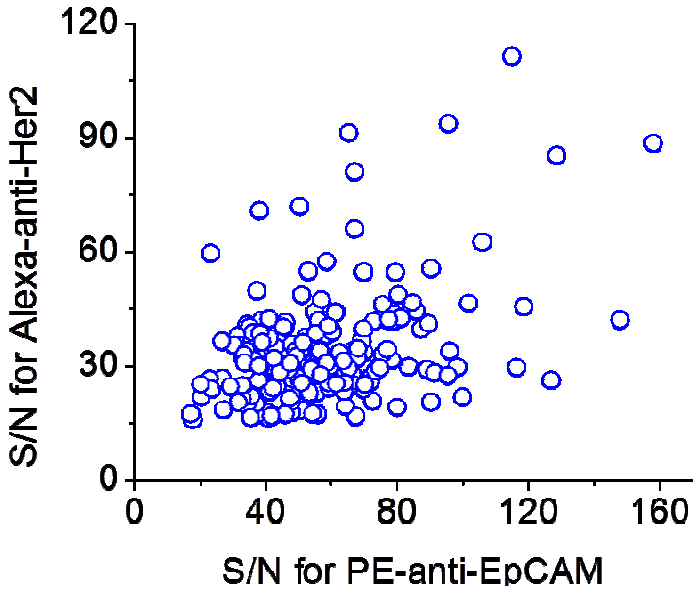


Figure 4S The distribution of signal-to-noise ratio (S/N) of a breast-cancer sample analyzed by EpCAM\Her2\CD45 method. All the data points were two-color events (EpCAM positive, Her2 positive, and CD45 negative).

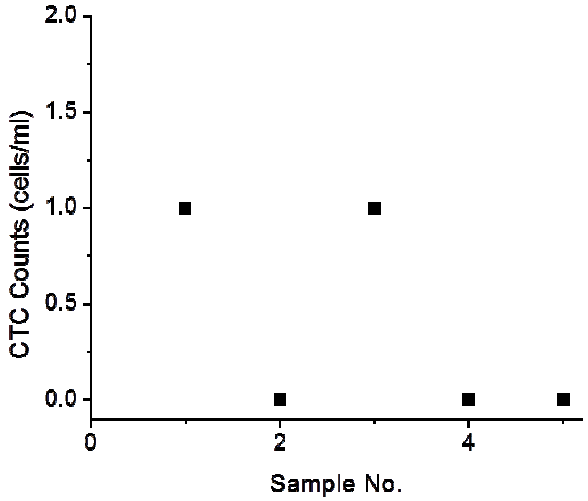


Figure 5S False positive performance for the EpCAM/Her2/CD45 scheme. On average 0.6 cells were found per mL of healthy donor blood with 60% of the samples reporting zero cells.

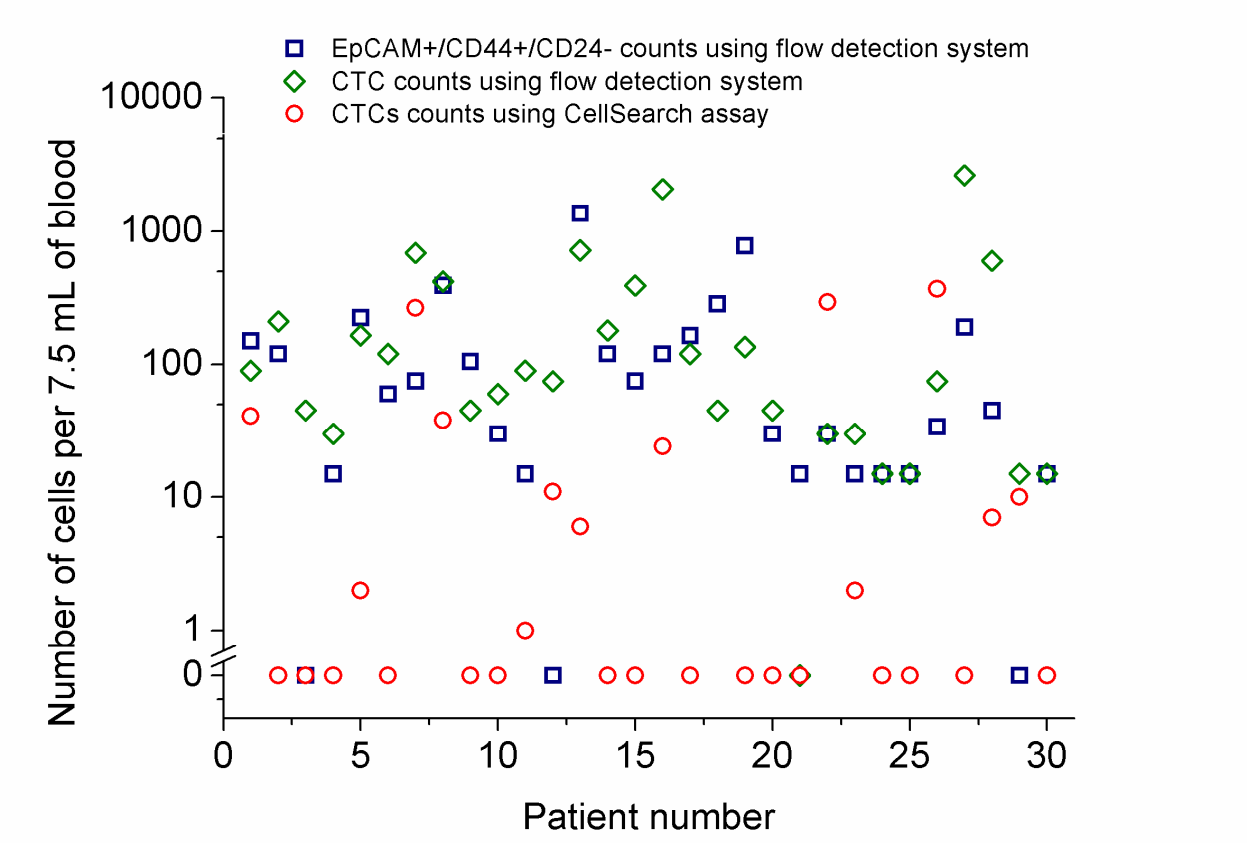


Figure 6S *Side-by-side clinical results for regular CTCs and circulating cells with EpCAM+/CD44+/CD24- expression from the CTC flow detection system and CellSearch. Normal CTCs were determined by both flow detection and CellSearch method. The average number of the EpCAM+/CD44+/CD24- cells for these 30 patient samples is 150 cells/7.5mL.*

Sample ID	CTC counts by flow detection /7.5 mL	CTC counts by CellSearch /7.5 mL	Sample ID	CTC counts by flow detection /7.5 mL	CTC counts by CellSearch /7.5 mL
1	90	41	46	0	0
2	210	0	47	120	0
3	45	0	48	60	0
4	30	0	49	180	0
5	165	2	50	45	90
6	120	0	51	30	3
7	690	265	52	15	0
8	420	38	53	15	0
9	45	3	54	105	457
10	60	0	55	255	0
11	90	0	56	20	0
12	75	1	57	30	0
13	720	11	58	15	3
14	180	6	59	15	0
15	390	0	60	30	0
16	2085	0	61	0	426
17	120	24	62	120	1
18	135	0	63	38	2
19	90	1	64	158	2
20	45	0	65	3015	0
21	0	0	66	90	0
22	30	294	67	56	1
23	30	2	68	30	0
24	15	0	69	90	0
25	75	368	70	315	5
26	2637	0	71	90	0
27	0	13	72	30	4
28	0	0	73	150	2
29	60	78	74	135	0
30	600	7	75	465	0
31	15	10	76	435	846
32	15	0	77	743	1
33	150	0	78	165	0
34	15	0	79	300	1
35	330	28	80	630	0
36	2655	0	81	90	0
37	195	1	82	89	0
38	0	0	83	285	0
39	3375	3	84	240	1
40	0	0	85	200	0

41	105	0	86	270	0
42	525	0	87	950	8
43	53	194	88	165	0
44	0	2	89	105	0
45	45	0	90	315	1

Table 1S High-throughput flow detection and CellSearch results obtained for 90 breast-cancer patient samples. The side-by-side comparison between the commercial CellSearch system and our eDAR platform using blood samples drawn from Stage IV metastatic breast cancer patients.