Supplementary Materials for

A platform for Artificial Intelligence based identification of the extravasation potential of cancer cells into the brain metastatic niche

C. Ryan Oliver[1,2†], Megan A. Altemus[1†], Trisha M. Westerhof[1], Hannah Cheriyan[1], Xu, Cheng[1], Michelle Dziubinski[1], Zhifen Wu[1], Joel Yates[1], Aki Morikawa[1], Jason Heth[3], Maria G. Castro[3,4], Brendan M. Leung[1,2††], Shuichi Takayama[2†††], Sofia D. Merajver[1]*

[1]Department of Internal Medicine, University of Michigan, Ann Arbor, MI 48109, USA.

[2] Department of Biomedical Engineering, University of Michigan, Ann Arbor, MI 48109, USA.

[3] Department of Neurosurgery, University of Michigan, Ann Arbor, MI 48109, USA.

[4] Department of Cell and Developmental Biology, University of Michigan, Ann Arbor, MI 48109, USA.

* Corresponding author: smerajve@med.umich.edu, 734.936.6884

† Equal contribution

†† Current affiliation: Department of Applied Oral Sciences and School of Biomedical Engineering, Dalhousie University, Nova Scotia, Canada
††† Current affiliation: The Wallace H Coulter Department of Biomedical Engineering, Georgia Institute of Technology and Emory School of Medicine, Atlanta GA 30332 USA.

Correspondence to: smerajve@med.umich.edu

This PDF file includes:

Materials and Methods

Mold fabrication

Using methods described by Xia et al., we designed a double chamber PDMS microfluidic device separated by a polycarbonate membrane via soft lithography. SU-8 (MicroChem 2075) master molds were fabricated to then cast PDMS microfluidic devices (41, 42). To fabricate the lower and upper chamber molds two 4 in Si wafer were cleaned in a Piranha solution to remove any organic residue. Then 4ml of SU-8 2075 was spun at 1250 rpm and 2100 rpm for 30 s, for a 200 µm and 100 µm layer thickness respectively on each wafer. This was followed by a pre-bake and softbake for 7 and 5 minutes at 65 ˚C and 35 and 12 min at 95 ˚C respectively. The mask (to define the channels) was then aligned and exposed for 47 s and 30 s with 20 mW/cm2 on an HTG 3000 HR. The mask was removed, and the wafer was put in the oven for a two-step post bake with 1 min at 65 ˚C and 6 min at 95 ˚C. The wafer was then developed by rinsing it with SU-8 developer until clean followed by an IPA rinse and drying under a nitrogen gun. After drying, PDMS (1:10) was cast to a thickness of 5 mm on the upper chamber mold and 1 mm on the lower chamber mold and then allowed to cure for 45 min in the oven at 85 ˚C. The inlets and outlets were opened using a 1.5 mm biopsy punch prior to assembly. Then a 5 mm x 10 mm polycarbonate membrane (5 µm pore size) was cut using a scalpel. The chip was assembled using a PDMS-Toluene glue (2:3). Each PDMS chamber was placed chamber down on a glass slide with the PDMS-Toluene glue spun on it (1000 rpm for 30 s). The lower chamber was lifted and the membrane was placed on top of it centered between the inlets. Finally, the second chamber was placed on top of the membrane to create the multi-level device. The PDMS device was then cured in an oven overnight at 65 ˚C. The device was then affixed onto a glass slide that had been plasma treated at 50W for 40 seconds. The completed device then had cut 200 µL pipette tips placed in each inlet/outlet and was plasma treated for 10 minutes at 100 W immediately before seeding cells.

The height of the SU-8 master mold was measured using a profilometer (KLA-Tencor P-16+). The profilometer was fitted with a 12.5 µm tip with a length of 500 µm. The SU-8 patterns on the first device measured 198 µm tall while the second measured 117 µm.

After assembly the upper channel is rectangular measuring 0.8 mm wide, 100 µm tall, and 17.5 mm long. The lower channel is 3 mm wide, 200 µm tall, and 12 mm long. The inlets and outlets of the device are 1.5 mm in diameter and formed using a biopsy punch. The membrane is made of polycarbonate and placed to separate the upper and lower channel.

Supplementary Text

Microfluidic chip design

Two masks are made from the below design. 1) The top channel 100 µm tall and 2) The bottom channel 200 µm deep. The file uBBN_Mask_Design.dwg is also available upon request.

Co-culture media composition

The co-culture media composition was found by comparing the growth rates of the endothelial cells and normal human astrocytes with different proportions of their respective medias. Figures S2 shows how the resulting cultures looked after 2 days.

Statistical summary data

Here we provide summary statistics of the cells by each measurement taken including the cell count, mean, standard deviation and p-value according to the statistical test described in the methods. Table S1 describes the % extravasated, distance extravasated and sphericity measurements for the cell lines at 24 hrs and 48 hrs. Table S2 describes the measured number of

cells and their volumes at 24 hrs for cells that extravasated < 90% and greater than 90% across the membrane. The p-values provide a comparison of the distributions of the cells against the MCF10A control. Table S3 describes the summary statistical data of the PDX samples including the number of cells, %volume extravasate of each cell, distance extravasated and sphericity at 24 hrs and 48 hrs. Table S4 describes the measured number of cells and their volumes at 24 hrs for cells that extravasated < 90% and greater than 90% across the membrane. It also provides data for 48 hrs between both TNBC samples from the primary and brain metastatic patients. The p-values provide a comparison of the distributions of the cells against the primary TNBC tumor site.

Diagnostic algorithm performance data

Table 5 illustrates an extended version of the tables shown in the manuscript with the addition of Precision and Recall. These were not included in the manuscript as they are redundant with the F1 measure but are useful to know regardless. The data shown in Table S6 was used to calculate the positive and negative predictive values discussed in the manuscript and give an illustration of the results shown in Table 1. Table 7 illustrates an extended version of the tables shown in the manuscript with the addition of Precision and Recall. These were not included in the manuscript as they are redundant with the F1 measure but are useful to know regardless. The data shown in Table S8 was used to calculate the positive and negative predictive values discussed in the manuscript and give an illustration of the results shown in Table 2.

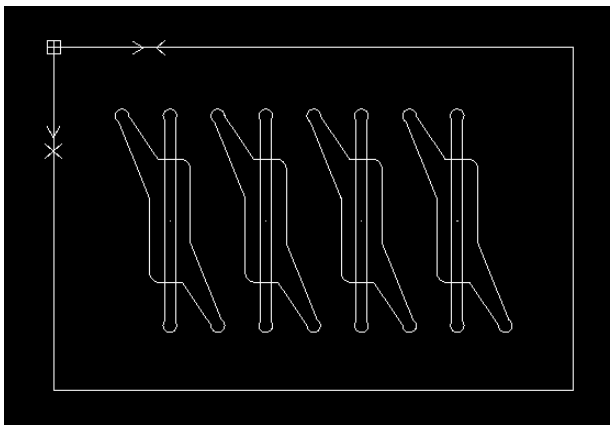

Fig. S1. Mask design of four channel chip.

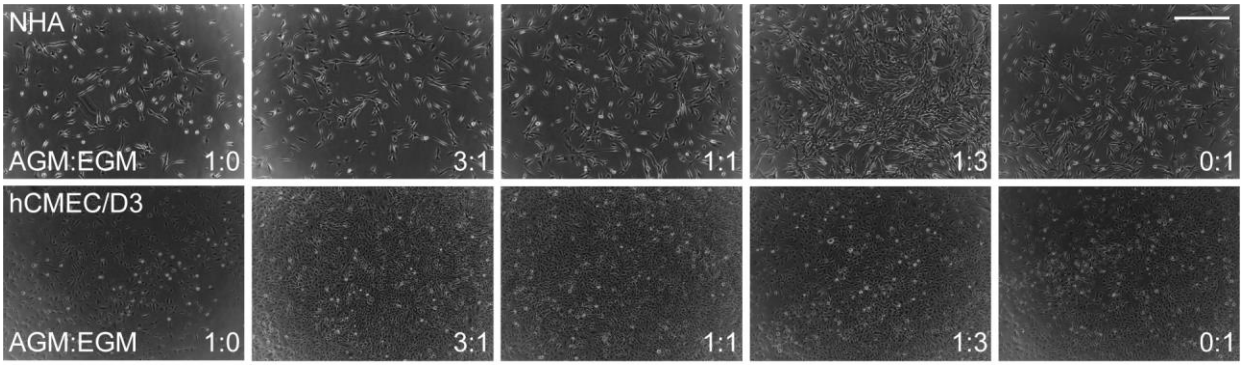

Fig. S2. Optimizing media composition for Co-culture.

Table S1. Summary of metrics measured for each cell line.

24 hrs

Cell line	Cell count	% Extravasated by volume	p-value	Distance extravasated (μm)	p-value	Sphericity	p-value
MCF10A	671	36.45±34.83	-	-14.74±34.23	-	0.45±0.14	-
MDA-MB-231	3780	43.44±30.90	8e-05	-0.77±33.87	6e-09	0.60±0.14	2e-16
MDA-MB-231-BR	6459	56.91±20.12	2e-16	6.91±14.03	8e-06	0.74±0.13	2e-16

48 hrs

Cell line	Cell count	% Extravasated by volume	p-value	Distance extravasated (μm)	p-value	Sphericity	p-value
MCF10A	2951	41.55±22.30	-	-2.18±20.62	-	0.47±0.14	-
MDA-MB-231	11125	42.25±28.47	2e-13	0.89±25.73	2e-16	0.57±0.17	0.068
MDA-MB-231-BR	5290	51.35±31.60	2e-16	18.34±6.33	2e-16	0.52±0.11	3e-06

Table S2. Cell counts and volumes by location.

Cell line	Cell count < 90%	Volume < 90% (voxels)	p-value	Cell count > 90%	Volume > 90% (voxels)	p-value
MCF10A	472	503.6±502.5	-	43	370.2±361.9	-
MDA-MB-231	3068	429.0±435.6	2e-16	387	285.9±284.5	1e-09
MDA-MB-231-BR	5986	167.7±201.8	2e-15	440	89.2±137.0	0.011

Table S3. Summary of metrics measured for each PDX type.

24 hrs

PDX	Cell count	% Extravasated by volume	p-value	Distance extravasated (μm)		Sphericity	p-value
Primary breast (PDX9040C1)	389	34.63±15.41	-	-4.22±14.36	-	0.57±0.11	-
TNBC (PDXbrC1)	517	53.12±21.61	2e-16	10.90±16.39	2e-16	0.65±0.14	2e-16
Tongue (PDXTonC1)	1423	78.30±21.26	2e-16	30.79±24.27	2e-16	0.59±0.14	6e-08
Lung (PDXLuC1)	677	70.90±31.37	2e-16	-16.68±14.57	2e-16	0.53±0.11	1e-07
Ovarian (PDXOvC1)	239	51.75±16.86	2e-16	-29.50±6.05	2e-16	0.61±0.13	7e-03

48 hrs

PDX	Cell count	% Extravasated by volume	p-value	Distance extravasated (μm)		Sphericity	p-value
Primary breast (PDX9040C1)	520	39.18±22.62	8e-06	4.30±17.69	1e-14	0.55±0.10	0.02
TNBC (PDXbrC1)	322	72.34±33.17	2e-16	17.59±18.59	2e-16	0.58±0.15	2e-08

Table S4. PDX type counts and volumes by location.

24 hrs

PDX	Cell count < 90%	Volume < 90% (voxels)	p-value	Cell count > 90%	Volume > 90% (voxels)	p-value
Primary breast (PDX9040C1)	324	519.2±479.7	-	2	474.1±684.1	-
TNBC (PDXbrC1)	417	589.0±472.8	1e-04	19	304.2±328.6	0.63
Tongue	455	687.8±500.1	6e-15	460	538.5±493.3	0.70
Lung	377	485.2±322.1	9e-11	275	420.3±299.6	0.50
Ovarian	199	613.5±419.0	6e-08	7	233.4±93.6	0.75

48 hrs

PDX	Cell count < 90%	Volume < 90% (voxels)	p-value	Cell count > 90%	Volume > 90% (voxels)	p-value
Primary breast (PDX9040C1)	448	460.2±433.8	4e-05	29	110±78.9	0.24
TNBC (PDXbrC1)	173	440.1±414.3	3e-06	127	297.6±311.2	0.67

Table S5. Comparison of methods to classify cancer cells by brain met potential.

Method	AUC	CA	F1	Precision	Recall
Neural Network	0.951	0.871	0.871	0.867	0.876
AdaBoost	0.950	0.876	0.876	0.874	0.877
Random Forest	0.946	0.874	0.874	0.873	0.875
Tree	0.917	0.843	0.839	0.857	0.823
kNN	0.868	0.787	0.776	0.817	0.739
Logistic Regression	0.848	0.779	0.783	0.769	0.796
Naïve Bayes	0.833	0.751	0.757	0.740	0.774
SGD	0.774	0.774	0.778	0.763	0.795

Table S6. Confusion matrix for random forest

		Predicted		
		0	1	∑
Actual	0	698	102	800
	1	100	700	800
	∑	798	802	1600

Table S7. Comparison of methods to classify breast PDX cancer cells by brain met potential.

Method	AUC	CA	F1	Precision	Recall
Neural Network	0.972	0.881	0.878	0.910	0.847
Random Forest	0.964	0.888	0.887	0.900	0.875
AdaBoost	0.957	0.881	0.879	0.899	0.861
Tree	0.954	0.867	0.865	0.884	0.847
Logistic Regression	0.897	0.832	0.831	0.843	0.819
Naïve Bayes	0.896	0.846	0.849	0.838	0.861
kNN	0.882	0.818	0.814	0.838	0.792
SGD	0.861	0.860	0.853	0.906	0.806

Table S8. Confusion matrix for random forest using PDX cancer cells.

		Predicted		
		0	1	∑
Actual	0	93	7	100
	1	8	92	100
	∑	101	99	200