

Table 1: Distribution of clinicopathological characteristics in the BMM cohort and associations to manual (n ≤ 411) and script-based (n ≤ 377) assessment of fraction of VEGF-C expressing stromal cells.

	VEGF-C Stromal Cell					<i>P</i> -value ^b	VEGF-C Stromal Cell	
	Total <i>n</i> (col %)	0% <i>n</i> (%)	1-5% <i>n</i> (%)	6-25% <i>n</i> (%)	≥26% <i>n</i> (%)		Script-based, % <i>median</i> (<i>IQR</i>)	<i>P</i> -value ^d
No. of patients	411	49	239	90	33		≤ 377	
Age, years^a	57.0 (30.1-85.4)	66.0 (42.0-81.3)	57.6 (35.1-84.4)	54.7 (33.0-75.9)	52.6 (45.1-73.8)	< 0.001	-0.26 ^c	<0.001
Detection						0.751		0.903
Screening	174 (42.4)	15 (8.6)	108 (62.1)	40 (23.0)	11 (6.3)		0.33 (1.10-1.20)	
Symptomatic	236 (57.6)	34 (14.4)	131 (55.5)	50 (21.2)	21 (8.9)		0.39 (0.10-1.10)	
Missing	1	0	0	0	1			
Tumor size						0.124		0.651
T1	274 (66.7)	34 (12.4)	165 (60.2)	56 (20.4)	19 (6.9)		0.39 (0.09-1.07)	
T2	137 (33.3)	15 (11.0)	74 (54.0)	34 (24.8)	14 (10.2)		0.33 (0.11-1.20)	
Tumor size, mm^a	18 (6-50)	16 (13-22)	18 (15-25)	19 (15-27)	20 (14-25)	0.053	0.07 ^c	0.159
Nodal status						0.001		<0.001
Negative	248 (60.3)	38 (15.3)	149 (60.1)	45 (18.2)	16 (6.4)		0.27 (0.05-0.85)	
Positive	163 (39.7)	11 (6.8)	90 (55.2)	45 (27.6)	17 (10.4)		0.56 (0.19-1.47)	
Multifocal						0.887		0.212
No	375 (91.2)	46 (12.3)	216 (57.6)	83 (22.1)	30 (8.0)		0.36 (0.09-1.10)	
Yes	36 (8.8)	3 (8.3)	23 (63.9)	7 (19.4)	3 (8.3)		0.45 (0.15-1.71)	
VI						0.001		<0.001
No	266 (72.1)	37 (13.9)	160 (60.2)	52 (19.6)	17 (6.4)		0.28 (0.08-0.86)	
Yes	103 (27.9)	7 (6.8)	52 (50.5)	31 (30.1)	13 (12.6)		0.57 (0.26-1.49)	
Missing	42	5	27	7	3			
TILs						<0.001		<0.001
0-9%	177 (48.0)	29 (16.4)	106 (59.9)	35 (19.8)	7 (4.0)		0.32 (0.08-0.98)	
10-49%	148 (40.1)	14 (9.5)	88 (59.5)	31 (21.0)	15 (10.1)		0.28 (0.10-0.91)	
50-74%	32 (8.7)	1 (3.1)	12 (37.5)	14 (43.8)	5 (15.6)		0.84 (0.39-1.67)	
≥75%	12 (3.2)	0 (0.0)	6 (50.0)	3 (25.0)	3 (25.0)		1.44 (0.60-2.33)	
Missing	42	5	27	7	3			
NHG						< 0.001		0.194
I	85 (20.8)	14 (16.5)	52 (61.2)	16 (18.8)	3 (3.5)		0.41 (0.11-1.30)	
II	211 (51.7)	25 (11.8)	134 (63.5)	38 (18.0)	14 (6.6)		0.30 (0.09-0.91)	
III	112 (27.4)	9 (8.0)	52 (46.4)	36 (32.1)	15 (13.4)		0.48 (0.15-1.23)	
Missing	3	1	1	0	1			
ER-status						0.024		0.002

Negative	49 (12.2)	1 (2.0)	30 (61.2)	11 (22.4)	7 (14.3)	0.84 (0.26-1.89)	
Positive	353 (87.8)	48 (13.6)	203 (57.5)	77 (21.8)	25 (7.1)	0.31 (0.09-1.00)	
Missing	9	0	6	2	1		
PR-status						0.022	0.003
Negative	86 (22.7)	7 (8.1)	47 (54.6)	20 (23.3)	12 (14.0)	0.61 (0.21-1.54)	
Positive	293 (77.3)	41 (14.0)	168 (57.3)	65 (22.2)	19 (6.5)	0.30 (0.08-1.03)	
Missing	32	1	24	5	2		
HER2-status						0.220	0.447
Negative	327 (80.3)	37 (11.3)	197 (60.2)	73 (22.3)	20 (6.1)	0.34 (0.08-1.10)	
Positive	80 (19.7)	12 (15.0)	40 (50.0)	16 (20.0)	12 (15.0)	0.44 (0.15-1.13)	
Missing	4	0	2	1	1		
Ki67 ≥ 20%						< 0.001	0.029
Negative	266 (67.0)	42 (15.8)	159 (59.8)	52 (19.6)	13 (4.9)	0.32 (0.08-1.04)	
Positive	131 (33.0)	7 (5.3)	70 (53.4)	36 (27.5)	18 (13.7)	0.44 (0.15-1.21)	
Missing	14	0	10	2	2		
Histological type						0.125	0.121
NST	298 (72.7)	30 (10.1)	177 (59.4)	66 (22.2)	25 (8.4)	0.39 (0.11-1.10)	
ILC	75 (18.3)	14 (18.7)	40 (53.3)	19 (25.3)	2 (2.7)	0.27 (0.02-0.98)	
Other	37 (9.0)	5 (13.5)	22 (59.5)	5 (13.5)	5 (13.5)	0.63 (0.14-2.26)	
Missing	1	0	0	0	0		
St Gallen Classification						0.013	0.029
Luminal A/B	275 (72.6)	36 (13.1)	163 (59.3)	62 (22.6)	14 (5.1)	0.31 (0.08-0.99)	
HER2+	75 (19.8)	12 (16.0)	36 (48.0)	16 (21.3)	11 (14.7)	0.42 (0.15-1.14)	
TNBC	29 (7.6)	0 (0.0)	17 (58.6)	7 (24.1)	5 (17.2)	1.01 (0.25-2.47)	
Missing	32	1	23	5	3		

Abbreviations and definitions: VEGF-C, Vascular endothelial growth factor C; IQR, Inter quartile range; Multifocal, > 2cm between tumors; VI, Vascular invasion; TILs, Tumor infiltrating lymphocytes; NHG, Nottingham histological grade; ER, Estrogen receptor >1%; PR, Progesterone receptor >1%; HER2, Human epidermal growth factor receptor 2; Ki67, a proliferation marker; NST, No specific type; ILC, Invasive lobular carcinoma; TNBC, triple negative breast cancer

^a Median (IQR)

^b Tests of zero slope in linear regression models with VEGF-C as the dependent variable and each of the clinicopathological variables as independent variables were used to assess evidence for association. Integer coding 1-4 was used for the four VEGF-C categories from the manual assessment. For non-ordinal variables Pearson's Chi² test was used.

^c Spearman's ρ

^d Test of $\rho=0$ for age and tumor size. Kruskal-Wallis test for categorical variables.

Table 2: Distribution of clinicopathological characteristics in the BMM cohort and associations to manual (n ≤ 411) and script-based (n ≤ 377) assessment of cancer cell VEGF-C intensity.

	Total n (col %)	Cancer cell VEGF-C Manual			P-value ^c	Cancer cell VEGF-C Script-based	
		Negative n (%)	Weak n (%)	Moderate/ Strong ^b n (%)		median (IQR) ^d	P- value ^e
No. of patients	411	17	333	61		≤ 375	
Age, years^a	57.0 (30.1-85.4)	58.4 (56.5-64.1)	57.8 (51.6-67.7)	53.2 (47.4-64.0)	0.010	-0.18 ^f	<0.001
Detection					0.138		<0.001
Screening	174 (42.4)	9 (5.2)	144 (82.8)	21 (12.1)		1.68 (1.23-2.50)	
Symptomatic	236 (57.6)	8 (3.4)	189 (80.1)	39 (16.5)		2.09 (1.58-2.78)	
Missing	1	0	0	1			
Tumor size					0.284		0.038
T1	274 (66.7)	15 (5.5)	219 (79.9)	40 (14.6)		1.85 (1.33-2.60)	
T2	137 (33.3)	2 (1.5)	114 (83.2)	21 (15.3)		2.08 (1.57-2.81)	
Tumor size, mm^a	18 (6-50)	14 (11-16)	18 (15-25)	17 (15-25)	0.269	0.12 ^f	0.016
Nodal status					0.896		0.316
Negative	248 (60.3)	12 (4.8)	198 (79.8)	38 (15.3)		1.88 (1.39-2.57)	
Positive	163 (39.7)	5 (3.1)	135 (82.8)	23 (14.1)		2.05 (1.44-2.72)	
Multifocal					0.112		0.022
No	375 (91.2)	14 (3.7)	303 (80.8)	58 (15.5)		1.99 (1.44-2.71)	
Yes	36 (8.8)	3 (8.3)	30 (83.3)	3 (8.3)		1.53 (1.20-2.20)	
VI					0.041		0.008
No	266 (72.1)	11 (4.1)	221 (83.1)	34 (12.8)		1.87 (1.33-2.57)	
Yes	103 (27.9)	2 (1.9)	80 (77.7)	21 (20.4)		2.14 (1.58-2.86)	
Missing	42	4	32	6			
TILs					0.078		0.067
0-9%	177 (48.0)	10 (5.6)	142 (80.2)	25 (14.1)		1.76 (1.32-2.49)	
10-49%	148 (40.1)	3 (2.0)	124 (83.8)	21 (14.2)		2.03 (1.49-2.81)	
50-74%	32 (8.7)	0 (0.0)	26 (81.2)	6 (18.8)		2.22 (1.72-2.85)	
≥75%	12 (3.2)	0 (0.0)	9 (75.0)	3 (25.0)		1.85 (1.62-2.78)	
Missing	42	4	32	6			
NHG					0.886		0.396
I	85 (20.8)	1 (1.2)	68 (80.0)	16 (18.8)		1.85 (1.41-2.55)	
II	211 (51.7)	13 (6.2)	174 (82.5)	24 (11.4)		1.93 (1.33-2.62)	
III	112 (27.4)	3 (2.7)	89 (79.5)	20 (17.9)		1.99 (1.54-2.69)	
Missing	3	0	2	1			

ER-status					0.001		0.051
Negative	49 (12.2)	0 (0.0)	36 (70.6)	15 (29.4)		2.22 (1.63-2.83)	
Positive	353 (87.8)	17 (4.8)	288 (82.0)	46 (13.1)		1.94 (1.39-2.62)	
Missing	9	0	9	0			
PR-status					0.048		0.140
Negative	86 (22.7)	1 (1.2)	68 (79.1)	17 (19.8)		2.12 (1.55-2.74)	
Positive	293 (77.3)	16 (5.5)	237 (80.9)	40 (13.6)		1.90 (1.33-2.68)	
Missing	32	0	28	4			
HER2-status					0.284		0.312
Negative	327 (80.3)	12 (3.7)	264 (80.7)	51 (15.6)		1.95 (1.37-2.63)	
Positive	80 (19.7)	5 (6.2)	65 (81.2)	10 (12.5)		1.96 (1.57-2.68)	
Missing	4	0	4	0			
Ki67 ≥ 20%					0.004		<0.001
Negative	266 (67.0)	16 (6.0)	216 (81.2)	34 (12.8)		1.81 (1.27-2.53)	
Positive	131 (33.0)	1 (0.8)	103 (78.6)	27 (20.6)		2.21 (1.62-2.90)	
Missing	14	0	14	0			
Histological type					<0.001		0.011
NST	298 (72.7)	8 (2.7)	247 (82.9)	43 (14.4)		2.00 (1.41-2.62)	
ILC	75 (18.3)	9 (12.0)	61 (81.3)	5 (6.7)		1.70 (1.18-2.30)	
Other	37 (9.0)	0 (0.0)	25 (67.6)	12 (32.4)		2.44 (1.71-2.91)	
Missing	1	0	0	1			
St Gallen Classification					<0.001		0.059
Luminal A/B	275 (72.6)	12 (4.4)	228 (82.9)	35 (12.7)		1.94 (1.35-2.58)	
HER2+	75 (19.8)	5 (6.7)	61 (81.3)	9 (12.0)		1.96 (1.56-2.72)	
TNBC	29 (7.6)	0 (0.0)	16 (55.2)	13 (44.8)		2.49 (1.79-3.14)	
Missing	32	0	28	4			

Abbreviations and definitions: VEGF-C, Vascular endothelial growth factor C; IQR, Inter quartile range; Multifocal, > 2cm between tumors; VI, Vascular invasion; TILs, Tumor infiltrating lymphocytes; NHG, Nottingham histological grade; ER, Estrogen receptor >1%; PR, Progesterone receptor >1%; HER2, Human epidermal growth factor receptor 2; Ki67, a proliferation marker; NST, No specific type; ILC, Invasive lobular carcinoma; TNBC, triple negative breast cancer

^a Median (IQR)

^b Only two patients with strong intensity, why the classes moderate and strong were merged.

^c Tests of zero slope in linear regression models with VEGF-C as the dependent variable and each of the clinicopathological variables as independent variables were used to assess evidence for association. Integer coding 1-4 was used for the three VEGF-C categories from the manual assessment. For non-ordinal variables Pearson's Chi² test was used.

^d Measured in arbitrary units. Values divided by 1 SD. Hence, a difference of 1 unit between groups corresponds to a 1 SD difference.

^e Test of $p=0$ for age and tumor size. Kruskal-Wallis test for categorical variables.

^f Spearman's ρ

Table 3: Univariable logistic regression analyses of clinicopathological variables on the outcome of positive lymph node status.

	<i>n</i>	OR (95% CI)	P-value
Age, years	411	0.968 (0.951-0.985)	< 0.001
Detection	410		< 0.001
Screening		1.00	
Symptomatic		2.08 (1.38)	
Tumor size, mm	411	1.101 (1.071-1.132)	< 0.001
Multifocal	411		0.002
No		1.00	
Yes		2.97 (1.46-6.05)	
VI	369		< 0.001
No		1.00	
Yes		3.41 (2.13-5.48)	
NHG	408		0.086 ^a
I		1.00	
II		1.16 (0.68-1.96)	0.591
III		1.80 (1.00-3.22)	0.048
ER-status	402		0.847
Positive		1.00	
Negative		1.06 (0.58-1.95)	
PR-status	379		0.854
Positive		1.00	
Negative		1.05 (0.64-1.71)	
HER2-status	407		0.269
Negative		1.00	
Positive		1.32 (0.81-2.17)	
Ki67 ≥ 20%	397		0.032
Negative		1.00	
Positive		1.59 (1.04-2.44)	
Histological type	410		0.453 ^a
NST		1.00	
ILC		0.74 (0.44-1.26)	0.274
Other		0.76 (0.37-1.55)	0.451
Stromal VEGF-C fraction, manual	411		0.005 ^a
0%		1.00	
1-5%		2.09 (1.02-4.29)	0.045
6-25%		3.45 (1.57-7.60)	0.002
≥26%		3.67 (1.41-9.56)	0.008
Stromal VEGF-C fraction, %, SB	377	1.140 (1.011-1.286)	0.019
Stromal VEGF-C fraction, SB log2-scale	377	1.212 (1.115- 1.318) ^b	< 0.001
Cancer cells Cytoplasmic	411		0.612 ^a

VEGF-C intensity, manual			
Negative		1.00	
Weak		1.64 (0.56-4.75)	0.365
Moderate/Strong		1.45 (0.45-4.66)	0.530
Cancer cells VEGF-C intensity, SB	375	1.04 (0.85-1.29) ^c	0.654

Abbreviations and definitions: *n*, Number of patients; OR, Odds ratio; CI, Confidence Interval; Multifocal, > 2cm between tumors; VI, Vascular invasion; NHG, Nottingham histological grade; ER, Estrogen receptor >1%; PR, Progesterone receptor >1%; HER2, Human epidermal growth factor receptor 2; Ki67, a proliferation marker; NST, No specific type; ILC, Invasive lobular carcinoma; VEGF-C, Vascular endothelial growth factor C; SB, script based

^a Test of overall effect: assesses whether the odds of nodal metastases are the same across all categories.

^b Modelled on log₂-scale; therefore, the odds ratio corresponds to a doubling of the fraction of VEGF-C-stained stromal cells.

^c Odds ratio per one standard deviation increase in intensity (measured in arbitrary units).

Table 4: Two multivariable logistic regression analyses for positive lymph node status, including stromal VEGF-C fraction assessed manually or script-based, respectively, along with other established clinicopathological predictors of this outcome. Two complete-case analyses based on data from the same 332 patients are presented. The reduction from 377 patients, who were assessed for stromal VEGF-C fraction by both methods, is mainly due to missing data on vascular invasion (VI), a factor that is difficult to assess preoperatively.

Variable	Manual Stromal VEGF-C			Script-based Stromal VEGF-C		
	OR	95% CI	P-value	OR	95% CI	P-value
VEGF-C			0.027 ^a	1.176 ^b	1.055-1.311	0.004
0%	1.00					
1-5%	2.55	0.86-7.54	0.092			
6-25%	4.00	1.26-12.7	0.019			
≥26%	3.72	0.95-14.6	0.059			
Age, years	0.971	0.949-0.993	0.009	0.972	0.950-0.994	0.014
Detection						
Screening	1.00			1.00		
Symptomatic	1.35	0.79-2.30	0.272	1.29	0.76-2.20	0.349
Tumor size, mm	1.102	1.064-1.142	<0.001	1.103	1.064-1.144	<0.001
Histological type						
NST	1.00			1.00		
ILC	0.52	0.24-1.13	0.101	0.52	0.24-1.13	0.097
Other	0.62	0.23-1.69	0.348	0.54	0.20-1.47	0.228
Multifocality						
No	1.00			1.00		
Yes	3.44	1.32-8.97	0.012	3.40	1.31-8.82	0.012
VI						
No	1.00			1.00		
Yes	2.14	1.21-3.78	0.009	1.94	1.09-3.45	0.024
NHG						
I	1.00			1.00		
II	0.77	0.37-1.59	0.477	0.82	0.39-1.70	0.587
III	0.57	0.25-1.35	0.202	0.68	0.29-1.59	0.375
ER-status						
Negative	1.00			1.00		
Positive	1.26	0.54-2.94	0.594	1.42	0.61-3.33	0.417
Constant^c	0.15	0.02-1.24		0.44	0.08-2.49	

Abbreviations and definitions: OR, Odds ratio; CI, Confidence interval; VEGF-C, Vascular endothelial growth factor C; NST, No special type; ILC, Invasive lobular carcinoma; Multifocality, >2 cm between tumors; VI, Vascular invasion; NHG, Nottingham histological grade; ER, Estrogen receptor >1%

^a Test of linear trend on the log-odds scale over the four manual VEGF-C categories conducted by assigning the scores 1 to 4 to the categories and adding this variable to the multivariable model instead of the corresponding factor variable with 3 degrees of freedom, thereby forcing the multiplicative increase in the odds of lymph node metastasis to be the same for all the three one-unit steps from 1 to 4.

^b Modelled on \log_2 -scale; therefore, the odds ratio corresponds to a doubling of the fraction of VEGF-C-stained stromal cells.

^c The constants presented in the OR columns represent reference odds and are part of each prediction formula. They have no meaningful interpretation without centering of non-categorical variables. Hence, no *P*-values are presented.