

Table 1: Summary of Results

Group <sup>†</sup>			Metric <sup>‡</sup>		
Repeats	Filters	Loss	Dice Score	Sensitivity	Positive Predictive Value
1-2	exp-16	BCE	0.65, 0.86 (IQR 0.49-0.90)	0.87, 0.88 (IQR 0.85-0.89)	0.86, 0.85 (IQR 0.84-0.87)
		Focal	0.67, 0.84 (IQR 0.57-0.89)	0.90, 0.90 (IQR 0.88-0.91)	0.82, 0.82 (IQR 0.80-0.84)
	lin-16	BCE	0.66, 0.83 (IQR 0.61-0.90)	0.86, 0.87 (IQR 0.84-0.88)	0.84, 0.84 (IQR 0.82-0.86)
		Focal	0.64, 0.85 (IQR 0.28-0.90)	0.83, 0.84 (IQR 0.80-0.86)	0.87, 0.87 (IQR 0.85-0.88)
2-2	exp-16	BCE	0.65, 0.84 (IQR 0.40-0.88)	0.88, 0.88 (IQR 0.86-0.90)	0.83, 0.83 (IQR 0.82-0.84)
		Focal	0.63, 0.86 (IQR 0.28-0.91)	0.85, 0.86 (IQR 0.84-0.88)	0.88, 0.88 (IQR 0.87-0.89)
	lin-16	BCE	0.61, 0.83 (IQR 0.28-0.88)	0.84, 0.84 (IQR 0.81-0.86)	0.83, 0.83 (IQR 0.81-0.85)
		Focal	0.65, 0.82 (IQR 0.50-0.88)	0.86, 0.87 (IQR 0.84-0.88)	0.82, 0.82 (IQR 0.80-0.83)
2-3	exp-16	BCE	0.64, 0.85 (IQR 0.18-0.90)	0.85, 0.86 (IQR 0.83-0.87)	0.86, 0.86 (IQR 0.85-0.88)
		Focal	0.70, 0.84 (IQR 0.64-0.90)	0.88, 0.89 (IQR 0.86-0.91)	0.85, 0.85 (IQR 0.84-0.86)
	lin-16	BCE	0.65, 0.87 (IQR 0.58-0.90)	0.85, 0.85 (IQR 0.82-0.87)	0.84, 0.84 (IQR 0.82-0.86)
		Focal	0.61, 0.84 (IQR 0.00-0.91)	0.83, 0.85 (IQR 0.81-0.86)	0.87, 0.87 (IQR 0.85-0.88)

<sup>†</sup> Results reported for: valid cohort, tumor = positive, class weights = 1, gamma foc = 2.0, resnet = False, convnext = False, norm = layernorm, func = leakyrelu

<sup>‡</sup> Table values are shown as: mean, median (interquartile range)













Table 2: Comparison Between Experiments - Dice Score

Group <sup>†</sup>				Violin	Exp <sup>‡</sup>											
Repeats	Filters	Loss	Exp		1	2	3	4	5	6	7	8	9	10	11	12
1-2	exp-16	BCE	1		–	N.S.	N.S.	N.S.	N.S.	N.S.	N.S.	N.S.	N.S.	N.S.	N.S.	N.S.
		Focal	2		N.S.	–	N.S.	N.S.	*	N.S.	*	*	N.S.	N.S.	N.S.	*
	lin-16	BCE	3		N.S.	N.S.	–	N.S.	N.S.	N.S.	N.S.	N.S.	N.S.	N.S.	N.S.	N.S.
		Focal	4		N.S.	N.S.	N.S.	–	N.S.	N.S.	N.S.	N.S.	N.S.	*	N.S.	N.S.
2-2	exp-16	BCE	5		N.S.	*	N.S.	N.S.	–	N.S.	N.S.	N.S.	N.S.	N.S.	N.S.	N.S.
		Focal	6		N.S.	N.S.	N.S.	N.S.	N.S.	–	N.S.	N.S.	N.S.	*	N.S.	N.S.
	lin-16	BCE	7		N.S.	*	N.S.	N.S.	N.S.	N.S.	–	N.S.	N.S.	*	N.S.	N.S.
		Focal	8		N.S.	*	N.S.	N.S.	N.S.	N.S.	N.S.	–	N.S.	N.S.	N.S.	N.S.
2-3	exp-16	BCE	9		N.S.	N.S.	N.S.	N.S.	N.S.	N.S.	N.S.	N.S.	–	*	N.S.	N.S.
		Focal	10		N.S.	N.S.	N.S.	*	N.S.	*	*	N.S.	*	–	N.S.	*
	lin-16	BCE	11		N.S.	N.S.	N.S.	N.S.	N.S.	N.S.	N.S.	N.S.	N.S.	–	N.S.	N.S.
		Focal	12		N.S.	*	N.S.	N.S.	N.S.	N.S.	N.S.	N.S.	N.S.	*	N.S.	–

<sup>†</sup> Results reported for: valid cohort, tumor = positive, class weights = 1, gamma foc = 2.0, resnet = False, convnext = False, norm = layernorm, func = leakyrelu








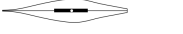




<sup>‡</sup> N.S. = not significant, \* =  $p < 0.05$

Table 3: Comparison Between Experiments - Sensitivity

Group <sup>†</sup>				Violin	Exp <sup>‡</sup>												
Repeats	Filters	Loss	Exp	Sensitivity	1	2	3	4	5	6	7	8	9	10	11	12	
1-2	exp-16	BCE	1		—	**	**	**	**	**	**	**	**	**	**	**	
		Focal	2		**	—	**	**	**	**	**	**	**	**	**	**	
	lin-16	BCE	3		**	**	—	**	**	**	**	<i>N.S.</i>	**	**	**	**	
		Focal	4		**	**	**	—	**	**	**	**	**	**	**	*	
2-2	exp-16	BCE	5		**	**	**	**	—	**	**	**	**	**	**	**	
		Focal	6		**	**	**	**	**	—	**	**	**	**	**	**	
	lin-16	BCE	7		**	**	**	**	**	**	—	**	**	**	**	<i>N.S.</i>	
		Focal	8		**	**	<i>N.S.</i>	**	**	**	**	**	—	**	**	**	
2-3	exp-16	BCE	9		**	**	**	**	**	**	**	**	—	**	**	**	
		Focal	10		**	**	**	**	**	**	**	**	**	—	**	**	
	lin-16	BCE	11		**	**	**	**	**	**	**	**	**	**	—	**	
		Focal	12		**	**	**	*	**	**	<i>N.S.</i>	**	**	**	**	—	

<sup>†</sup> Results reported for: valid cohort, tumor = positive, class weights = 1, gamma foc = 2.0, resnet = False, convnext = False, norm = layernorm, func = leakyrelu  
<sup>‡</sup> *N.S.* = not significant, \* =  $p < 0.05$ , \*\* =  $p < 0.005$

Table 4: Comparison Between Experiments - Positive Predictive Value

Group <sup>†</sup>				Violin		Exp <sup>‡</sup>											
Repeats	Filters	Loss	Exp	Positive Predictive Value		1	2	3	4	5	6	7	8	9	10	11	12
1-2	exp-16	BCE	1			—	**	**	**	**	**	**	**	**	**	**	**
		Focal	2		**	—	**	**	**	**	**	**	<i>N.S.</i>	**	**	**	**
	lin-16	BCE	3		**	**	—	**	**	**	**	**	**	**	**	**	**
		Focal	4		**	**	**	—	**	**	**	**	*	**	**	**	<i>N.S.</i>
2-2	exp-16	BCE	5		**	**	**	**	—	**	<i>N.S.</i>	**	**	**	**	**	**
		Focal	6		**	**	**	**	**	—	**	**	**	**	**	**	**
	lin-16	BCE	7		**	**	**	**	<i>N.S.</i>	**	—	**	**	**	**	**	**
		Focal	8		**	<i>N.S.</i>	**	**	**	**	**	**	—	**	**	**	**
2-3	exp-16	BCE	9		**	**	**	*	**	**	**	**	**	—	**	**	*
		Focal	10		**	**	**	**	**	**	**	**	**	**	—	*	**
	lin-16	BCE	11		**	**	**	**	**	**	**	**	**	**	*	—	**
		Focal	12		**	**	**	<i>N.S.</i>	**	**	**	**	*	**	**	**	—

<sup>†</sup> Results reported for: valid cohort, tumor = positive, class weights = 1, gamma foc = 2.0, resnet = False, convnext = False, norm = layernorm, func = leakyrelu  
<sup>‡</sup> *N.S.* = not significant, \* =  $p < 0.05$ , \*\* =  $p < 0.005$

Figure 1: Comparison Between Experiments - Violin Plots

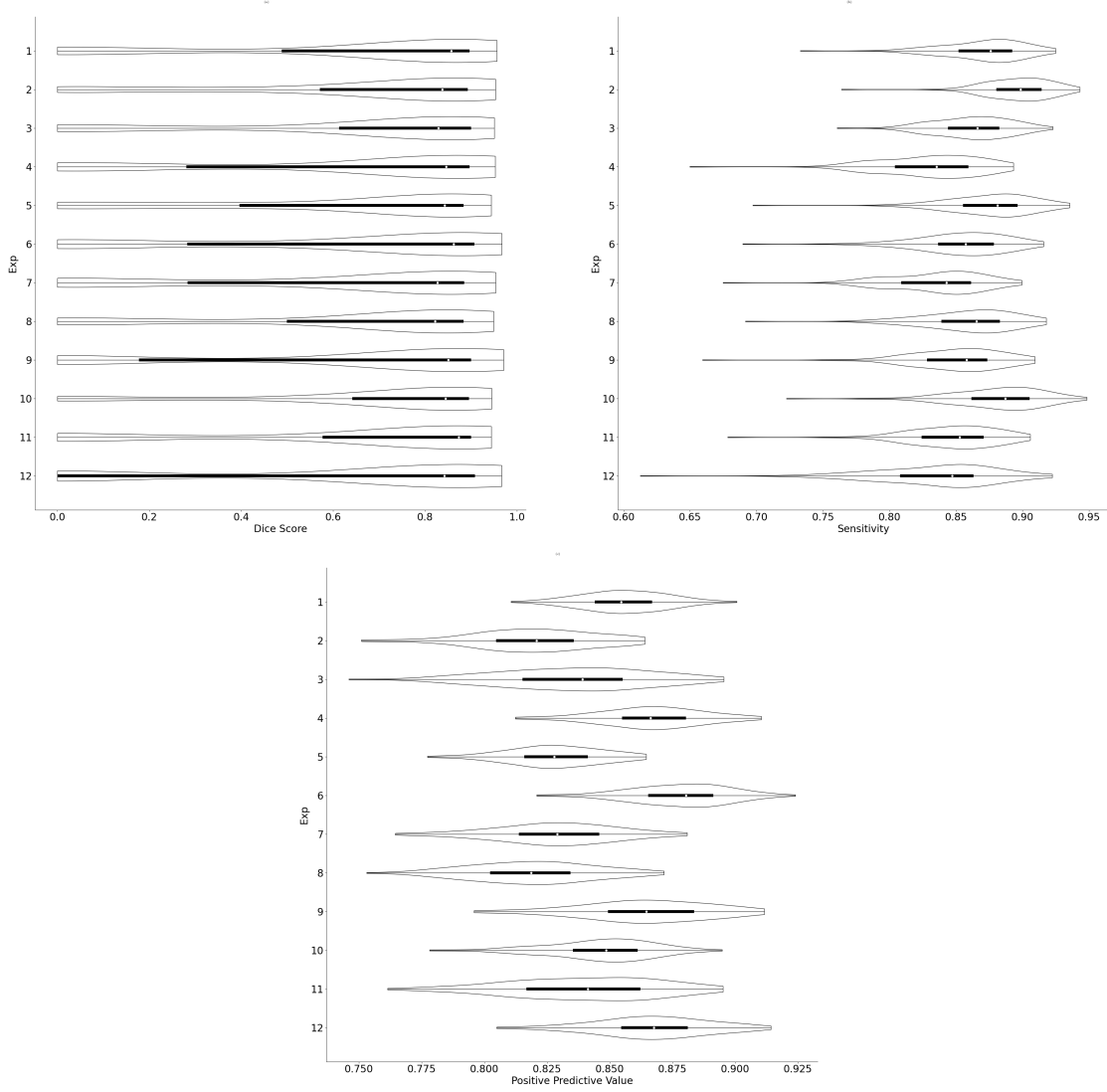


Table 5: Mean Differences - Dice Score

Group <sup>†</sup>				Exp											
Repeats	Filters	Loss	Exp	1	2	3	4	5	6	7	8	9	10	11	12
1-2	exp-16	BCE	1	-	0.02 (p = 0.473)	0.00 (p = 0.858)	-0.02 (p = 0.572)	-0.01 (p = 0.758)	-0.02 (p = 0.548)	-0.04 (p = 0.227)	-0.01 (p = 0.804)	-0.02 (p = 0.576)	0.05 (p = 0.226)	-0.00 (p = 0.891)	-0.04 (p = 0.169)
		Focal	2	-0.02 (p = 0.473)	-	-0.03 (p = 0.453)	-0.02 (p = 0.146)	-0.04 (p = 0.025)	-0.04 (p = 0.140)	-0.06 (p = 0.036)	-0.02 (p = 0.025)	-0.03 (p = 0.128)	0.03 (p = 0.282)	-0.02 (p = 0.185)	-0.06 (p = 0.026)
	lin-16	BCE	3	-0.00 (p = 0.858)	0.01 (p = 0.453)	-	-0.02 (p = 0.509)	-0.01 (p = 0.644)	-0.03 (p = 0.425)	-0.04 (p = 0.128)	-0.01 (p = 0.593)	-0.02 (p = 0.377)	0.05 (p = 0.203)	-0.01 (p = 0.765)	-0.05 (p = 0.171)
		Focal	4	0.02 (p = 0.572)	0.03 (p = 0.146)	0.02 (p = 0.509)	-	0.01 (p = 0.640)	-0.01 (p = 0.858)	-0.02 (p = 0.352)	0.01 (p = 0.633)	-0.00 (p = 0.997)	0.07 (p = 0.018)	0.01 (p = 0.497)	-0.03 (p = 0.072)
	2-2	exp-16	BCE	5	0.01 (p = 0.758)	0.02 (p = 0.025)	0.01 (p = 0.644)	-0.01 (p = 0.640)	-	-0.01 (p = 0.574)	-0.03 (p = 0.205)	0.00 (p = 0.960)	-0.01 (p = 0.675)	0.06 (p = 0.062)	-0.03 (p = 0.840)
		Focal	6	0.02 (p = 0.548)	0.04 (p = 0.140)	0.03 (p = 0.425)	0.01 (p = 0.858)	0.01 (p = 0.574)	-	-0.02 (p = 0.463)	0.02 (p = 0.547)	0.01 (p = 0.840)	0.07 (p = 0.042)	0.02 (p = 0.469)	-0.02 (p = 0.547)
	lin-16	BCE	7	0.04 (p = 0.227)	0.06 (p = 0.036)	0.04 (p = 0.128)	0.02 (p = 0.352)	0.03 (p = 0.205)	0.02 (p = 0.463)	-	0.03 (p = 0.186)	0.02 (p = 0.346)	0.09 (p = 0.007)	0.04 (p = 0.151)	-0.00 (p = 0.989)
		Focal	8	0.01 (p = 0.804)	0.02 (p = 0.025)	0.01 (p = 0.593)	-0.01 (p = 0.633)	-0.00 (p = 0.960)	-0.02 (p = 0.547)	-0.03 (p = 0.186)	-	-0.01 (p = 0.614)	0.06 (p = 0.053)	0.00 (p = 0.836)	-0.04 (p = 0.132)
2-3	exp-16	BCE	9	0.02 (p = 0.576)	0.03 (p = 0.128)	0.02 (p = 0.377)	0.00 (p = 0.997)	0.01 (p = 0.675)	-0.01 (p = 0.840)	-0.02 (p = 0.346)	0.01 (p = 0.614)	-	0.07 (p = 0.048)	0.01 (p = 0.444)	-0.03 (p = 0.372)
		Focal	10	-0.05 (p = 0.226)	-0.03 (p = 0.282)	-0.05 (p = 0.203)	-0.07 (p = 0.018)	-0.06 (p = 0.062)	-0.07 (p = 0.042)	-0.09 (p = 0.007)	-0.06 (p = 0.053)	-0.07 (p = 0.048)	-	-0.05 (p = 0.056)	-0.09 (p = 0.009)
	lin-16	BCE	11	0.00 (p = 0.891)	0.02 (p = 0.185)	0.01 (p = 0.765)	-0.01 (p = 0.497)	-0.00 (p = 0.840)	-0.02 (p = 0.469)	-0.04 (p = 0.151)	-0.00 (p = 0.836)	-0.01 (p = 0.444)	0.05 (p = 0.056)	-	-0.04 (p = 0.073)
		Focal	12	0.04 (p = 0.169)	0.06 (p = 0.026)	0.05 (p = 0.171)	0.03 (p = 0.072)	0.03 (p = 0.128)	0.02 (p = 0.547)	0.00 (p = 0.989)	0.04 (p = 0.132)	0.03 (p = 0.372)	0.09 (p = 0.009)	0.04 (p = 0.073)	-

<sup>†</sup> Results reported for: valid cohort, tumor = positive, class weights = 1, gamma foc = 2.0, resnet = False, convnext = False, norm = layernorm, func = leakyrelu

Table 6: Mean Differences - Sensitivity

Group <sup>†</sup>				Exp											
Repeats	Filters	Loss	Exp	1	2	3	4	5	6	7	8	9	10	11	12
1-2	exp-16	BCE	1	–	0.03 (p < 0.001)	-0.01 (p < 0.001)	-0.04 (p < 0.001)	0.00 (p = 0.001)	-0.02 (p < 0.001)	-0.04 (p < 0.001)	-0.01 (p < 0.001)	-0.02 (p < 0.001)	0.01 (p < 0.001)	-0.02 (p < 0.001)	-0.04 (p < 0.001)
			2	-0.03 (p < 0.001)	–	-0.03 (p < 0.001)	-0.07 (p < 0.001)	-0.02 (p < 0.001)	-0.04 (p < 0.001)	-0.06 (p < 0.001)	-0.04 (p < 0.001)	-0.05 (p < 0.001)	-0.02 (p < 0.001)	-0.05 (p < 0.001)	-0.06 (p < 0.001)
		lin-16	3	0.01 (p < 0.001)	0.03 (p < 0.001)	–	-0.03 (p < 0.001)	0.01 (p < 0.001)	-0.01 (p = 0.001)	-0.03 (p < 0.001)	-0.00 (p = 0.288)	-0.01 (p < 0.001)	0.02 (p < 0.001)	-0.01 (p < 0.001)	-0.03 (p < 0.001)
			4	0.04 (p < 0.001)	0.07 (p < 0.001)	0.03 (p < 0.001)	–	0.05 (p < 0.001)	0.02 (p < 0.001)	0.00 (p < 0.001)	0.03 (p < 0.001)	0.02 (p < 0.001)	0.05 (p < 0.001)	0.02 (p < 0.001)	0.00 (p = 0.047)
	exp-16	BCE	5	-0.00 (p = 0.001)	0.02 (p < 0.001)	-0.01 (p < 0.001)	-0.05 (p < 0.001)	–	-0.02 (p < 0.001)	-0.04 (p < 0.001)	-0.02 (p < 0.001)	-0.03 (p < 0.001)	0.01 (p = 0.003)	-0.03 (p < 0.001)	-0.04 (p < 0.001)
			6	0.02 (p < 0.001)	0.04 (p < 0.001)	0.01 (p = 0.001)	-0.02 (p < 0.001)	0.02 (p < 0.001)	–	-0.02 (p < 0.001)	0.01 (p < 0.001)	-0.00 (p < 0.001)	0.03 (p < 0.001)	-0.01 (p < 0.001)	-0.02 (p < 0.001)
		lin-16	7	0.04 (p < 0.001)	0.06 (p < 0.001)	0.03 (p < 0.001)	-0.00 (p < 0.001)	0.04 (p < 0.001)	0.02 (p < 0.001)	–	-0.02 (p < 0.001)	0.02 (p < 0.001)	0.05 (p < 0.001)	0.01 (p < 0.001)	-0.00 (p = 0.850)
			8	0.01 (p < 0.001)	0.04 (p < 0.001)	0.00 (p = 0.288)	-0.03 (p < 0.001)	0.02 (p < 0.001)	-0.01 (p < 0.001)	-0.02 (p < 0.001)	–	-0.01 (p < 0.001)	0.02 (p < 0.001)	-0.01 (p < 0.001)	-0.02 (p < 0.001)

<sup>†</sup> Results reported for: valid cohort, tumor = positive, class weights = 1, gamma foc = 2.0, resnet = False, convnext = False, norm = layernorm, func = leakyrelu

Table 7: Mean Differences - Positive Predictive Value

Group <sup>†</sup>				Exp											
Repeats	Filters	Loss	Exp	1	2	3	4	5	6	7	8	9	10	11	12
1-2	exp-16	BCE	1	–	-0.04 (p < 0.001)	-0.02 (p < 0.001)	0.01 (p < 0.001)	-0.03 (p < 0.001)	0.02 (p < 0.001)	-0.03 (p < 0.001)	-0.04 (p < 0.001)	0.01 (p < 0.001)	-0.01 (p < 0.001)	-0.02 (p < 0.001)	0.01 (p < 0.001)
			2	0.04 (p < 0.001)	–	0.02 (p < 0.001)	0.05 (p < 0.001)	0.01 (p < 0.001)	0.06 (p < 0.001)	0.01 (p < 0.001)	-0.00 (p = 0.361)	0.04 (p < 0.001)	0.03 (p < 0.001)	0.02 (p < 0.001)	0.05 (p < 0.001)
		lin-16	3	0.02 (p < 0.001)	-0.02 (p < 0.001)	–	0.03 (p < 0.001)	-0.01 (p < 0.001)	0.04 (p < 0.001)	-0.01 (p < 0.001)	-0.02 (p < 0.001)	0.03 (p < 0.001)	0.01 (p < 0.001)	0.00 (p < 0.001)	0.03 (p < 0.001)
			4	-0.01 (p < 0.001)	-0.05 (p < 0.001)	-0.03 (p < 0.001)	–	-0.04 (p < 0.001)	0.01 (p < 0.001)	-0.04 (p < 0.001)	-0.05 (p < 0.001)	-0.00 (p = 0.044)	-0.02 (p < 0.001)	-0.03 (p < 0.001)	0.00 (p = 0.788)
	exp-16	BCE	5	0.03 (p < 0.001)	-0.01 (p < 0.001)	0.01 (p < 0.001)	0.04 (p < 0.001)	–	0.05 (p < 0.001)	-0.00 (p = 0.992)	-0.01 (p < 0.001)	0.04 (p < 0.001)	0.02 (p < 0.001)	0.01 (p < 0.001)	0.04 (p < 0.001)
			6	-0.02 (p < 0.001)	-0.06 (p < 0.001)	-0.04 (p < 0.001)	-0.01 (p < 0.001)	-0.05 (p < 0.001)	–	-0.05 (p < 0.001)	-0.06 (p < 0.001)	-0.01 (p < 0.001)	-0.03 (p < 0.001)	-0.04 (p < 0.001)	-0.01 (p < 0.001)
		lin-16	7	0.03 (p < 0.001)	-0.01 (p < 0.001)	0.01 (p < 0.001)	0.04 (p < 0.001)	0.00 (p = 0.992)	0.05 (p < 0.001)	–	-0.01 (p < 0.001)	0.04 (p < 0.001)	0.02 (p < 0.001)	0.01 (p < 0.001)	0.04 (p < 0.001)
			8	0.04 (p < 0.001)	0.00 (p = 0.361)	0.02 (p < 0.001)	0.05 (p < 0.001)	0.01 (p < 0.001)	0.06 (p < 0.001)	0.01 (p < 0.001)	–	0.04 (p < 0.001)	0.03 (p < 0.001)	0.02 (p < 0.001)	0.05 (p < 0.001)
2-3	exp-16	BCE	9	-0.01 (p < 0.001)	-0.04 (p < 0.001)	-0.03 (p < 0.001)	0.00 (p = 0.044)	-0.04 (p < 0.001)	0.01 (p < 0.001)	-0.04 (p < 0.001)	-0.04 (p < 0.001)	–	-0.02 (p < 0.001)	-0.02 (p < 0.001)	0.00 (p = 0.010)
			10	0.01 (p < 0.001)	-0.03 (p < 0.001)	-0.01 (p < 0.001)	0.02 (p < 0.001)	-0.02 (p < 0.001)	0.03 (p < 0.001)	-0.02 (p < 0.001)	-0.03 (p < 0.001)	0.02 (p < 0.001)	–	-0.01 (p = 0.006)	0.02 (p < 0.001)
		lin-16	11	0.02 (p < 0.001)	-0.02 (p < 0.001)	-0.00 (p < 0.001)	0.03 (p < 0.001)	-0.01 (p < 0.001)	0.04 (p < 0.001)	-0.01 (p < 0.001)	-0.02 (p < 0.001)	0.02 (p < 0.001)	0.01 (p = 0.006)	–	0.03 (p < 0.001)
			12	-0.01 (p < 0.001)	-0.05 (p < 0.001)	-0.03 (p < 0.001)	-0.00 (p = 0.788)	-0.04 (p < 0.001)	0.01 (p < 0.001)	-0.04 (p < 0.001)	-0.05 (p < 0.001)	-0.00 (p = 0.010)	-0.02 (p < 0.001)	-0.03 (p < 0.001)	–

<sup>†</sup> Results reported for: valid cohort, tumor = positive, class weights = 1, gamma foc = 2.0, resnet = False, convnext = False, norm = layernorm, func = leakyrelu