

Supplementary Table SI. Correlation between clinicopathological features and NAT efficacy

Clinicopathological characteristics	Total (n = 201)	Pathological response to NAC		χ^2	P-value
		RCB0~I	RCBII~ III		
Age					
≤ 50	98	49	49	0.574	0.448
> 50	103	57	46		
Histological grade					
1~2	69	24	45	13.588	< 0.001
3	132	82	50		
Histological type					
Invasive carcinoma of no special type	183	103	80	10.320	0.001
Other invasive histology	18	3	15		
Estrogen Receptor (ER) status					
Negative	87	63	24	23.829	< 0.001
Positive	114	43	71		
Progesterone Receptor (PR) status					
Negative	111	79	32	33.800	< 0.001
Positive	90	27	63		
Androgen Receptor (AR) status *					
Negative	37	21	16	0.287	0.592
Positive	160	83	77		
HER2					
Negative	135	45	90	62.099	< 0.001
Positive	66	61	5		
Ki67 expression status (cut-of 30%)					
Low (< 30%)	47	14	33	12.962	< 0.001
High (≥ 30%)	154	92	62		

* Some cases with AR status were unknown and did not participate in statistical analysis.

Supplementary Table SII. Clinical characteristics of patients in training set and validation set

Clinicopathological characteristics	Training set (n = 140)		Validation set (n = 61)		χ^2	P-value
	No.	%	No.	%		
Age					0.708	0.400
≤ 50	71	50.71	27	44.26		
> 50	69	49.29	34	55.74		
Histological grade					2.673	0.102
1~2	43	30.71	26	42.62		
3	97	69.29	35	57.38		
Histological type					0.083	0.773
Invasive carcinoma of no special type	128	91.43	55	90.16		
Other invasive histology	12	8.57	6	9.84		
ER status					3.931	0.047
Negative	67	47.86	20	32.79		
Positive	73	52.14	41	67.21		
PR status					7.182	0.007
Negative	86	61.43	25	40.98		
Positive	54	38.57	36	59.02		
AR status *					0.134	0.714
Negative	25	18.12	12	20.34		
Positive	113	81.88	47	79.66		
HER2					0.113	0.737
Negative	93	66.43	42	69.85		
Positive	47	33.57	19	31.15		
Ki67 expression status (cut-of 30%)					5.962	0.015
Low (< 30%)	26	18.57	21	34.43		
High (≥ 30%)	114	81.43	40	65.57		
RCB grade					2.523	0.112
RCB0~I	79	56.43	27	44.26		
RCBII~III	61	43.57	34	55.74		

Supplementary Table SIII. The selected pathomics features

Pathomics features (*n* = 12)

Granularity_12_Hematoxylin

Granularity_8_Hematoxylin

Granularity_9_Hematoxylin

ImageQuality_MaxIntensity_Hematoxylin

Texture_Correlation_Hematoxylin_3_00_256

Mean_IdentifyPrimaryObjects_AreaShape_MaximumRadius

Mean_IdentifyPrimaryObjects_AreaShape_Zernike_1_1

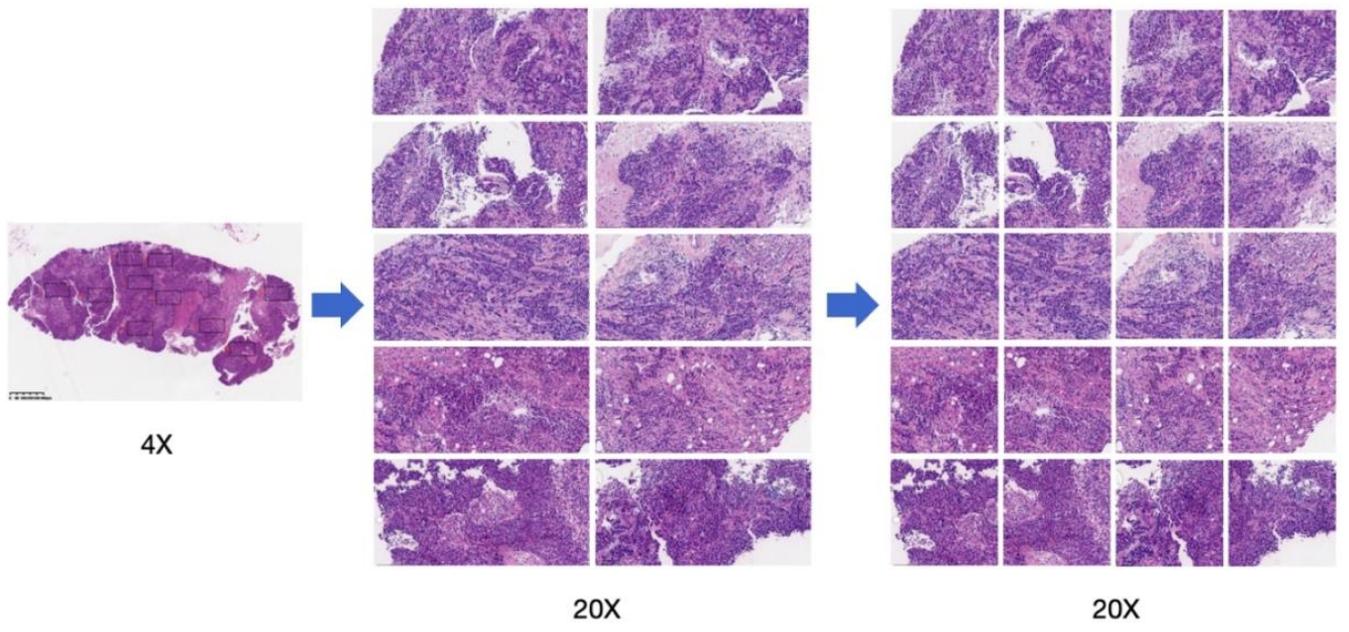
Mean_IdentifyPrimaryObjects_AreaShape_Zernike_6_6

Mean_IdentifyPrimaryObjects_Texture_InfoMeas2_Hematoxylin_3_00_256

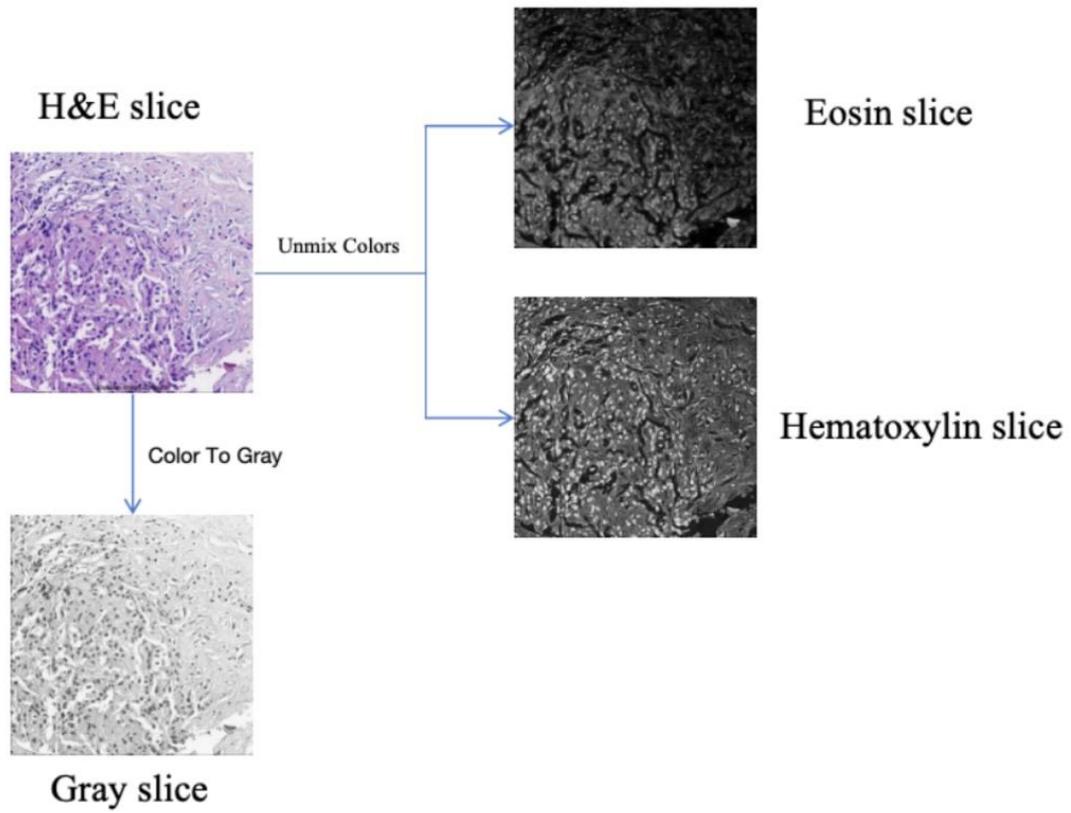
Mean_IdentifySecondaryObjects_AreaShape_Zernike_3_1

StDev_IdentifyPrimaryObjects_Intensity_MADIntensity_Hematoxylin

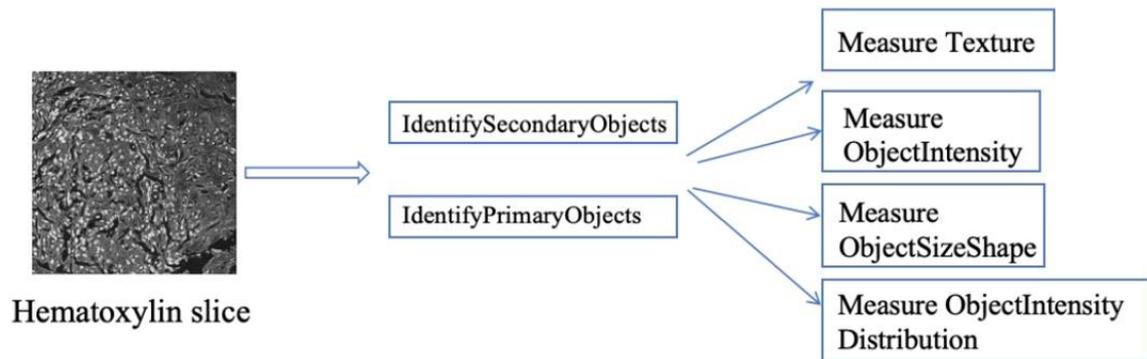
StDev_IdentifySecondaryObjects_AreaShape_Zernike_0_0



Supplementary Figure S1. Segmentation process of WSI



Supplementary Figure S2. Image preprocessing in CellProBased on the “Unmix Colors” module to separate H&E-stained images and convert them into haematoxylin-stained and eosin-stained greyscale images, The H&E-stained images were also converted to greyscale images using the “ColorToGray” module



Supplementary Figure S3. Haematoxylin-stained images were segmented via ‘Identify Primary Objects’ module and ‘Identify Secondary Objects’ module ,Quantitative image features of object shape, size, texture, and pixel intensity distribution were further extracted via multiple modules, including measure models of ‘Object Intensity Distribution’, ‘Object Intensity’, ‘Texture’, and ‘Object Size Shape’