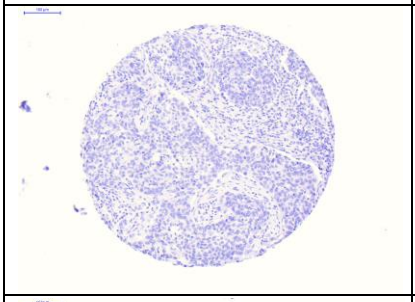
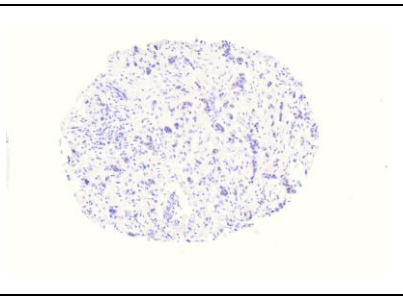
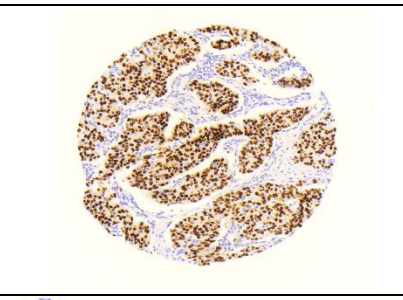
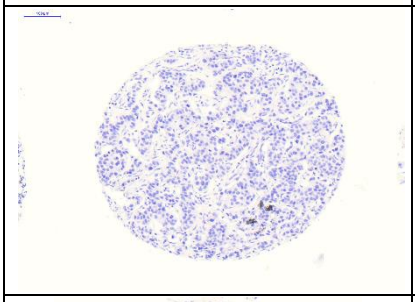
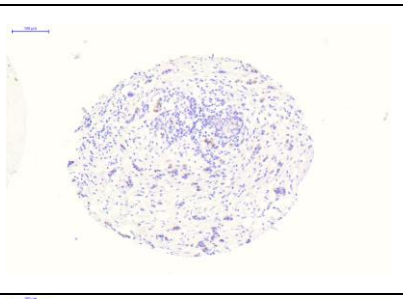
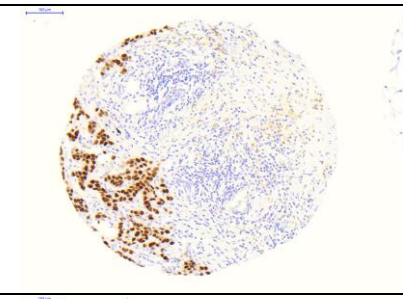
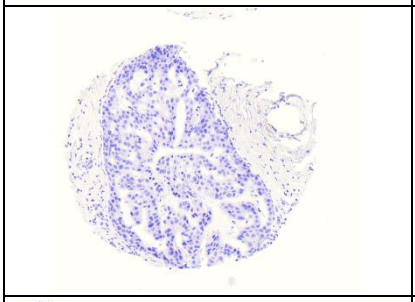
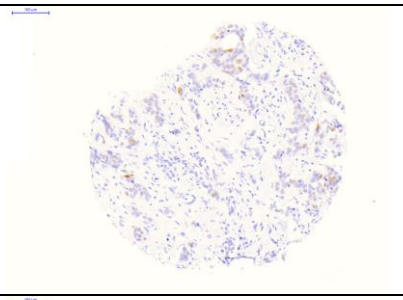
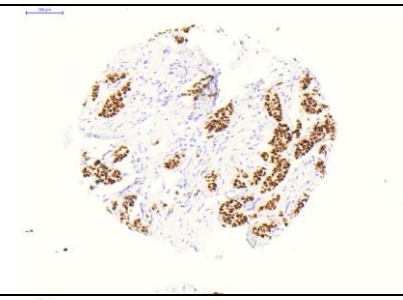
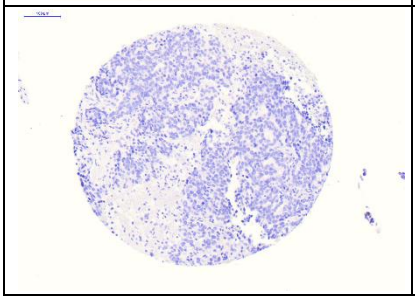
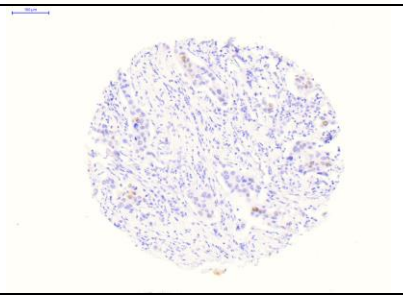
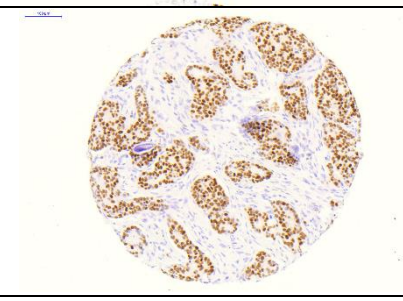
Supplementary 1. Core information about the four tissue microarrays (TMAs) stained for estrogen receptor (ER), progesterone receptor (PR), human epidermal growth factor receptor 2 (HER2), epithelial growth factor receptor (EGFR), and cytokeratin 5/6 (CK5/6).

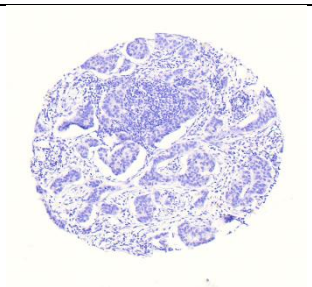
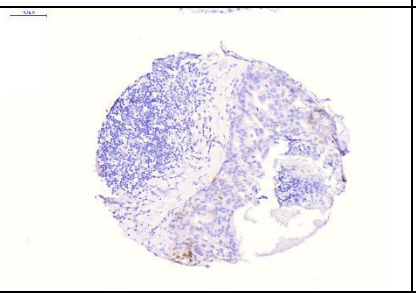
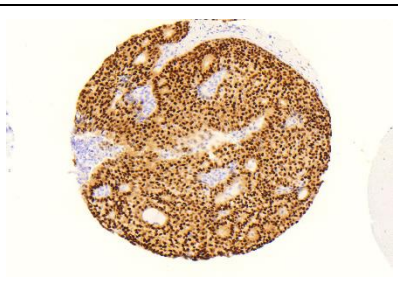
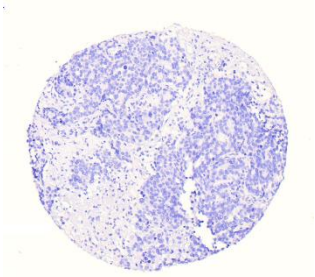
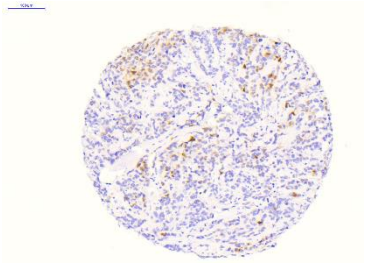
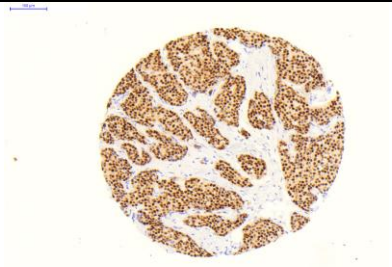
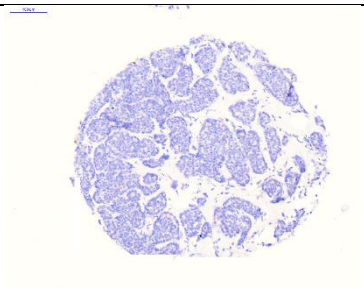
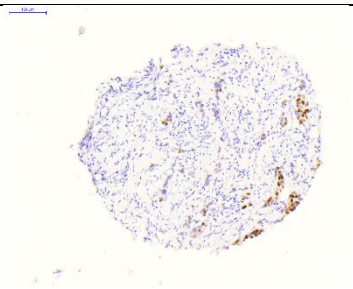
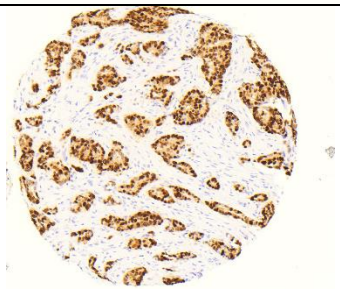
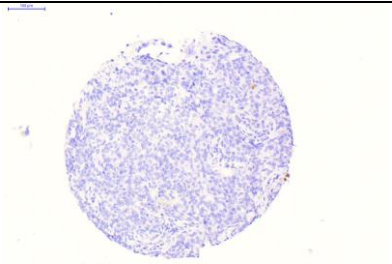
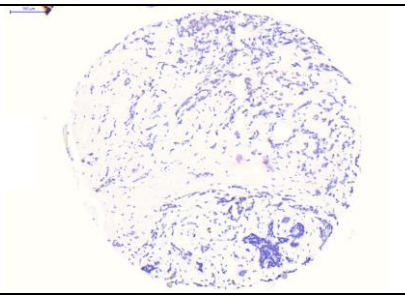
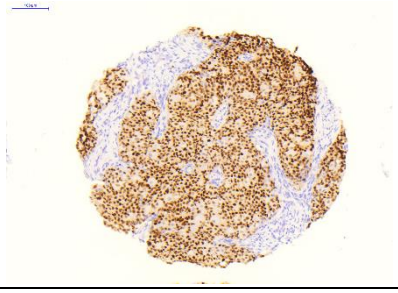
	ER	**PR**	**HER2**	**EGFR**	**CK5/6**
Total cores across the 4 TMAs, *n*	1197	1197	1197	1197	1197
Total women across the 4 TMAs, *n*	385	385	385	385	385
Cores with no tumor after pathologist's assessment, *n*	167	157	151	131	128
Cores determined as missing or not evaluable by pathologist, *n*	98	109	119	105	146
Cores evaluated by pathologist, *n*	932	931	927	961	923
Resultant women in this study, *n*	358	359	357	361	359

Supplementary 2. Twelve cores were chosen to train tissue segmentation and establish immunohistochemistry thresholds for each clinical marker. A board-certified pathologist scored each of these markers as negative (<1%), low positive (1-10%), and positive (>10%).

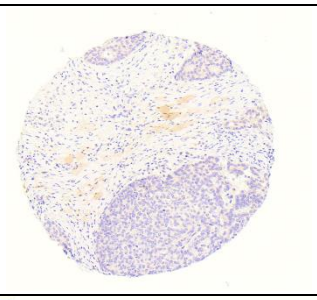
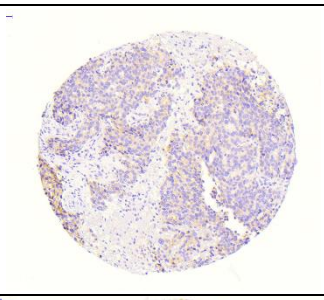
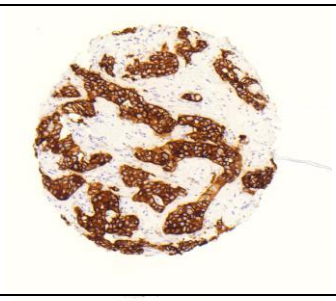
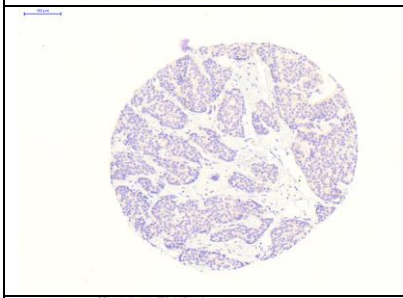
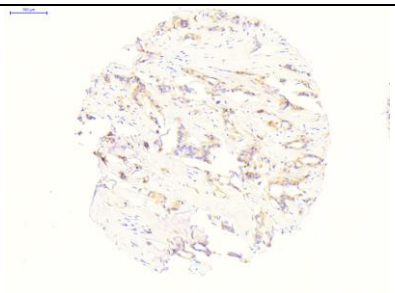
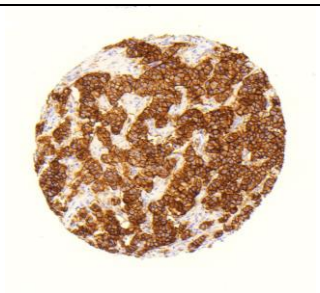
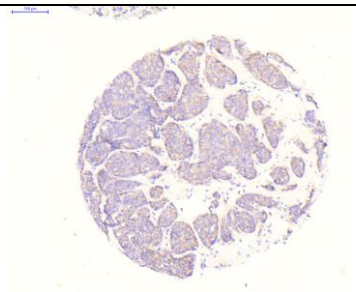
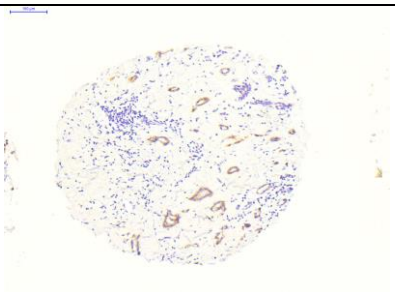
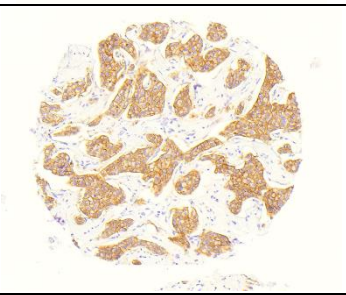
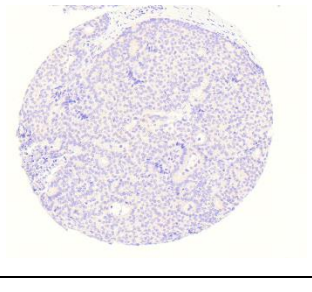
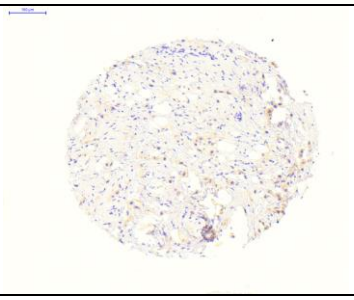
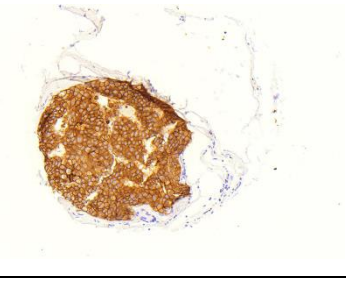
A. Estrogen receptor.

Negative (<1%)	**Low Positive (1-10%)**	**Positive (>10%)**
		
		
		
		

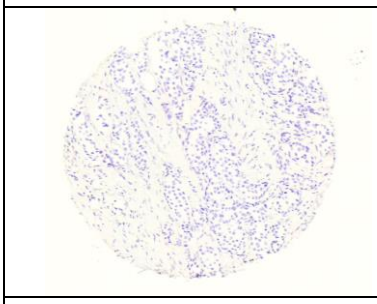
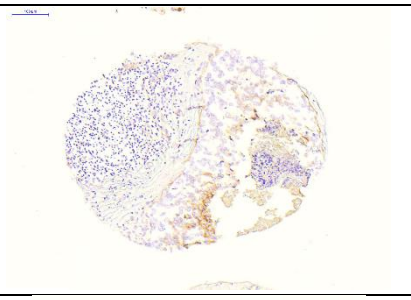
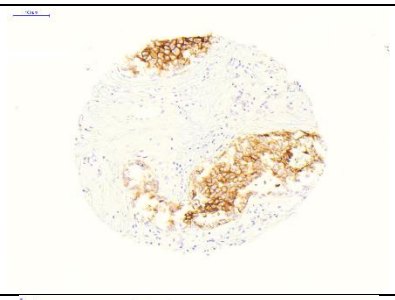
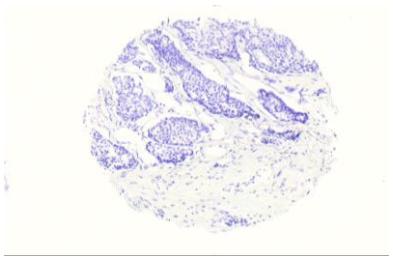
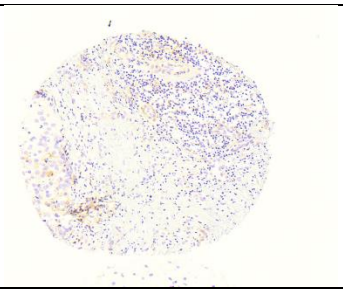
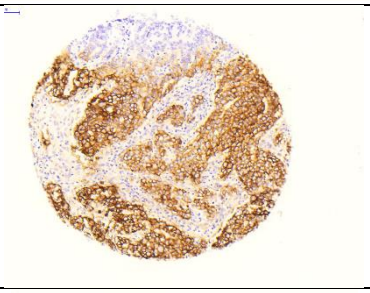
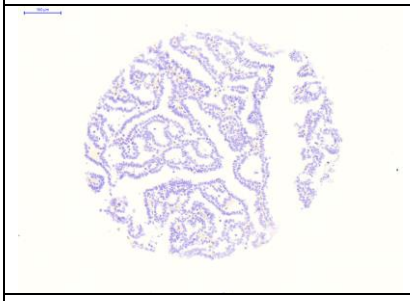
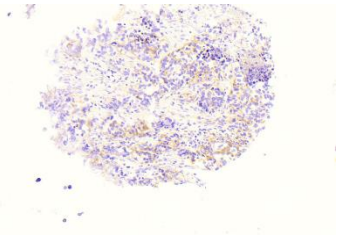
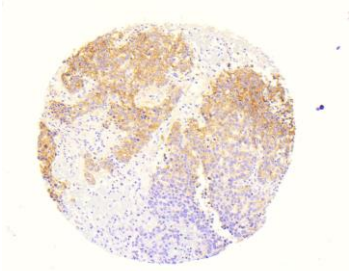
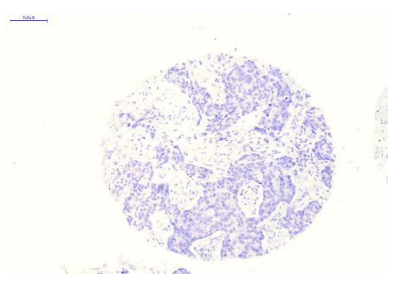
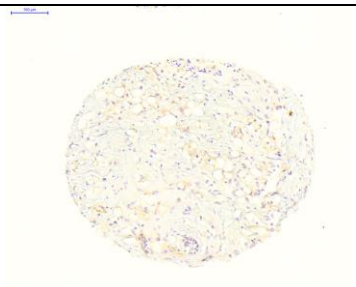
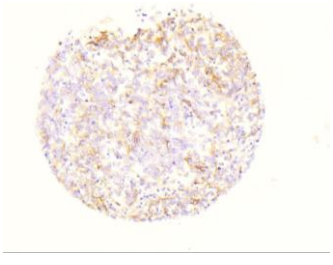
B. Progesterone receptor.

Negative (<1%)	Low Positive (1-10%)	Positive (>10%)
		
		
		
		

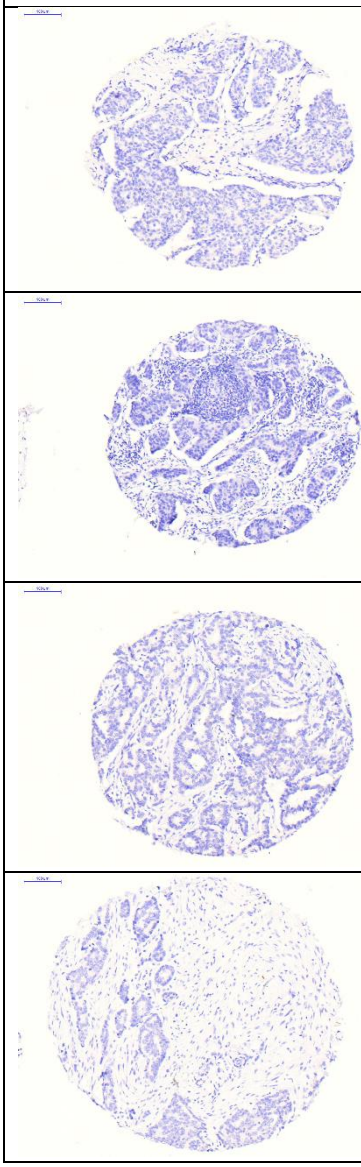
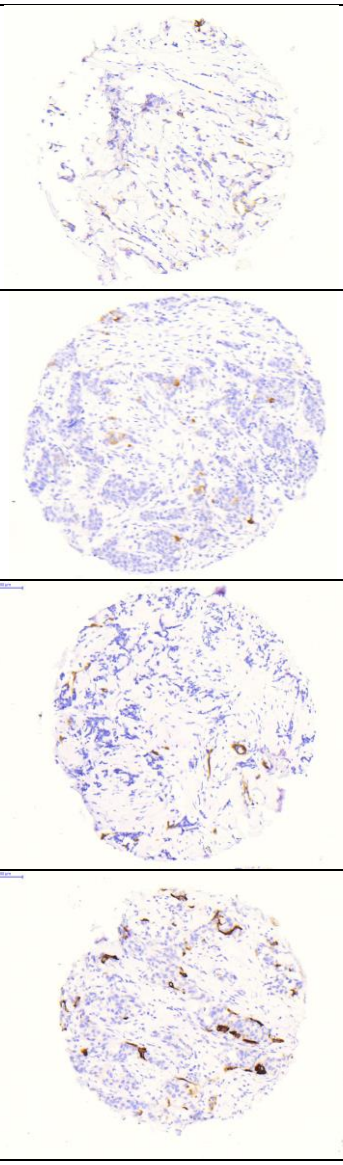
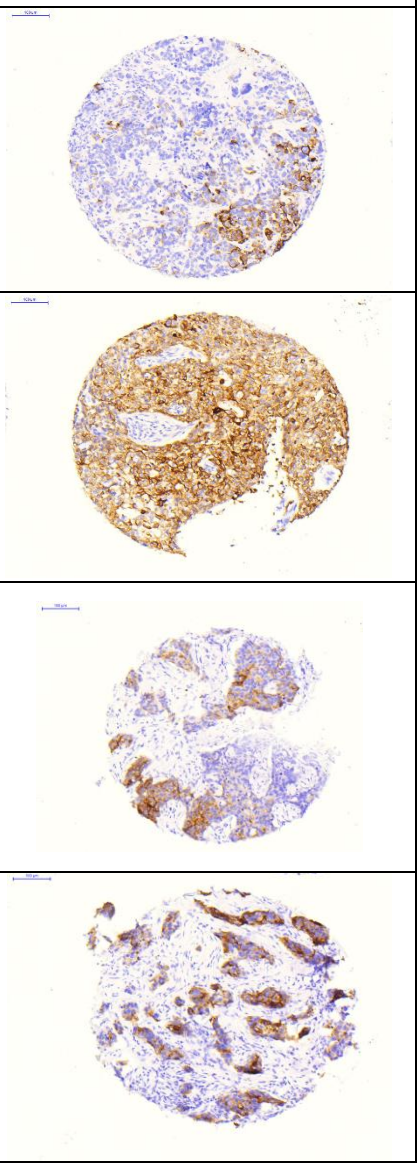
C. Human epidermal growth factor receptor 2.

Negative (<1%)	**Low Positive (1-10%)**	**Positive (>10%)**
		
		
		
		

D. Epithelial growth factor receptor.

Negative (<1%)	Low Positive (1-10%)	Positive (>10%)
		
		
		
		

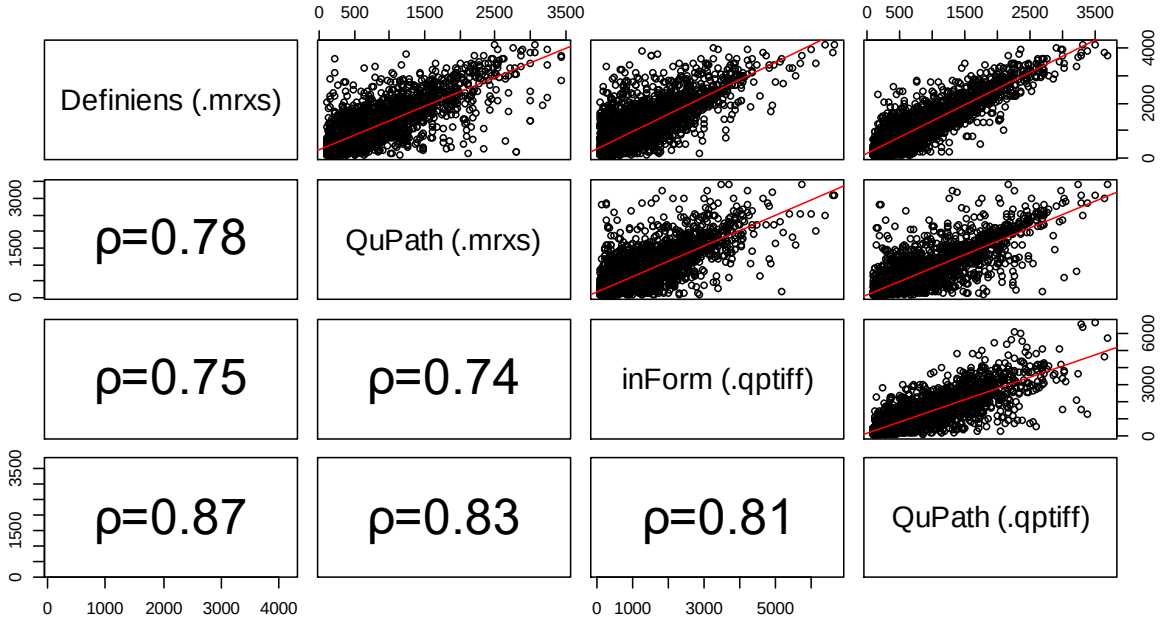
E. Cytokeratin 5/6.

Negative (<1%)	Low Positive (1-10%)	Positive (>10%)
		

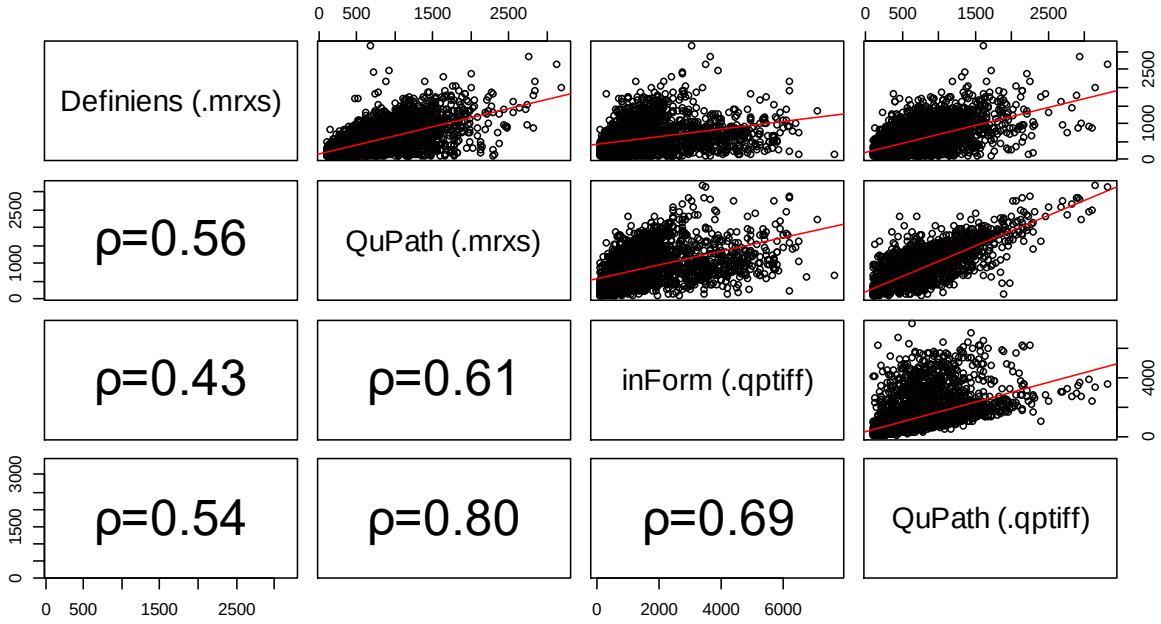
Supplementary 3. Parameters used in training the immunohistochemistry thresholds for each software.

	Definiens (.mrxs)	inForm (.qptiff)	QuPath (.mrxs)	QuPath (.qptiff)
Estrogen receptor				
Training TMA	359	359	359	359
IHC minimum detection value	0.20	0.10	0.20	0.20
H&E threshold	0.10	0.22	0.10	0.10
Progesterone receptor				
Training TMA	359	359	359	359
IHC minimum detection value	0.15	0.12	0.15	0.15
H&E threshold	0.07	0.28	0.10	0.08
Human epidermal growth factor receptor 2				
Training TMA	359	359	359	359
IHC minimum detection value	0.15	0.09	0.15	0.15
H&E threshold	0.07	0.16	0.07	0.05
Epithelial growth factor receptor				
Training TMA	359	359	359	359
IHC minimum detection value	0.08	0.06	0.12	0.12
H&E threshold	0.04	0.20	0.07	0.05
Cytokeratin 5/6				
Training TMA	359	359	359	359
IHC minimum detection value	0.08	0.06	0.06	0.06
H&E threshold	0.07	0.16	0.07	0.07

Supplementary 4A. Pair-wise correlation of the number of detected tumor cells using Spearman's ρ.



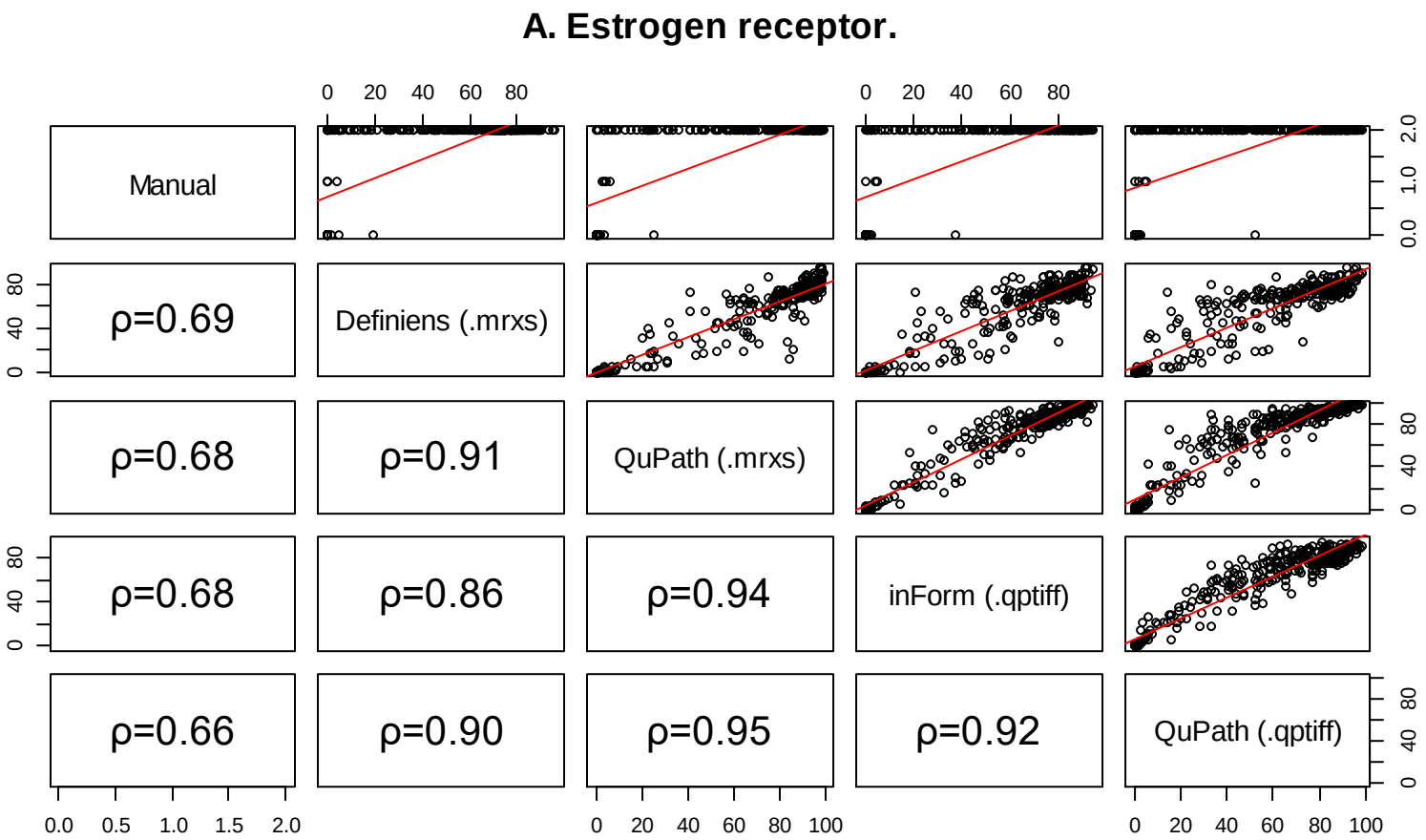
Supplementary 4B. Pair-wise correlation of the number of detected stromal cells using Spearman's ρ.



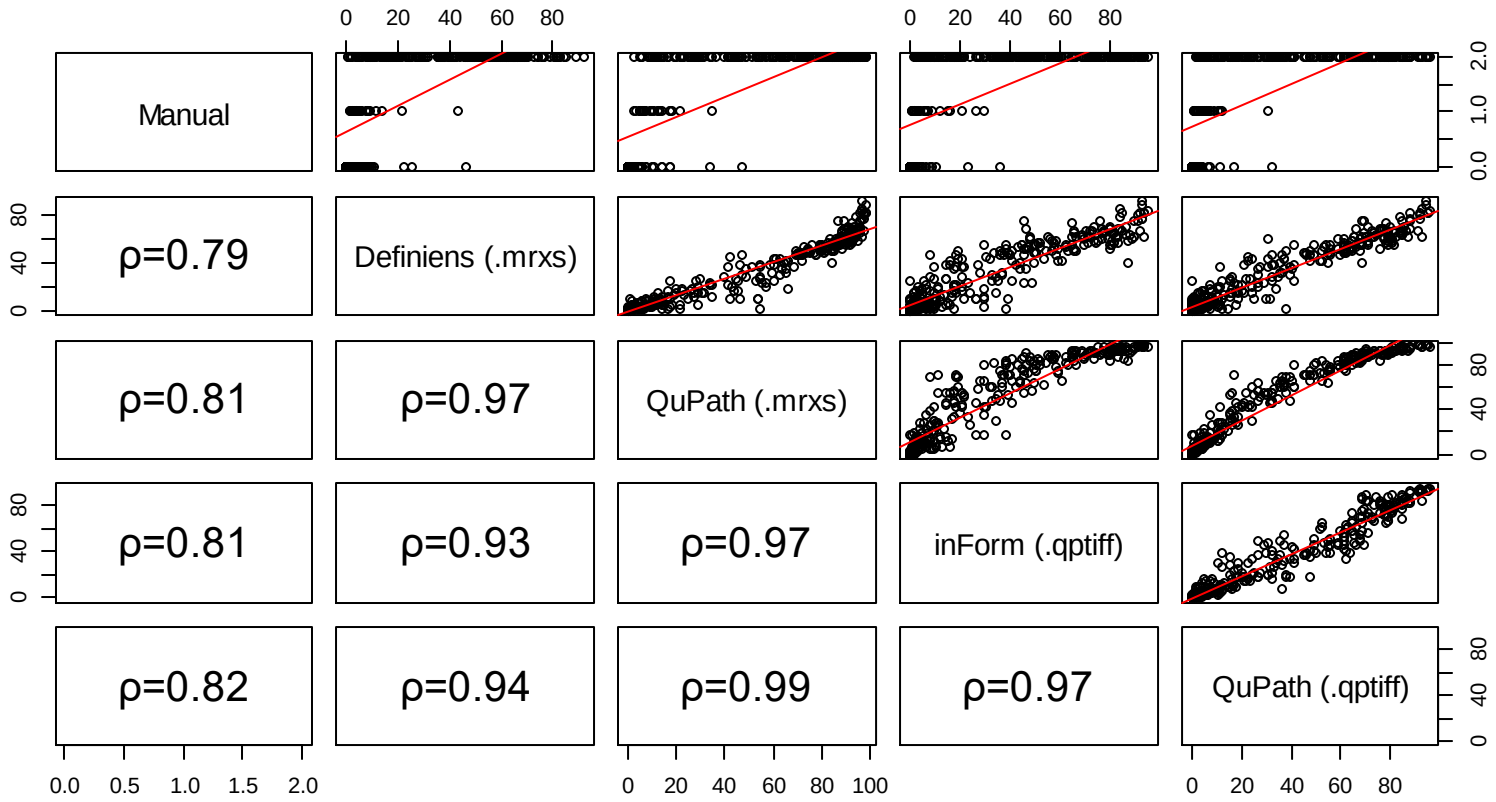
Supplementary 5. Intra-rater reliability of each software to classify tumor or stromal cells across 2 to 15 cores on 5 tissue microarrays pertaining to each woman, assessed using intraclass correlation (ICC).

	ICC (95% confidence interval)
Number of cells classified as tumor	
Definiens (.mrxs)	0.92 (0.91-0.93)
QuPath (.mrxs)	0.94 (0.93-0.95)
inForm (.qptiff)	0.92 (0.91-0.94)
QuPath (.qptiff)	0.91 (0.89-0.92)
Number of cells classified as stromal	
Definiens (.mrxs)	0.80 (0.75-0.83)
QuPath (.mrxs)	0.94 (0.93-0.95)
inForm (.qptiff)	0.79 (0.75-0.83)
QuPath (.qptiff)	0.91 (0.90-0.92)

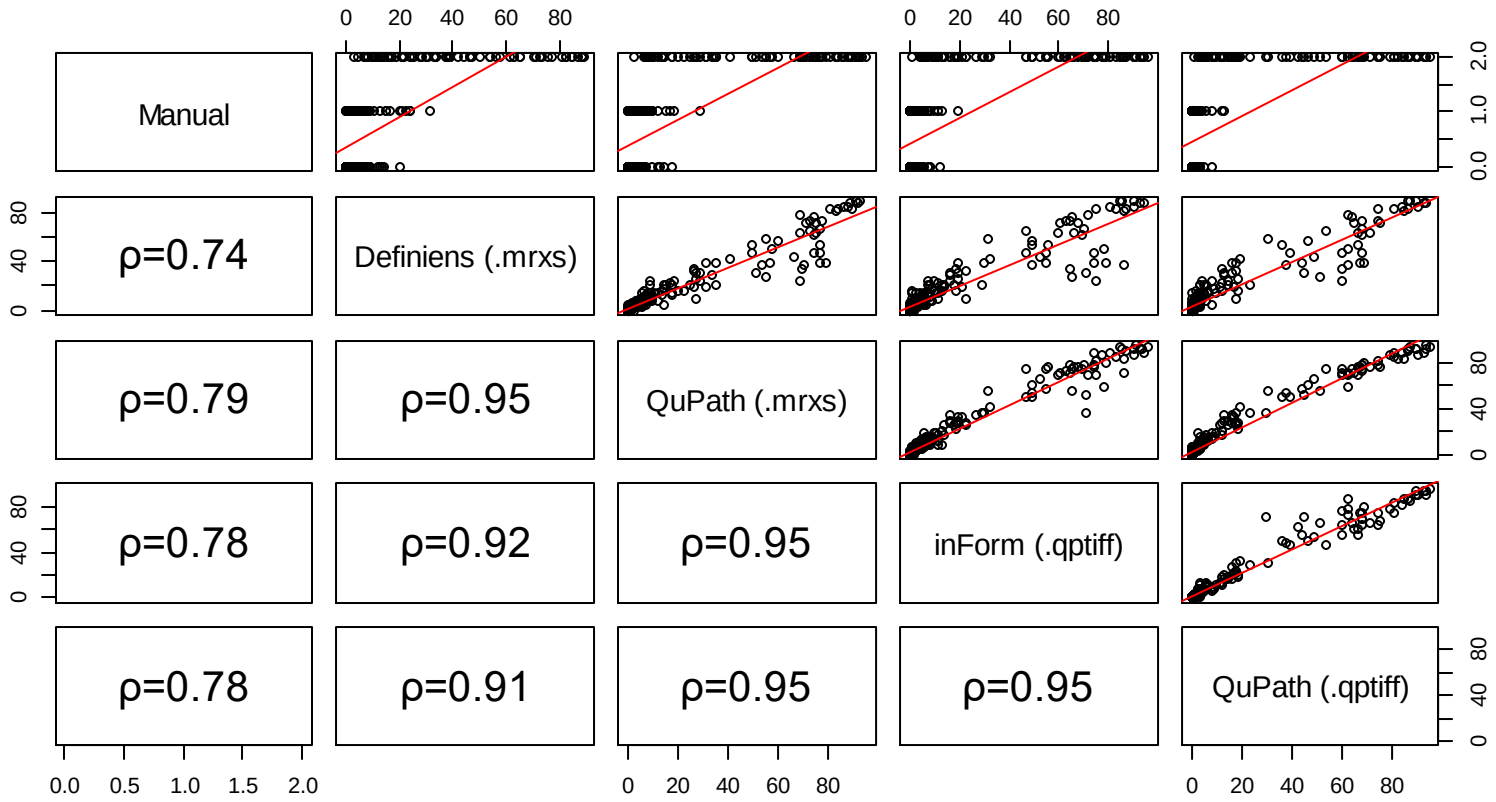
Supplementary 6. Spearman's correlations between manual and software-derived scores for each marker at case level whereby each woman is represented by 1-3 cores. For manual assessment, each case was represented by the highest ordinal category across its cores (negative, low positive, and positive). For software applications, a weighted average percent positive was calculated to represent each case.



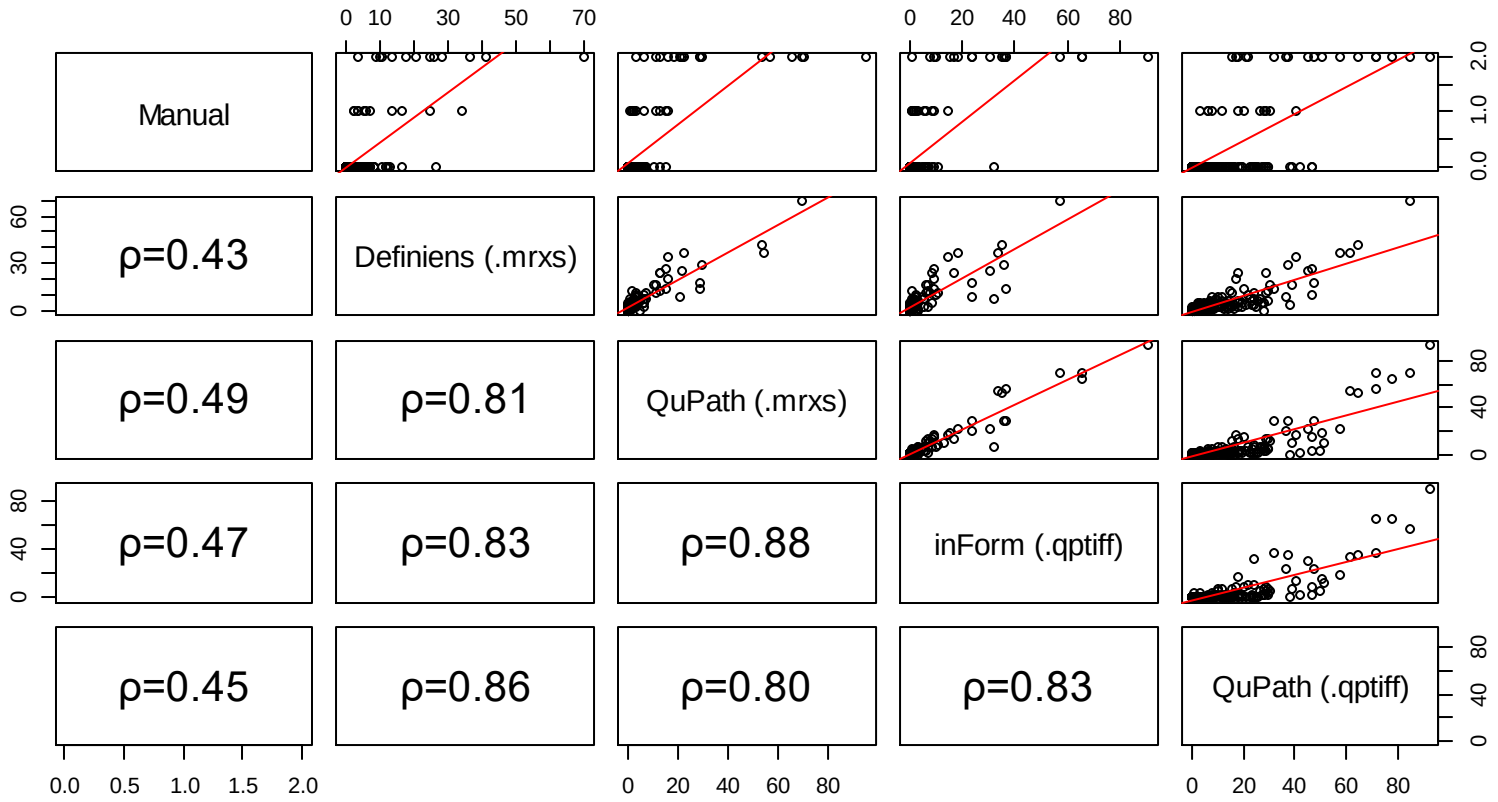
B. Progesterone receptor.



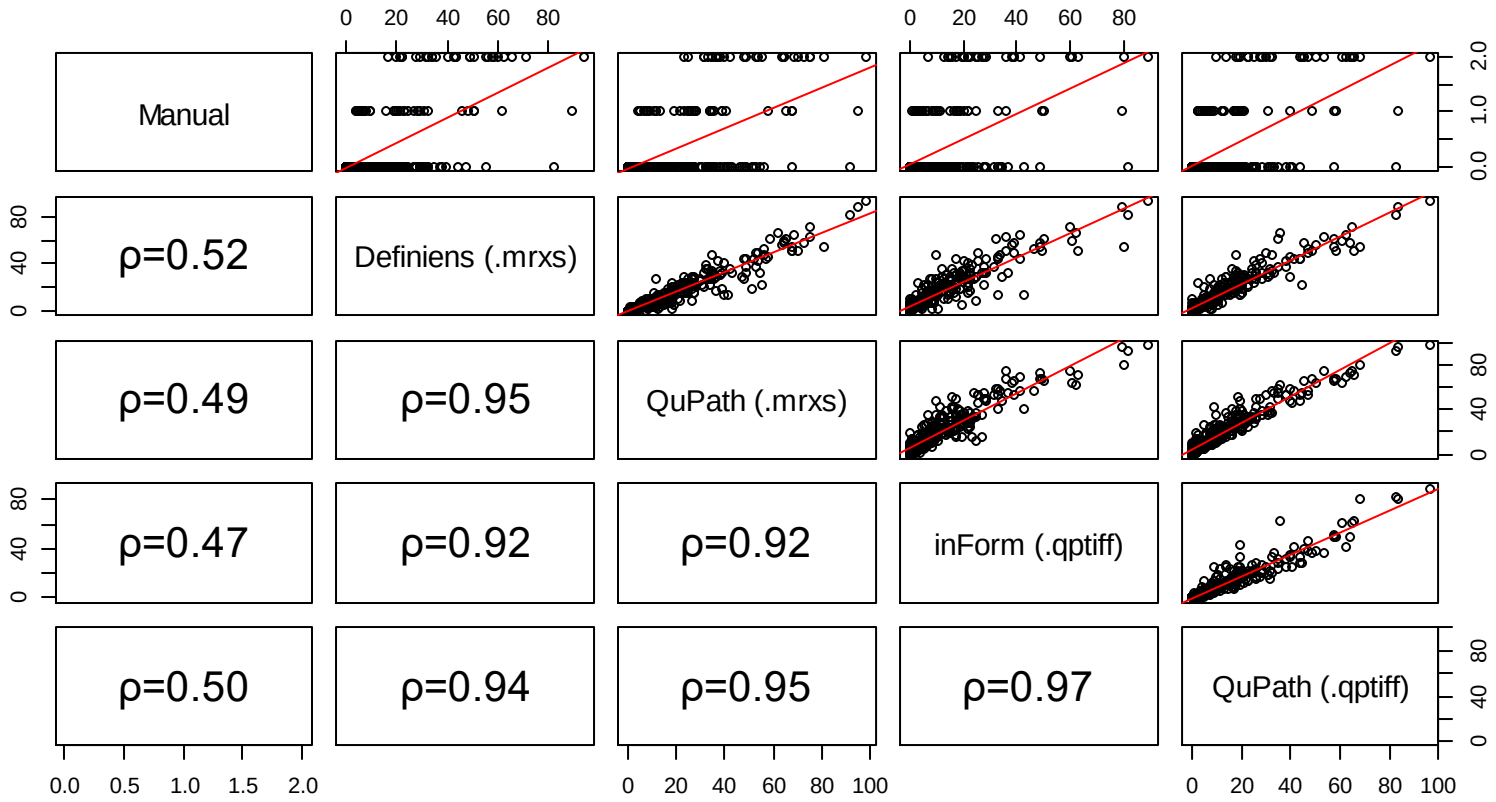
C. Human epidermal growth factor receptor 2.



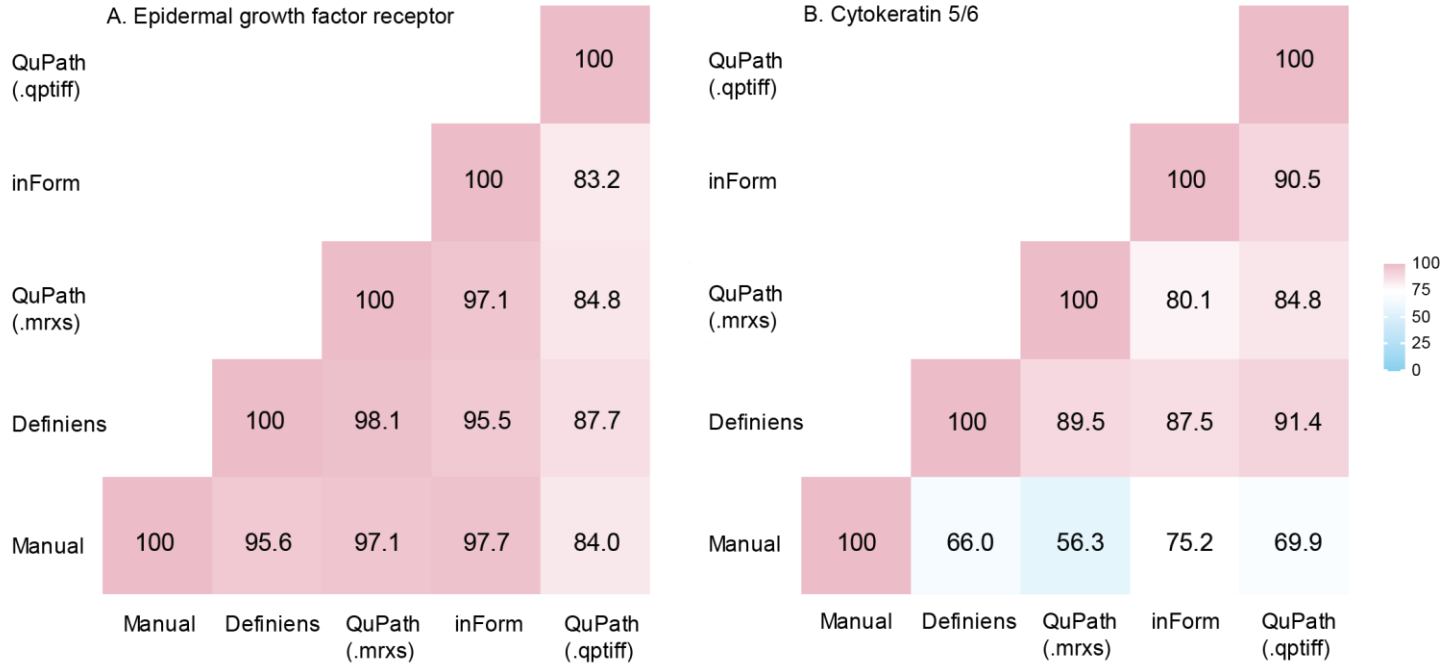
D. Epidermal growth factor receptor.



E. Cytokeratin 5/6.



Supplementary 7. Case-level percentage agreements between manual and software-derived scores for **A**. epidermal growth factor receptor and **B.** cytokeratin 5/6. For manual assessment, each case was represented by the highest ordinal category across its cores and dichotomized into negative/low positive ≤10% or positive (>10%). For software applications, the weighted average percent positive was calculated to represent each case and also dichotomized into negative/low positive ≤10% or positive (>10%).



Supplementary 8. Software ranking based on highest ρ or percentage agreement when benchmarked against manual scores.

	Definiens	QuPath (.mrxs)	inForm	QuPath (.qptiff)	Overall Highest Ranked Software
Core-level comparison—Spearman's ρ (Figure 3)					
ER	3	1	1	2	
PR	3	1	2	1	
HER2	3	1	2	1	
EGFR	3	1	1	2	
CK5/6	1	3	4	2	
					QuPath
Case-level comparison—Spearman's ρ (Supplementary 6)					
ER	1	2	2	3	
PR	3	2	2	1	
HER2	3	1	2	2	
EGFR	4	1	2	3	
CK5/6	1	3	4	2	
					QuPath
Case-level comparison—percentage agreements (Figure 4 for ER, PR and HER 2; Supplementary 7 for EGFR and CK5/6)					
ER	2	3	1	3	
PR	4	2	3	1	
HER2	4	2	3	1	
EGFR	3	2	1	4	
CK5/6	3	4	1	2	
					inForm®
Case-level comparison—area under receiver-operator characteristics curves (Figure 5)					
ER	2	1	1	2	
PR	3	1	2	1	
HER2	2	1	1	1	
EGFR	2	1	2	3	
CK5/6	1	2	3	2	
					QuPath

Estrogen receptor, ER; progesterone receptor, PR; human epidermal growth factor receptor 2, HER2; epithelial growth factor receptor, EGFR; cytokeratin 5/6, CK5/6.