Contents

Supplementary Table S1. Characteristics of Ovarian Cancers and Benign Pelvic Masses.

See excel

Supplementary Table S2. Estimated coefficients for the combined model of the 7MetP plus ROMA.

	Intercept	ROMA	7MetP
Coefficients	-3.15	4.06	4.12

Supplementary Table S3. Predictive performance of quantified metabolites for distinguishing OvCa from BPM in the Training Set.

See excel

Supplementary Table S4. Individual predictive performance of selected metabolites in the Training Set.

Features	AUC (95% CI)	Sensitivity (95% CI) @ 99% specificity	Specificity (95% CI) @ 99% sensitivity
Diacetylspermine (DAS)	0.82 (0.76 - 0.88)	0.28 (0.05 - 0.50)	0.07 (0.03 - 0.34)
N-acetylneuraminate (NANA)	0.65 (0.58 - 0.73)	0.03 (0.00 - 0.08)	0.05 (0.01 - 0.15)
N-acetyl-mannosamine (NAcMan)	0.58 (0.50 - 0.65)	0.05 (0.01 - 0.12)	0.04 (0.00 - 0.17)
N-acetyl-lactosamine (NAcLac)	0.55 (0.48 - 0.63)	0.03 (0.00 - 0.07)	0.04 (0.01 - 0.08)
Diacetylspermidine (DiAcSpmd)	0.67 (0.60 - 0.74)	0.03 (0.00 - 0.18)	0.05 (0.01 - 0.10)
N-(3-acetamidopropyl)pyrrolidin-2-one (N3AP)	0.56 (0.49 - 0.64)	0.05 (0.00 - 0.13)	0.03 (0.00 - 0.08)
Hydroxyisobutyrate (HBA)	0.71 (0.64 - 0.77)	0.08 (0.03 - 0.20)	0.16 (0.00 - 0.25)

Supplementary Table S5. Performance of the 7-marker metabolite panel (7MetP) for distinguishing OvCa cases from individuals with BPM in the Training Set, the independent Test Set, and the combined Training+Testing Specimen Set.

	Training Set		Testing Set		Combined Dataset	
	Early-Stage (n=39)	**All Stage (n=101)**	**Early-Stage (n=20)**	**All Stage (n=118)**	**Early-Stage (n=59)**	**All Stage (n=219)**
AUC	0.75 (0.66 - 0.85)	0.83 (0.78 - 0.89)	0.86 (0.76 - 0.95)	0.88 (0.82 - 0.93)	0.81 (0.76 - 0.86)	0.85 (0.81 - 0.88)
Sensitivity at 99% specificity	0.21 (0.08 - 0.40)	0.41 (0.17 - 0.50)	0.15 (0.00 - 0.35)	0.40 (0.31 - 0.62)	0.20 (0.11 - 0.28)	0.39 (0.20 - 0.47)
Sensitivity at 95% specificity	0.24 (0.11 - 0.47)	0.47 (0.35 - 0.66)	0.30 (0.05 - 0.65)	0.51 (0.33 - 0.75)	0.25 (0.15 - 0.44)	0.46 (0.38 - 0.62)
Specificity at 99% sensitivity	0.54 (0.40 - 0.78)	0.16 (0.06 - 0.30)	0.36 (0.24 - 0.80)	0.11 (0.02 - 0.42)	0.13 (0.08 - 0.35)	0.12 (0.05 - 0.28)
Specificity at 95% sensitivity	0.10 (0.05 - 0.28)	0.29 (0.14 - 0.47)	0.55 (0.27 - 0.87)	0.40 (0.07 - 0.62)	0.32 (0.19 - 0.50)	0.31 (0.20 - 0.44)

Supplementary Table S6. Stability check of the deep learning model (DLM) in the Training Set

	Scenario #1	Scenario #2	Scenario #3	Scenario #4	Scenario #5
	Early-stage OvCa cases and	All stage OvCa cases and BPM	500 random samples with	173 random samples with	Late-stage OvCa cases and BPM
N0 (# of BPM)	134	134	406	143	134
N1 (# of cases)	39	100	94	30	61
AUC	0.75 (0.66 - 0.85)	0.83 (0.78 - 0.89)	0.76 (0.70 - 0.82)	0.76 (0.68 - 0.84)	0.88 (0.83 - 0.94)
Sensitivity at 95% specificity	0.26 (0.11 - 0.47)	0.46 (0.36 - 0.64)	0.35 (0.17 - 0.56)	0.20 (0.10 - 0.45)	0.61 (0.47 - 0.78)
Specificity at 95% sensitivity	0.16 (0.05 - 0.47)	0.29 (0.16 - 0.49)	0.18 (0.12 - 0.28)	0.34 (0.14 - 0.50)	0.33 (0.22 - 0.68)

Supplementary Table S7. Predictive performance of the 7MetP for distinguishing OvCa cases stratified into serous and non-serous from BPM in the Training Set.

	Serous carcinoma		Non-serous carcinoma	
	Early-Stage (n=13)	**All Stage (n=69)**	**Early-Stage (n=25)**	**All Stage (n=31)**
AUC	0.69 (0.52 - 0.86)	0.85 (0.79 - 0.91)	0.79 (0.68 - 0.89)	0.80 (0.71 - 0.89)
Sensitivity at 99% specificity	0.23 (0.00 - 0.46)	0.48 (0.17 - 0.61)	0.20 (0.04 - 0.36)	0.23 (0.10 - 0.39)
Sensitivity at 95% specificity	0.23 (0.00 - 0.54)	0.55 (0.39 - 0.71)	0.24 (0.08 - 0.44)	0.29 (0.13 - 0.55)
Specificity at 99% sensitivity	0.10 (0.05 - 0.48)	0.11 (0.05 - 0.28)	0.19 (0.11 - 0.44)	0.19 (0.11 - 0.40)
Specificity at 95% sensitivity	0.11 (0.06 - 0.51)	0.26 (0.14 - 0.51)	0.33 (0.13 - 0.59)	0.30 (0.12 - 0.52)

Supplementary Table S8. Performance estimates of ROMA and the combined 7MetP+ROMA model for all OvCa in the combined specimen set.

All OvCa (n=219) vs BPM (N=190)					
		ROMA	**ROMA + 7MetP**	**Difference**	***P***
	AUC (95% CI)	0.94 (0.92 - 0.95)	0.94 (0.93 - 0.96)	0.01 (0.00 to 0.01)	.001
At 11.4% risk threshold for premenopausal and 29.9% for postmenopausal (same risk as ROMA)	**Sensitivity**	0.92 (0.89 - 0.94)	0.90 (0.87 - 0.92)	-0.02 (-0.04 to 0.00)	.01
	Specificity	0.78 (0.75 - 0.82)	0.89 (0.86 - 0.91)	0.11 (0.08 to 0.14)	**< .001**
	PPV	0.83 (0.80 - 0.86)	0.90 (0.88 - 0.92)	0.07 (0.05 to 0.10)	**< .001**
	NPV	0.89 (0.86 - 0.92)	0.88 (0.86 - 0.91)	-0.01 (-0.03 to 0.01)	.21

Abbreviations: PPV: positive predictive value; NPV: negative predictive value. P-values for comparison of AUCs represent likelihood ratio tests. Risk threshold corresponding to 11.4% in premenopausal women and 29.9% for postmenopausal were chosen based on reported findings from Ortiz-Munoz and colleagues.[1] 1-sided P-values are reported as we expect that the combined 7MetP+ROMA will yield improved performance estimates compared to ROMA alone.

References

1. Ortiz-Muñoz B, Aznar-Oroval E, García AG, et al: HE4, Ca125 and ROMA algorithm for differential diagnosis between benign gynaecological diseases and ovarian cancer. Tumor Biology 35:7249-7258, 2014