Table of Contents

M1. RNA Sequencing Data Generated by the GYN-COE

RNA sequencing (RNAseq) data, tumor grade and myometrial invasion data were repurposed from a legacy research project from the Gynecologic Cancer Center of Excellence (GYN-COE) (Bateman NW et al. *Cancer* 2017;123(20):4004-4012). Briefly, frozen primary tumors from 15 patients with stage I, IIIC or IV endometrioid endometrial carcinomas excised at the time of the hysterectomy were sectioned at 10 µm in thickness and collected on polyethylene naphthalate membrane slides. Laser capture microdissection (LMD) was performed to enrich for tumor cellularity prior to RNA extraction. Total RNA was isolated (RNeasy Micro Kit, Qiagen, Valencia, CA) according manufacturer's instructions, including DNase treatment, and quantified spectrophotometrically (NanoDrop™ 2000 Spectrophotometer, ThermoFisher Scientific, Waltham, MA) and quality checked by electrophoresis (2100 Bioanalyzer, Agilent Technologies, Santa Clara, CA using the RNA 6000 Pico Chip Kit, Agilent Technologies, Santa Clara, CA). The External RNA Control Consortium (ERCC) spike-in control mix (Ambion, Foster City, CA) was added to 500 ng of total RNA, which was subjected to mRNA enrichment (mRNA Catcher™ PLUS plate, ThermoFisher Scientific, Waltham, MA). Due to low input, next generation sequencing (NGS) libraries were constructed from the entire sample of poly(A) RNA (Library Builder™ Whole Transcriptome Core Kit for 5500 Genetic Analysis Systems, ThermoFisher Scientific, Waltham, MA) according to manufacturer's protocol using an automated fashion (AB Library Builder™ System, ThermoFisher Scientific, Waltham, MA). The NGS libraries were barcoded (SOLiD® RNA Barcoding Kit, ThermoFisher Scientific, Waltham, MA) and amplified for 18 cycles. Each sample's NGS library size and yield were assessed by spectrophotometry (NanoDrop Spectrophotometer) and electrophoresis (Agilent® 2100 Bioanalyzer). The barcoded NGS libraries were pooled in groups of 8 for templated bead preparation according to the manufacturer's instructions

(SOLiD™ EZ Bead® Emulsifier, Amplifier, and Enricher, ThermoFisher Scientific, Waltham, MA). Two full NGS flow chips (SOLiD 6-lane FlowChips, ThermoFisher Scientific, Waltham, MA) were prepared for NGS using 350 million beads per lane and loaded onto the next generation sequencer (5500xl Series Genetic Analyzer, ThermoFisher Scientific, Waltham, MA). Paired-end sequencing was performed with both forward (75bp) and reverse (35bp) primers. After primary data acquisition, the eXtensible SeQuence (.xsq) files were imported into the NGS vendor's analysis software (LifeScope Genomic Analysis Software 2.5.1, ThermoFisher Scientific, Waltham, MA). LifeScope software Whole Transcriptome Analysis (WTA) pipeline was applied to grouped, .xsq files followed by alignment to the human reference genome (hg19). As the result of the color-space alignment, Binary Alignment sequence Map (BAM) files were generated. The BAM files were further used to filter, count, generate Reads Per Kilobase of exon model per Million mapped reads (RPKM), and annotate how many exons were expressed in a transcript.

M2. Microarray Data Generated by the Gynecologic Cancer Center of Excellence

Hybridization-based microarray data were generated de novo for 64 Gynecologic Oncology Group (GOG) patients in Training-1 cohort for this investigation. In addition, Affymetrix microarray data were repurposed for 81 patients in Validation-2 from legacy research projects from the GYN-COE (Maxwell GL *et al. Gynecol Oncol*. 2013;130(1):169-73; Risinger JI *et al. Front Oncol*. 2013;3:139) for this study. Frozen primary tumors excised at the time of the hysterectomy were embedded in optimal cutting temperature (OCT) compound and thin sections (8 µm) were cut and mounted on glass slides for pathology review. A board-certified pathologist centrally reviewed hematoxylin and eosin stained sections from GOG (NR) or the GYN-COE (CZ), respectively, to select cases with at least 10% tumor cellularity, endometrioid histotype, and ≤50% necrosis. Malignant epithelial cells were enriched from tissue specimens using LMD or macrodissection/scraping. All tumor samples processed by the GYN-COE underwent LMD except for a subset of the legacy research samples with highly purified, dense tumor cellularity that were macrodissected / scraped. When LMD was required, tissue sections were prepared on PEN membrane slides, crosslinked for 30 min using a UV crosslinker (Spectronics Corporation, Westbury, NY), stained in Mayer's hematoxylin, air dried for 24 h, and underwent laser microdissection (LMD 6500, Leica Microsystems, Wetzlar, Germany). Staining of frozen tissue sections was required for successful collections of regions of interest and enrichment of tumor cellularity, and was shown in controlled pilot experiment to not have a negative impact on down-stream processing or assessment of transcript expression. Malignant epithelial cells were collected into RLT Buffer (Qiagen, Valencia, CA) plus 1% β-mercaptoethanol and stored at -80°C. Total RNA was isolated (RNeasy Micro Kit) and spectrophotometrically quantified (Nanodrop 1000, ThermoFischer Scientific, Inc., Wilmington, DE). The RNA Integrity Number (RIN, Bioanalyzer 2100, Agilent Technologies, Santa Clara,

CA) was utilized as a surrogate of RNA (and tissue specimen) integrity. Only those cases with a RIN > 5 were included in the transcriptomic analysis. Cases were excluded if they had poor quality RNA (RIN<5) or a high GAPDH 3':5' ratio (>10). One hundred ng of total RNA was amplified (Two-Cycle cDNA Synthesis Kit, Affymetrix Inc., Santa Clara, CA) from which biotinylated aRNA was prepared (Affymetrix Labeling Kit, Affymetrix Inc). Approximately 5 μg of total RNA from each sample was labeled using the high yield transcript labeling kit (ENZO) and labeled RNAs were hybridized to U133A Plus 2.0 microarrays (Affymetrix Inc), washed, and scanned according to the manufacturers specifications (Affymetrix Inc., 2001).

M3. Transcript Selection, and Classifier Development and Evaluation

RNAseq- and microarray-derived transcripts were screened for their association with nodal and distant metastasis in the Training-1 and Training-2 Cohorts, respectively, and then a multi-transcript classifier was developed in these two Cohorts.

- Univariate logistic regression modeling of nodal and distant metastasis was performed using R with individual transcripts included as the explanatory variable. We screened 17,265 transcripts to identify 1,630 transcripts associated with metastasis with $p<0.05$ in the 75 Training-1 patients with RNAseq data (**Supplement S1**).
- Randomized subsampling was implemented using the sample function in R/Bioconductor (Version 2.13) to create 100 subsamples of Training-1 each with 60 of the 75 cases. Univariate logistic regression modeling was then implemented with each of the 1,630 transcripts included as the explanatory variable in the 100 subsamples of Training-1. There were 311 transcripts associated with metastasis with $p<0.05$ in at least 80 of the 100 randomly selects subsamples of Training-1 cases (**Supplement S2**) that were selected for further consideration.
- Mapping was performed that identified Affymetrix Plus 2.0 probesets for 268 of the 311 candidate transcripts. There were 43 transcripts were dropped from further consideration. R was used to implement univariate logistic regression modeling of metastasis with each of the 268 transcripts included as the explanatory variable. There were 33 transcripts associated with metastasis with $p<0.05$ in the 64 Training-2 cases (**Supplement S3)**.
- Results for univariate logistic regression modeling in the 33 candidate transcripts in 75 Training-1 and 64 Training-2 cases were compared and 23 of these transcripts had a consistent relationship with metastasis in both cohorts. Eleven transcripts had a positive

coefficient (direct relationship with risk of metastasis) and 12 transcripts had a negative coefficient (inverse relationship with risk of metastasis) in both Training-1 and Training-2 (**Supplement S4**).

- A four-step selection process was implemented to screen the unique combinations of the 23 genes using WEKA version 3.6.12. The goal was to select the top metastasis classification model in the 75 Training-1 cases, the 64 Training-2 cases and 10-fold cross validation of both Training-1 and -2. WEKA calculated the weighted average for true positive rate (TPR), false positive rate (FPR), receiver operator characteristics (ROC) curve area and precision-recall curve (PRC) area for each of the unique combinations of transcripts which allowed us to rank and select the top multi-transcript classifier for nodal and distant metastasis in EEC patients.
    - During Step 1, the top 39,788 models were selected from the pool of 8,388,607 ($1\text{-}2^{23}$) models with a TPR >0.8, an area under the ROC curve >0.8, and an FPR <0.2 in all 75 Training-1 cases.
    - During Step 2, the top 27 models were selected based on the largest TPR in the 10-fold cross-validation (CV) of Training-2 cases.
    - During Step 3, the top 7 models were selected based on an ROC curve area >0.8 and PRC area >0.8 in the 10-fold CV of Training-2.
    - During Step 4, the top model was then selected based on largest TPR in Training-1, 10-fold CV of Training-1, Training-2 and 10-fold CV of Training-2
- The top model included BDNFOS also known as BDNF-AS (Gene ID 497258), APOL4 (Gene ID 80832), RSRC1 (Gene ID 51319), PDLIM3 (Gene ID 27295), TBRG1 (Gene ID 84897), ZNF596 (Gene ID 169270), and MLLT10 (Gene ID 8028), and was

integrated into the MS7 metastasis score as follows using SAS. Optimal platform-centric cut points were determined using Youden index.

- The algorithm for MS7 using RNAseq data = (TBRG1 x -1.6575) + (BDNFOS x -0.5445) + (APOL4 x -0.4643) + (ZNF596 x 0.2676) + (PDLIM3 x 0.6627) + (RSRC1 x 0.8427) + (MLLT10 x 1.0058). Fixed coefficients from modeling in Training-1. Optimal cut point for the RNAseq-based MS7 score was -2.00356.
- The algorithm for MS7 using microarray data = (TBRG1 x -1.5770) + (BDNFOS x -1.9217) + (APOL4 x -0.6422) + (ZNF596 x -0.7243) + (PDLIM3 x 1.0483) + (RSRC1 x 1.6294) + (MLLT10 x 1.7836). Fixed coefficients from Training-2. The optimal cut point for the microarray-based MS7 was -4.25324.

The MS7 Classifier was then evaluated with select clinical characteristics and molecular features in the Training-1 Cohort, and then with grade 3 disease (G3) and ≥50% myometrial invasion (MI) in the Validation-1, Training-2 and Validation-2 cohorts. Training-1 and Validation-1 Cohorts utilized RNAseq-based transcripts for the MS7 classifier whereas the Training-2 and Validation-2 utilized microarray-based transcripts for the MS7 classifier.

- Logistic regression modeling of nodal and distant metastasis and ROC curve analyses with calculations for area under the curve (AUC), standard error (SE), and 95% confidence intervals (CIs) were implemented were performed using SAS to compare the MS7 classifier with other clinical and/or molecular predictors, and to evaluate candidate companion diagnostic models using the MS7 classifier ± G3 ± MI.
- Classification performance was assessed and sensitivity (SN), specificity (SP), positive predictive value (PPV) and negative predictive value (NPV) were calculated in the four Cohorts and in the merged Validation-1/-2 Cohorts for companion diagnostic models based on MS7 alone, MS7+G3 or MS7+MI using SAS. MS7 score was categorized

using the platform-centric cut point. A negative test for the three companion diagnosis models was indicated by a low MS7 score for the MS7 model, a low MS7 score and grade 1 or 2 disease for the MS7+G3 model, or a low MS7 score and <50% MI for the MS7+MI model, respectively. PPV and NPV were estimated using an estimated prevalence of endometrioid endometrial cancer patients having vs. not having nodal or distant metastasis of 15% vs. 85%. The 95% CI for the PPV or NPV estimated using bootstrapping with 3000 repeats.

M4. Power Analysis and SAS Code

Power Analysis for Comparing the Area Under the Curve (AUC) for Nodal and Distant Metastasis in the Same Endometrioid Endometrial Cancer Cases: The AUC of the gold standard, grade 3 disease, was calculated both in SEER and The Cancer Genome Atlas (TCGA) cases in the Training-1 cohort to evaluate the stability of this estimate in a large vs. relatively sample size, respectively. A power analysis was then performed using PASS[13] (NCSS Statistical Software, Kaysville, Utah) in the TCGA Training-1 cohort with 46 cases with stage IIIC or IV disease in the positive group and 29 cases with stage I disease in the negative group. The AUC for the null hypothesis (grade 3 disease) was 0.656 and the AUC for the alternate hypothesis (multi-gene algorithm) was varied from 0.7 to 0.95 (see insert below). This sample achieved 80% power to detect a difference of 0.1406 between a diagnostic test based grade 3 disease and another based on a continuous multi-gene algorithm using a one-sided z-test at a significance level of 0.050 with AUC computed between false positive rates of 0.00 and 1.00, and the ratio of the standard deviation of the responses in the negative group to the standard deviation of the responses in the positive group is 1.00. A similar type of power analysis was also performed for evaluation of ROCs in GOG-210 with similar findings.

N- (stage I)	N+ (stage IIIC/IV)	N (total)	AUC1 Grade3	AUC2 Multi-Gene Algorithm	AUC difference	Power with alpha at 0.05
29	46	75	0.656	0.7000	0.0440	0.18285
29	46	75	0.656	0.7500	0.0940	0.49062
29	46	75	0.656	0.7966	0.1406	0.80022
29	46	75	0.656	0.8200	0.1640	0.90194
29	46	75	0.656	0.8500	0.1940	0.96997
29	46	75	0.656	0.9000	0.2440	0.99785
29	46	75	0.656	0.9500	0.2940	0.99990

Hanley JA, McNeil BJ. A method of comparing the areas under receiver operating characteristic curves derived from the same cases. Radiology 1983;148:839-43.
Obuchowski N, McClish D. Sample size determination for diagnostic accuracy studies involving binormal ROC curve indices. Statistics in Medicine 1997;16:1529-42.

/* SAS Codes for Evaluation of MS7 Signature in Prediction of Metastasis in Training and Validation Datasets ***/**

/* Datasets ***/**
/*** All patients were stage IA/IB (no metastasis) vs. stage IIIC/IV (metastasis) ***/

/*** RNA-seq platform ***/
/*** Training 1: selected TCGA RNA-seq data, n=75***/
/*** Validation 1: TCGA RNA-seq stage IA/IB and IIIC/IV cases not used for training 1
and excluding GOG submitted cases (n=230) plus 15 cases with RNA-seq data from GYN-COE; total N=245 ***/

/*** Affymetrix array platform***/
/*** Training 2: selected GOG 210 Microarray affymetrix data, n=64 ***/
/*** Validation 2: Duke Microarray affymetrix data, n=81 ***/

/*/**
/* Search for Optimal Cutoff from Traning Datasets based on Yonden Index ****/**
/* Adapt SAS Macro Program Developed by Mithat Gönen, (MSKCC) ********/**
/* Reference: Analyzing Receiver Operating Characteristics Curves with SAS ***/**
/*/**

/*** For RNA-seq Data from Training Dataset 1 (TCGA75) ***/

data TCGA75;
set TCGA75;

/*** Genetic score based on 7 genes from discovery stage***/
MS7 =
BDNFOS_497258 * (**-0.5445**) +
APOL4_80832 * (**-0.4643**) +
RSRC1_51319 * (**0.8427**) +
PDLIM3_27295 * (**0.6627**) +
TBRG1_84897 * (**-1.6575**) +
ZNF596_169270 * (**0.2676**) +
MLLT10_8028 * (**1.0058**);
run;

%inc 'D:\WHIRC\Metastasis\Program\rocplot.sas';
%***rocplot*** (TCGA75,MS7,stgrp,anno=**3**,tlist=**3 4 5 6**,round=**0.1**,optimal=**-9**);
data CUT7;
SET _plroc;
YJ=TPR+TNR-**1**;
RUN;

PROC SORT DATA=CUT7;
BY DESENDING YJ;
RUN;

PROC PRINT DATA=CUT7 (OBS=**1**);
var _x_;
RUN;

/*** MS7=-2.00356 defined as optimal cut-off for RNA-seq data using Youden index ***/

/*** For Microarray Data from Training Dataset 2 (GOG64) ***/

data GOG64;
set GOG64;

/*** Genetic score based on 7 genes from discovery stage***/
MS7=
BDNFOS_497258*(**-1.9217**) +
APOL4_80832 * (**-0.6422**) +
RSRC1_51319 *(**1.6294**) +
PDLIM3_27295 *(**1.0483**) +
TBRG1_84897 *(**-1.5770**) +
ZNF596_169270 *(**-0.7243**) +
MLLT10_8028 *(**1.7836**);
run;

```
%rocplot (GOG64,MS7,stgrp,anno=3,tlist=3 4 5 6,round=0.1,optimal=-9);

data CUT7; SET
_plroc;
YJ=TPR+TNR-1;
RUN;

PROC SORT DATA=CUT7; BY
DESENDING YJ;
RUN;

PROC PRINT DATA=CUT7 (OBS=1);
var _x_;
RUN;


/*************************************************************************/
/*** Evaluation of MS7 and Other Clinical Factors in Prediction of Metastasis Status ***/
/*** In Training and Validation Datasets (based on ROC Analyses) ******************/
/*************************************************************************/


/***** TABLES *****/

/*** MS7 / Grade /MI in Prediction of Metastasis (ROC analysis) ***/

/*** Training 1: TCGA75 ***/
/*** RNA-seq data ***/

data TCGA75; set TCGA75;
/*** Genetic score based on 7 genes from
discovery stage***/ MS7 =
BDNFOS_497258 * (-0.5445) + APOL4_80832
* (-0.4643) + RSRC1_51319 * (0.8427) +
PDLIM3_27295 * (0.6627) + TBRG1_84897 * (-
1.6575) + ZNF596_169270 * (0.2676) +
MLLT10_8028 * (1.0058);

/*** Optimal classification from traning 1 (based on Youden index) ***/
MS7GRP=MS7>=-2.00356;
run;

/*** MS7 / Grade in Prediction of Metastasis (ROC analysis) ***/

proc logistic data=TCGA75;
model stgrp (event='1')=grade3 ms7; roc
'MS7' ms7;
roc 'Grade' grade3;
roccontrast reference('Grade')/estimate;
run;

/*** MS7 / Grade / MI in Prediction of Metastasis ***/
/*** 3 patients with MI missing data excluded from analyses ***/
proc logistic data=TCGA75;
model stgrp (event='1')=grade3 myoinvcat1 ms7; roc
'MI' myoinvcat1;
roc 'Grade+MI' grade3 myoinvcat1; roc
'MI+Ms7' myoinvcat1 ms7;
roc 'Grade+MI+Ms7' grade3 myoinvcat1 ms7 ;
roccontrast /estimate=allpairs;
run;

/*** Validation 1: V245 ***/
/*** RNA-seq data ***/
data v245; set
v245;


/*** Genetic score based on 7 genes from discovery stage***/ MS7 =
BDNFOS_497258 * (-0.5445) + APOL4_80832
* (-0.4643) + RSRC1_51319 * (0.8427) +
```

```
PDLIM3_27295 * (0.6627) + TBRG1_84897 * (-
1.6575) + ZNF596_169270 * (0.2676) +
MLLT10_8028 * (1.0058);

/*** Optimal classification from traning 1 (based on Youden index) ***/
MS7GRP=MS7>=-2.00356;
run;

/*** MS7 / Grade in Prediction of Metastasis ***/
proc logistic data=v245;
model stgrp (event='1')=grade3 ms7; roc
'MS7' ms7;
roc 'Grade' grade3;
roc 'Grade+Ms7' grade3 ms7;
roccontrast /estimate=allpairs;
run;

/*** MS7 / Grade / MI in Prediction of Metastasis ***/
/*** 7 patients with MI missing data excluded from analyses ***/
proc logistic data=v245;
model stgrp (event='1')=grade3 myoinvcat1 ms7; roc
'Grade' grade3;
roc 'MI' myoinvcat1;
roc 'Grade+MI' grade3 myoinvcat1; roc
'MI+Ms7' myoinvcat1 ms7;
roc 'Grade+MI+Ms7' grade3 myoinvcat1 ms7;
roccontrast /estimate=allpairs;
run;

/*** Training 2: GOG64 ***/
/*** Micoarray data ***/

data GOG64; set
GOG64;

/*** Genetic score based on 7 genes (from discovery stage) ***/
MS7=
BDNFOS_497258*(-1.9217) +
APOL4_80832 * (-0.6422) +
RSRC1_51319 *(1.6294) +
PDLIM3_27295 *(1.0483) +
TBRG1_84897 *(-1.5770) +
ZNF596_169270 *(-0.7243) +
MLLT10_8028 *(1.7836);

/*** Optimal classification from training 2 (based on Youden index) ***/
ms7grp=ms7>=-4.25324;
run;

/*** MS7 / Grade / MI in Prediction of Metastasis ***/
/*** No patient with missing data ***/
proc logistic data=GOG64;
model stgrp (event='1')=grade3 myoinvcat1 ms7; roc
'MS7' ms7;
roc 'Grade' grade3;
roc 'Grade+Ms7' grade3 ms7;

roc 'MI' myoinvcat1;
roc 'Grade+MI' grade3 myoinvcat1; roc
'MI+Ms7' myoinvcat1 ms7;
roc 'Grade+MI+Ms7' grade3 myoinvcat1 ms7;
roccontrast /estimate=allpairs;
run;

/*** Validation 2: DK81 ***/
/*** Microarray data ***/

data DK81; set
DK81;

/*** Genetic score based on 7 genes (from discovery stage) ***/
```

```
MS7=
BDNFOS_497258*(-1.9217)                    +
APOL4_80832 * (-0.6422) + RSRC1_51319
*(1.6294)                                  +
PDLIM3_27295 *(1.0483)                     +
TBRG1_84897 *(-1.5770)                     +
ZNF596_169270 *(-0.7243) + MLLT10_8028
*(1.7836);

/*** Optimal classification from training 2 (based on Youden index) ***/
ms7grp=ms7>=-4.25324;
run;

/*** MS7 / Grade / MI in Prediction of Metastasis ***/
/*** No patient with missing data ***/
proc logistic data=DK81;
model stgrp (event='1')=grade3 myoinvcat1 ms7; roc
'MS7' MS7;
roc 'Grade' Grade3;
roc 'Grade+MS7' Grade3 Ms7;

roc 'MI' Myoinvcat1;
roc 'Grade+MI' Grade3 Myoinvcat1; roc
'MI+Ms7' Myoinvcat1 Ms7;
roc 'Grade+MI+Ms7' Grade3 Myoinvcat1 ms7 ;
roccontrast /estimate=allpairs;
run;


/***** FIGURES (ROC) *****/

/*** MS7 in Prediction of Metastasis (ROC curves) ***/
/*** in Training and Validation Datasets ***/

/*** for RNA-seq data ***/
proc logistic data=TCGA75;
      model stgrp(event='1') = ms7 / outroc=troc;
      score data=V245 out=valpred outroc=vroc;
      run;
data a;
set troc(in=train) vroc; data="valid";
if train then data="train";
run;

data b; set a;
if data='train' then y1=_sensit_;
else if data='valid' then y2=_sensit_;
run;

proc sgplot data=b;
      xaxis values=(0 to 1 by 0.25) valueattrs=(weight=bold size=12 color=black) label='1 - Specificity' labelattrs=(weight=bold size=12
color=black);
      yaxis values=(0 to 1 by 0.25) valueattrs=(weight=bold size=12 color=black) label='Sensitivity' labelattrs=(weight=bold size=12
color=black);
      lineparm x=0 y=0 slope=1 / lineattrs=(thickness=1 pattern=4 color=Black) CURVELABEL=' '; series
      x=_1mspec_ y=y1 /lineattrs=(thickness=2 pattern=1 color=Blue);
      series x=_1mspec_ y=y2 /lineattrs=(thickness=2 pattern=1color=red); label
Y1='Training (AUC=0.889)' Y2='Validation (AUC=0.754)';
keylegend / noborder title=' ' location=outside position=bottom across=2 valueattrs=(weight=bold size=10 color=black);
run;

/*** for Affymetrix array data ***/

proc logistic data=GOG64;
      model stgrp(event='1') = MS7 / outroc=troc;
      score data=DK81 out=valpred outroc=vroc;
      run;
data c;
set troc(in=train) vroc;
      data="valid";
if train then data="train";
```

```
run;

data d;
set c;
if data='train' then y1=_sensit_;
else if data='valid' then y2=_sensit_;
run;

proc sgplot data=d;
      xaxis values=(0 to 1 by 0.25) valueattrs=(weight=bold size=12 color=black) label='1 - Specificity' labelattrs=(weight=bold size=12
color=black);
      yaxis values=(0 to 1 by 0.25) valueattrs=(weight=bold size=12 color=black) label='Sensitivity' labelattrs=(weight=bold size=12
color=black);
      lineparm x=0 y=0 slope=1 / lineattrs=(thickness=1 pattern=4 color=Black) CURVELABEL=' ';
      series x=_1mspec_ y=y1 /lineattrs=(thickness=2 pattern=1 color=Blue);
      series x=_1mspec_ y=y2 /lineattrs=(thickness=2 pattern=1color=red);
label Y1='Training (AUC=0.894)' Y2='Validation(AUC=0.744)';
keylegend / noborder title=' ' location=outside position=bottom across=2 valueattrs=(weight=bold size=10 color=black);
run;


/*************************************************************************/
/*** Evaluation of Classification Accuracy in Training and Validation Datasets ***/
/****MS7 Optimal Classification Obtained based on Youden Index *************/
/*************************************************************************/


/*** Sensitivity (SE) and specificity (SP) on Training 1: TCGA75 (RNA-seq data) ***/


/*** Sensitivity ***/
proc freq data=TCGA75 order=freq;
tables ms7grp/binomial(exact) alpha=.05;
where stgrp=1;
run;

/*** Specificity ***/
proc freq data=TCGA75 order=freq;
tables ms7grp/binomial(exact) alpha=.05;
where stgrp=0;
run;

/*** Positive Predictive Value (PPV) and Negative Predictive Value (NPV) also Estimated by Assuming Metastasis Rate (P) = 15% in Clinical
Practice ***/
/*** 95% CI Estimated using Bootstrap Method (SAS program not shown here) ***/
/***

 P = 0.15;
 PPV = (Sen*P)/(Sen*P + (1-SP)*(1-P));
 NPV = (SP*(1-P))/(SP*(1-P)+(1-Sen)*P);

***/


/*** SE and SP on Validation 1: V245 (RNA-seq data)***/

proc freq data=V245 order=freq;
tables ms7grp/binomial(exact) alpha=.05;
where stgrp=1;
run;
```

```
proc freq data=V245 order=freq;
tables ms7grp/binomial(exact) alpha=.05; where
stgrp=0;
run;

/*** NPV and NPV also estimated by assuming metastasis rate (P) = 15% in clinical practice ***/
/*** 95% CI estimated using Bootstrap approach***/


/*** SE and SP on Training 2: GOG64 (Microarray data) ***/

proc freq data=GOG64 order=freq; tables
ms7grp/binomial(exact) alpha=.05; where
stgrp=1;
run;

proc freq data=GOG64 order=freq; tables
ms7grp/binomial(exact) alpha=.05; where
stgrp=0;
run;

/*** NPV and NPV also estimated by assuming metastasis rate (P) = 15% in clinical practice ***/
/*** 95% CI estimated using Bootstrap approach (not shown here) ***/

/*** SE and SP on Validation 2: DK81 (Microarray data) ***/

proc freq data=DK81 order=freq;
tables ms7grp/binomial(exact) alpha=.05; where
stgrp=1;
run;

proc freq data=DK81 order=freq;
tables ms7grp/binomial(exact) alpha=.05; where
stgrp=0;
run;

/*** NPV and NPV also estimated by assuming metastasis rate (P) = 15% in clinical practice ***/
/*** 95% CI estimated using Bootstrap approach***/


/*******************************************************/
/*** Combing Validation 1 and Validation 2***/

data v1; set
v245;
keep ms7 ms7grp grade grade3 myoinvcat1 stgrp;
run;

data v2; set
DK81;
keep ms7 ms7grp grade3 myoinvcat1 stgrp;
run;

data vall; set
v1 v2;
mi=myoinvcat1-1; /*** Recoded as 0/1 ***/
run;


/*** NPV and NPV also estimated by assuming metastasis rate (P) = 15% in general practice ***/
/*** 95% CI estimated using Bootstrap approach***/


/*** END ***/
```

M5. Evaluating the Relationship between the MS7 Classifier and Molecular Subtypes, Cancer Biomarkers and Functional Pathway Analysis

The relationship between categorized MS7 score and molecular subtype of endometrial cancer as defined by the Uterine Corpus Endometrial Carcinoma (UCEC) of the Cancer Genome Atlas Research Network was evaluated in the Training-1 Cohort (N=75) using Fisher's exact test, and the odds ratio and 95% confidence interval were calculated from logistic regression modeling. The relationship between MS7 score and RNAseq based transcript expression of some common cancer biomarkers was evaluated (**Supplement S5**) using Spearman's correlation coefficient in Training-1 (N=75) and correlation were validated in the independent Validation-1 cases with RNAseq data (N=245). The relationship between categorized MS7 score and mutation status of ARID1A, CTNNB1, KRAS, PIK3CA, TP53 and PTEN was evaluated in the 197 TCGA cases with mutation data (**Supplement S6**) using Fisher's exact test, and the odds ratio and 95% confidence interval were calculated from logistic regression modeling.

Functional pathway analysis of the MS7 classifier was performed by categorizing the patients in the Training-1 Cohort into quartiles (Q) by their MS7 score. There were 19 patients with the highest MS7 score (Q4) and another 19 with the lowest MS7 score (Q1). Differentially expressed transcripts with $q<0.05$ in these two groups were identified and functional pathway analysis was performed as follows. Significantly, differential abundant (q-value ≤ 0.05) gene identifiers were hand-curated using public data sources, such as NCBI to afford mapping of all significantly different genes in the Ingenuity Pathway Analysis (IPA) knowledgebase. Significant genes were uploaded to IPA along with fold-change data and gene lists were subjected to "Core Analyses" using default settings. The top enriched canonical pathways in high versus low risk patients were prioritized for further analyses. Predicted activation

states of canonical pathways were based on abundance trends of associated transcripts relative to the putative functional role of associated candidates in the pathway of interest. Results are presented in **Supplement S7** and **S8**.

M6. Download Data

The data utilized in this publication for the Training-2 and the Validation-2 Cohort have been deposited in NCBI's Gene Expression Omnibus and are accessible through GEO Series Accession Number GSE120490 (https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE120490). The date for release of this data is currently set for December 31, 2020 and will be reset to the date when this manuscript is published.