

Supplementary Appendix

This appendix has been provided by the authors to give readers additional information about their work.

Supplement to: Paik S, Shak S, Tang G, et al. A multigene assay to predict recurrence of tamoxifen-treated node-negative breast cancer. *N Engl J Med* 2004;351:2817-26.

Supplemental Background and Methods

Section I. Detailed Assay Methods

Sample Preparation

All sample preparation was performed at the NSABP Division of Pathology. We reviewed all the blocks from tamoxifen-treated patients in NSABP B-14 that were currently in the NSABP Tumor Bank. A section was cut, using a Histo Collimator (Microm International) and Paraffin Sectioning Aid (Instrumedics, Hackensack, NJ) to minimize tissue usage, stained with Hematoxylin and Eosin, and reviewed by the NSABP and Genomic Health pathologists. Cases with no cancer (depleted by prior tissue studies) or with cancer cells occupying <5% of the section area were excluded from the study. The blocks in 79 patients had insufficient tumor and did not meet the protocol eligibility criteria. Macro-dissection to obtain enriched tumor tissue was performed using a safety blade cleaned with RNaseZAP (Ambion, Austin, TX) on sections having non-tumor elements (such as smooth muscle, fibrosis, hemorrhage, normal breast stromal tissue, but not DCIS or LCIS or necrosis) that were both sufficiently localized to be amenable to macro-dissection and constituted >50% of the overall tissue area of the section. RNA was

extracted from three 10 micron thick sections without dissection or from macro-dissected tumor from six 10 micron sections. Cut sections were placed into 1.5 ml RNase/DNase-free microfuge tubes (PGC Scientific, Frederick, MD), labeled with a bar-code (blinding the sample), and shipped at ambient temperature to the reference laboratory at Genomic Health, Inc.

Gene Expression

Gene expression in fixed paraffin embedded tumor tissue was performed by the methods recently described by Cronin et al (22). Paraffin was removed from specimens by xylene extraction followed by ethanol washes. RNA was isolated using the Epicentre Technologies, Inc. MasterPure™ kit (Madison, WI). DNase I treatment is routinely included in the extraction step. Total RNA content was measured using Ribogreen® Fluorescence Quantitation Kit (Molecular Probes, Inc., Eugene, OR). Residual genomic DNA contamination was assayed by a quantitative TaqMan® PCR assay for β -actin DNA, which includes both positive and negative controls. Samples with contaminating DNA were re-subjected to DNase I treatment, and assayed again for DNA contamination.

Reverse transcription of the purified RNA was carried out with the Omniscript RT kit (Qiagen, Valencia, CA) for first strand synthesis, using simultaneous random hexamer and gene-specific priming. TaqMan® reactions were carried out in 384 well plates according to instructions of the manufacturer, using Applied Biosystems PRISM® 7900HT instruments. Expression of each gene was measured (the number of cycles required to achieve a threshold, or C_T) in triplicate, aggregated, and then normalized relative to a set of five reference genes [beta-actin (or ACTB), GAPDH (or GAPD), GUS (or GUSB), RPLPO, and TFRC] by subtracting the average of the expression of the 5 reference genes. Reference-normalized expression measurements range from 0 to 15, where a one unit increase reflects approximately a 2-fold increase in RNA (amplification efficiencies ranged from 75% to 112% for the 21 genes).

Stringent quality specifications were developed for each step of the assay and an RNA standard is used as a calibrator to ensure intra-laboratory accuracy. The accuracy, precision, linearity, analytical sensitivity, dynamic range and reproducibility of the assay as a function of day, machine, and operator were documented as part of the rigorous federally-regulated CLIA review process. Reproducibility within and between blocks was

assessed by measuring expression in 5 serial sections from 6 blocks in 2 patients. The standard deviation for expression of the 16 cancer-related genes in serial sections is 0.07 to 0.21 expression units. The standard deviation within block for the Recurrence Score is 0.72 RS units (95% CI, 0.55, 1.04). The total standard deviation within patient (includes between blocks and within block) is 2.2 RS units. Similar variability in the RS was observed in reanalysis of clinical trial samples on separate days with different reagent lots (data not shown).

Section 2: Development of the Recurrence Score Assay

To develop and clinically validate a tumor gene expression assay for use with tumor blocks that are routinely prepared following surgery, we used a multi-step approach.

First, a real time RT-PCR method to quantify the expression of hundreds of genes in RNA isolated from three 10 micron sections of fixed paraffin embedded tumor tissue was developed (22). Assay sensitivity, specificity, and reproducibility were characterized and optimized for measurement of fragmented RNA that is extracted from fixed tumor tissue stored for varying periods of time.

Second, we selected 250 candidate genes from the published literature, genomic databases, pathway analysis, and from microarray-based gene expression profiling experiments performed using fresh frozen tissue (17-19, 23). The 250 candidate genes include:

Table. Listing of the 250 Candidate Genes	
Gene Name	Accession Number
A-Catenin	NM_001903
ACTG2	NM_001615
AD024	NM_020675
AIB1	NM_006534
AK055699	AK055699
AKAP-2	NM_007203

AKT1	NM_005163
AKT2	NM_001626
AKT3	NM_005465
ALDH1A1	NM_000689
ALDH1A3	NM_000693
APC	NM_000038
AREG	NM_001657
B-actin	NM_001101
B-Catenin	NM_001904
BAD	NM_032989
BAG1	NM_004323
Bak	NM_001188
Bax	NM_004324
BBC3	NM_014417
Bcl2	NM_000633
Bclx	NM_001191
BECN1	NM_003766
BG675392	BG675392
BIN1	NM_004305
BRCA1	NM_007295
BRCA2	NM_000059
BRK	NM_005975
BRMS1	AF159141
BUB1	NM_004336
C20 orf1	NM_012112
C20orf103	NM_012261
CA9	NM_001216
CCNB1	NM_031966
CCND1	NM_001758
CCNE1	NM_001238
CCNE2	NM_057749
CD105	NM_000118
CD18	NM_000211
CD31	NM_000442
CD3z	NM_000734
CD44E	X55150
CD68	NM_001251
CD9	NM_001769
CDC20	NM_001255
CDC25B	NM_021874
CDH1	NM_004360
CEGP1	NM_020974
CHAF1B	NM_005441
Chk1	NM_001274
Chk2	NM_007194
CIAP1	NM_001166
clAP2	NM_001165
cMet	NM_000245
cMYC	NM_002467
CNN	NM_001299
COL1A1	NM_000088
COL1A2	NM_000089
Contig 27882	AK000618

Contig 36744	XM_087225
Contig 51037	XM_058945
Contig44799	Contig44799
Contig46653	Contig46653
Contig47405	Contig47405
COX2	NM_000963
CRBP	NM_002899
CSK	NM_004383
CSTF1	NM_001324
CTSB	NM_001908
CTSL	NM_001912
CTSL2	NM_001333
Cyclin G1	NM_004060
CYP	NM_006347
CYP2C8	NM_000770
CYP3A4	NM_017460
DAPK1	NM_004938
DCR3	NM_016434
DHPS	NM_013407
DIABLO	NM_019887
DKFZp564	XM_047080
DKFZp586M0723	AL050227
DPYD	NM_000110
DR4	NM_003844
DR5	NM_003842
EGFR	NM_005228
EGFRd27	EGFRd27
EIF4E	NM_001968
EIF4EL3	NM_004846
EMS1	NM_005231
EpCAM	NM_002354
EPHX1	NM_000120
ErbB3	NM_001982
EREG	NM_001432
EstR1	NM_000125
fas	NM_000043
fasl	NM_000639
FBXO5	NM_012177
FGF18	NM_003862
FGF2	NM_002006
FGFR1	NM_023109
FHIT	NM_002012
FKBP4	NM_002014
FLJ20354 (DEPDC1 official)	NM_017779
FLT1	NM_002019
FOXM1	NM_021953
FRP1	NM_003012
G-Catenin	NM_002230
GAPDH	NM_002046
GATA3	NM_002051
GCLC	NM_001498
GPC3	NM_004484
GRB7	NM_005310

GRO1	NM_001511
GSN	NM_000177
GSTM1	NM_000561
GSTM3	NM_000849
GSTp	NM_000852
GUS	NM_000181
HB-EGF	NM_001945
hENT1	NM_004955
HER2	NM_004448
HGF	M29145
HIF1A	NM_001530
HLA-DPB1	NM_002121
HLA-G	NM_002127
HNF3A	NM_004496
HNRPAB	NM_004499
ID1	NM_002165
ID2	NM_002166
IGF1	NM_000618
IGF1R	NM_000875
IGFBP2	NM_000597
IGFBP3	NM_000598
IGFBP5	NM_000599
IL6	NM_000600
ILT-2	NM_006669
IRS1	NM_005544
ITGA7	NM_002206
KDR	NM_002253
Ki-67	NM_002417
KIAA1209	AJ420468
KLK10	NM_002776
KNSL2	BC000712
KRT14	NM_000526
KRT17	NM_000422
KRT18	NM_000224
KRT19	NM_002276
KRT5	NM_000424
KRT8	NM_002273
LAMC2	NM_005562
LMNB1	NM_005573
LOT1 variant 1	NM_002656
LPL	NM_000237
Maspin	NM_002639
MCM2	NM_004526
MCM3	NM_002388
MCM6	NM_005915
MCP1	NM_002982
MDM2	NM_002392
MDSO28	NM_018463
MELK	NM_014791
MMP12	NM_002426
MMP2	NM_004530
MMP9	NM_004994
MT3	NM_005954

MTA1	NM_004689
MVP	NM_017458
MYBL2	NM_002466
MYH11	NM_002474
MYLK	NM_053025
MYRIP	NM_015460
NEK2	NM_002497
NFKBp65	NM_021975
NME1	NM_000269
NMU	NM_006681
NMYC	NM_005378
NPD009 (ABAT official)	NM_020686
NR4A1	NM_002135
ODC1	NM_002539
p14ARF	NM_000077
p14ARF	S78535
p21	NM_000389
p27	NM_004064
P40	NM_006824
P53	NM_000546
PAI1	NM_000602
PCNA	NM_002592
PDGFA	NM_002607
PDGFB	NM_002608
PDGFRa	NM_006206
PDGFRb	NM_002609
PI3KC2A	NM_002645
Pin1	NM_006221
PLAUR	NM_002659
PPM1D	NM_003620
PR	NM_000926
PRAME	NM_006115
PREP	NM_002726
PRO2000	NM_014109
pS2	NM_003225
PTEN	NM_000314
PTTG1	NM_004219
RAB27B	NM_004163
RAD51C	NM_058216
RAD54L	NM_003579
RARA	NM_000964
RASSF1	NM_007182
RB1	NM_000321
RBM5	NM_005778
RBP4	NM_006744
RFC	NM_003056
rhoC	NM_005167
RIZ1	NM_012231
RPLPO	NM_001002
RPS6KB1	NM_003161
RUNX1	NM_001754
SNRPF	NM_003095
STK15	NM_003600

STMY3	NM_005940
SURV	NM_001168
TBP	NM_003194
TERC	U86046
TERT	NM_003219
TFRC	NM_003234
TGFA	NM_003236
TGFB3	NM_003239
TGFBR2	NM_003242
TIMP1	NM_003254
TIMP2	NM_003255
TIMP3	NM_000362
TJP1	NM_003257
TK1	NM_003258
TOP2A	NM_001067
TOP2B	NM_001068
TP	NM_001953
TP53BP1	NM_005657
TP53BP2	NM_005426
TRAIL	NM_003810
TS	NM_001071
TULP3	NM_003324
upa	NM_002658
VDR	NM_000376
VEGF	NM_003376
VEGFB	NM_003377
VEGFC	NM_005429
VIM	NM_003380
WISP1	NM_003882
XIAP	NM_001167
YB-1	NM_004559
ZNF217	NM_006526

Third, we performed three independent breast cancer clinical studies in a total of 447 patients to test the relationship between the expression of the 250 candidate genes and recurrence (24-26). We reasoned that in any single gene expression study a considerable number of the genes may correlate with the outcome as a result of chance alone. To identify true positives, we performed the three independent studies to identify whether expression of any of the 250 candidate genes correlated with recurrence across

the studies. We hypothesized that the genes most highly correlated with recurrence would survive evaluation across diverse patients and treatments, and selected heterogeneous populations for development of the gene list:

Table. Three Breast Cancer Studies to Develop the Gene Panel and Algorithm for Prospective Validation

Site	N	Adjuvant Treatment	
NSABP Study B-20 Pittsburgh, PA	233 Node - ER+	Tamoxifen Chemotherapy	100% 0%
Rush University Medical Center Chicago, IL	78 ≥ 10 nodes ER+/-	Tamoxifen Chemotherapy	54% 80%
Providence St. Josephs Hospital Burbank, CA	136 Node +/- ER+/-	Tamoxifen Chemotherapy	41% 39%

There were 9 genes whose expression correlated with recurrence in all three studies at an unadjusted p-value < 0.05. There were 5 additional genes whose expression correlated with outcome in all three studies at an unadjusted p-value < 0.10. There were an additional 9 genes whose expression correlated with outcome in 2 studies at an unadjusted p-value < 0.05.

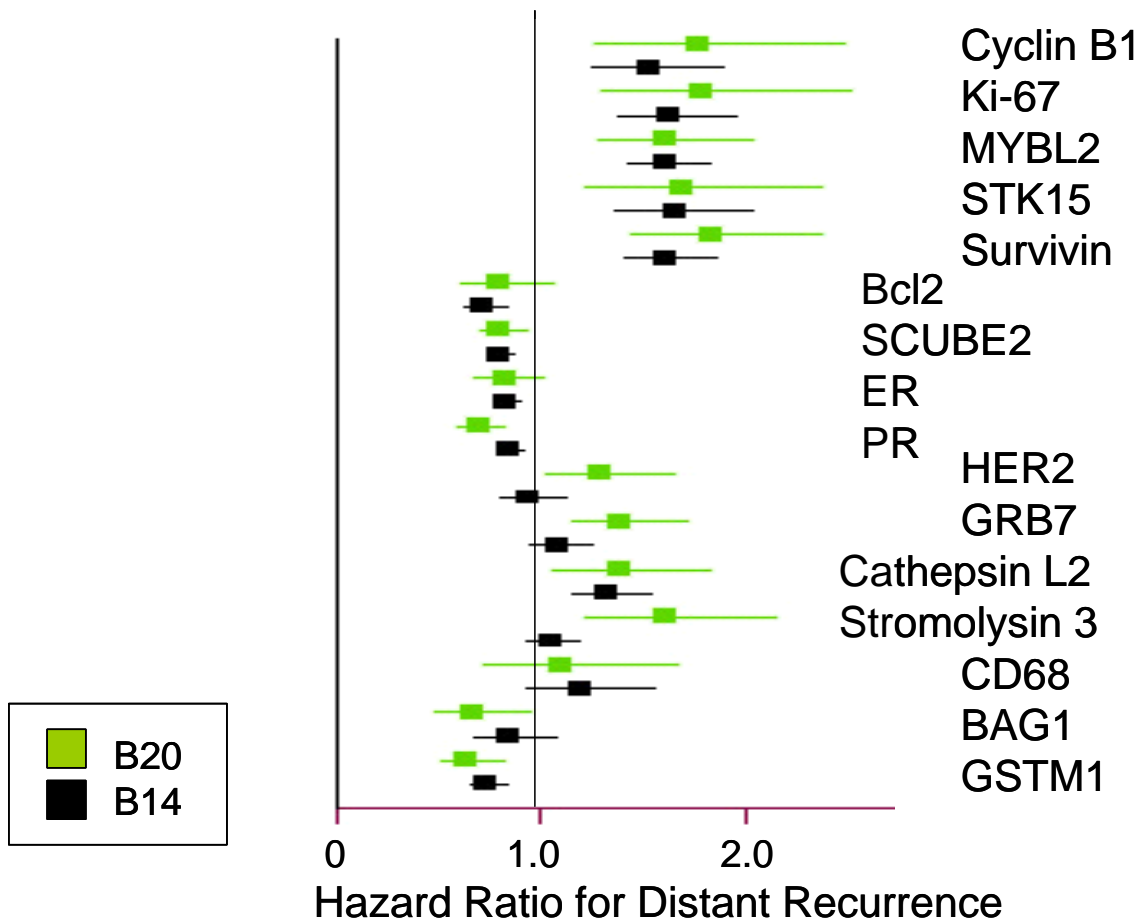
Fourth, we used the results from these three preliminary studies to select the final gene panel and to design an algorithm to compute a Recurrence Score (RS) for each tumor sample. Since we planned to validate the Recurrence Score in the large

homogeneous population of NSABP B-14 tamoxifen-treated patients, we weighted the NSABP B-20 study results most heavily in selecting the final gene list and algorithm for validation. Genes that were consistently significant across multiple studies provided the basis for developing of the Recurrence Score model. Analyses were then performed to determine the functional forms of the variables to include in that model and the appropriate number of terms to include in the model. We did not assume that linear relationships would be appropriate for all genes, and explored non-linear relationships. For this purpose, we used correlation analysis, dimension reduction, Martingale residual analysis, concordance measures of accuracy, and bootstrap resampling (to be published in detail separately).

The correlation of the expression of the genes with respect to each other was analyzed by unsupervised cluster analysis and by principal component analysis. A number of groups were identified—a proliferation group, including genes such as Ki-67 and Cyclin B1, an ER group, including genes such as ER and PR, a HER2 group, including HER2 and GRB7, and an invasion group, including genes such as Stromolysin 3 and Cathepsin L2. The selection of the final 16 cancer-related genes was based

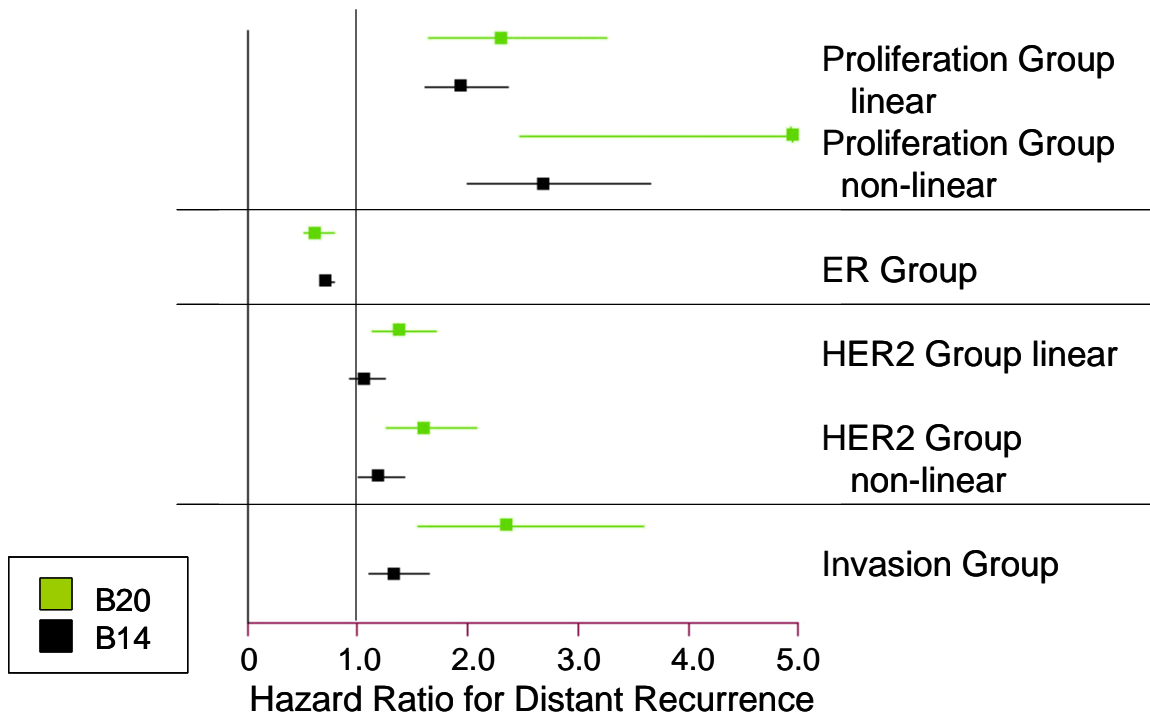
primarily on the strength of their performance across all three studies and the consistency of primer/probe performance in the assay.

The consistency between the results of the RT-PCR analysis in the NSABP B-20 samples that were used to develop the Recurrence Score and the results in the NSABP B-14 samples where the Recurrence Score was validated is notable. Shown below are forest plots of the Kaplan-Meier estimate for the Hazard Ratios for all 16 cancer-related genes in both studies. The 95% confidence intervals for the Hazard Ratios for all the individual genes overlap between studies.



Moreover, consistency between the B-20 and the B-14 studies was also observed

the hazard ratios for the proliferation, the ER, the HER2, and the invasion groups were compared.



The predictive power of the gene groups is greater than that of individual genes.

Figure in Supplemental Background and Methods – Section 2

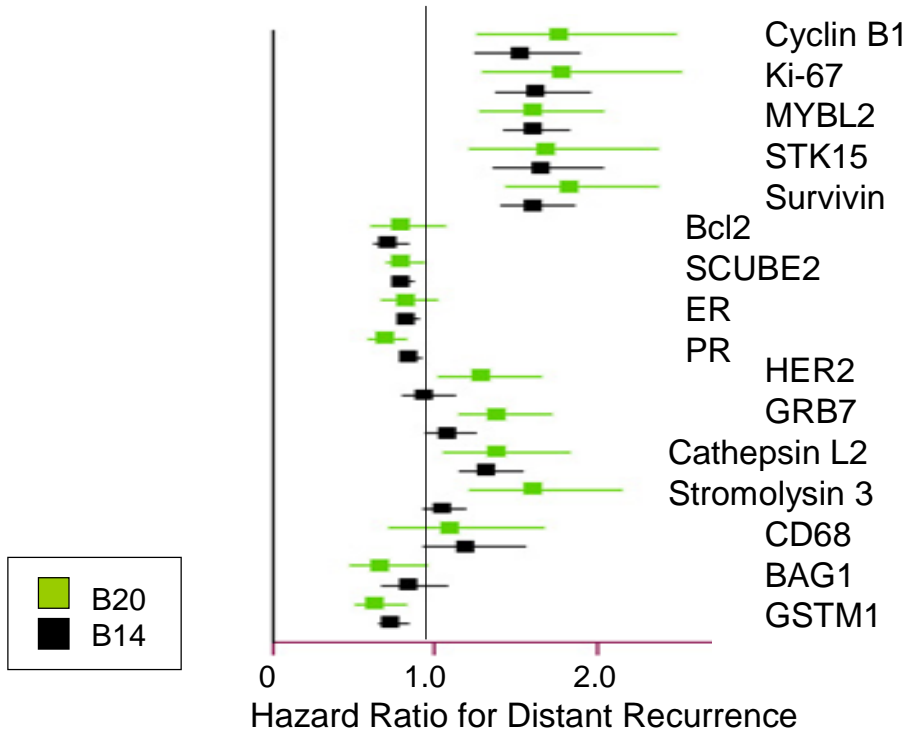
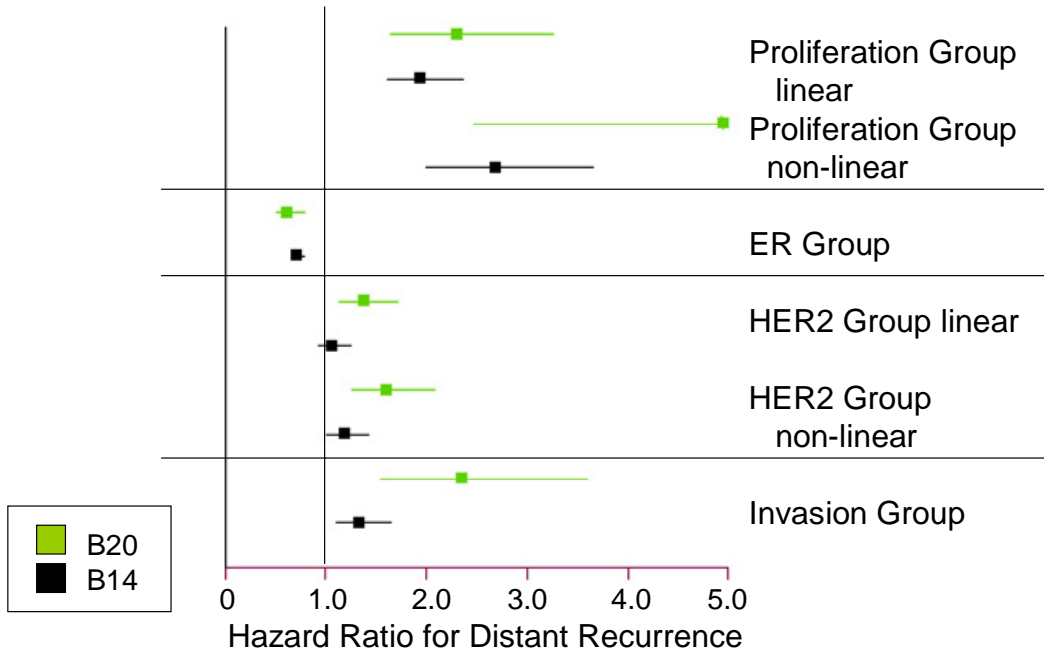


Figure in Supplemental Background and Methods – Section 2



Supplementary Tables and Figures

Supplementary Table 1. Patient Characteristics

Table 1. Baseline Characteristics of the 668 Patients Included in the Analysis			
Characteristic	Genomic Health Population (N = 668)	Remaining TAM-treated population (N = 1949)	NSABP B-14 TAM-treated population (N = 2617)
Treatment			
Randomized Pts	290 (43%)	1114 (57%)	1404 (54%)
Registered Pts	378 (57%)	835 (43%)	1213 (46%)
Age at Enrollment**			
< 50 yr of age	194 (29%)	618 (32%)	812 (31%)
50 ≤ age < 60	173 (26%)	583 (30%)	756 (29%)
≥ 60 yr of age	301 (45%)	748 (38%)	1049 (40%)
Race			
White	615 (92%)	1763 (90%)	2378 (91%)
Black	31 (5%)	87 (5%)	118 (4%)
Others	22 (3%)	99 (5%)	121 (5%)
Clinical Tumor Size			
0 – 1.0 cm	109 (16%)	366 (19%)	475 (18%)
1.1 – 2.0 cm	305 (46%)	877 (45%)	1182 (45%)
2.1 – 4.0 cm	220 (33%)	635 (32%)	855 (33%)
4.1+ cm	34 (5%)	71 (4%)	105 (4%)

** Likelihood Ratio test significant at p-Value < 0.01

Supplementary Table 2. Tumor Grade and Distant Recurrence for Three Pathologists

Supplementary Table 2A. Multivariate Cox-Proportional Model for Tumor Grade (Pathologist A) and Recurrence Score			
Without Recurrence Score			
Variable	Pr > χ^2	Hazard Ratio	95% C.I. for Hazard Ratio
Moderate ¹	0.04	1.84	(1.04, 3.25)
Poor ¹	<0.001	5.40	(3.12, 9.35)
With Recurrence Score			
Variable	Pr > χ^2	Hazard Ratio	95% C.I. for Hazard Ratio
Moderate ¹	0.15	1.54	(0.86, 2.74)
Poor ¹	<0.001	3.42	(1.86, 6.29)
Recurrence Score ²	<0.001	2.27	(1.51, 3.41)
Likelihood Ratio Test for Recurrence Score and Tumor Grade versus Tumor Grade Alone $\chi_1^2 = 14.3, p < 0.001$			

¹ Grade used as binary factors, poor relative to well and moderate relative to well.

² Recurrence Score used as a continuous variable, with HR for distant recurrence relative to an increment of 50 units.

Supplementary Table 2B. Multivariate Cox-Proportional Model for Tumor Grade (Pathologist B) and Recurrence Score			
Without Recurrence Score			
Variable	Pr > χ^2	Hazard Ratio	95% C.I. for Hazard Ratio
Moderate ¹	0.12	1.72	(0.88, 3.37)
Poor ¹	<0.001	7.20	(3.61, 14.36)
With Recurrence Score			
Variable	Pr > χ^2	Hazard Ratio	95% C.I. for Hazard Ratio
Moderate ¹	0.23	1.51	(0.77, 2.98)
Poor ¹	<0.001	4.34	(2.03, 9.25)
Recurrence Score ²	<0.001	2.12	(1.38, 3.25)
Likelihood Ratio Test for Recurrence Score and Tumor Grade versus Tumor Grade Alone $\chi_1^2 = 10.9, p < 0.001$			

¹ Grade used as binary factors, poor relative to well and moderate relative to well.

² Recurrence Score used as a continuous variable, with HR for distant recurrence relative to an increment of 50 units.

Supplementary Table 2C. Multivariate Cox-Proportional Model for Tumor Grade (Pathologist C) and Recurrence Score			
Without Recurrence Score			
Variable	Pr > χ^2	Hazard Ratio	95% C.I. for Hazard Ratio
Moderate ¹	0.25	1.49	(0.76, 2.93)
Poor ¹	<0.001	3.73	(1.88, 7.42)
With Recurrence Score			
Variable	Pr > χ^2	Hazard Ratio	95% C.I. for Hazard Ratio
Moderate ¹	0.56	1.23	(0.62, 2.43)
Poor ¹	0.06	2.02	(0.96, 4.24)
Recurrence Score ²	<0.001	2.83	(1.88, 4.24)
Likelihood Ratio Test for Recurrence Score and Tumor Grade versus Tumor Grade Alone $\chi_1^2 = 22.4, p < 0.001$			

¹ Grade used as binary factors, poor relative to well and moderate relative to well.

² Recurrence Score used as a continuous variable, with HR for distant recurrence relative to an increment of 50 units.

Recurrence score provides significant ($p < 0.001$) information beyond tumor grade (Pathologist provided the “middle” prognostic assessment of tumor grade).

Supplementary Table 3. Pairwise Comparison of Assessments of Tumor Grade by Three Pathologists.

Supplementary Table 3A. Pairwise Comparison of Assessments of Tumor Grade					
Pathologist A vs Pathologist B					
		Pathologist B			
		Well	Moderate	Poor	Total
Pathologist A	Well	105	114	5	224
	Moderate	24	241	31	296
	Poor	3	82	63	148
	Total	132	437	99	668
Concordance = 61%; Kappa = 0.37 (0.31, 0.42)					

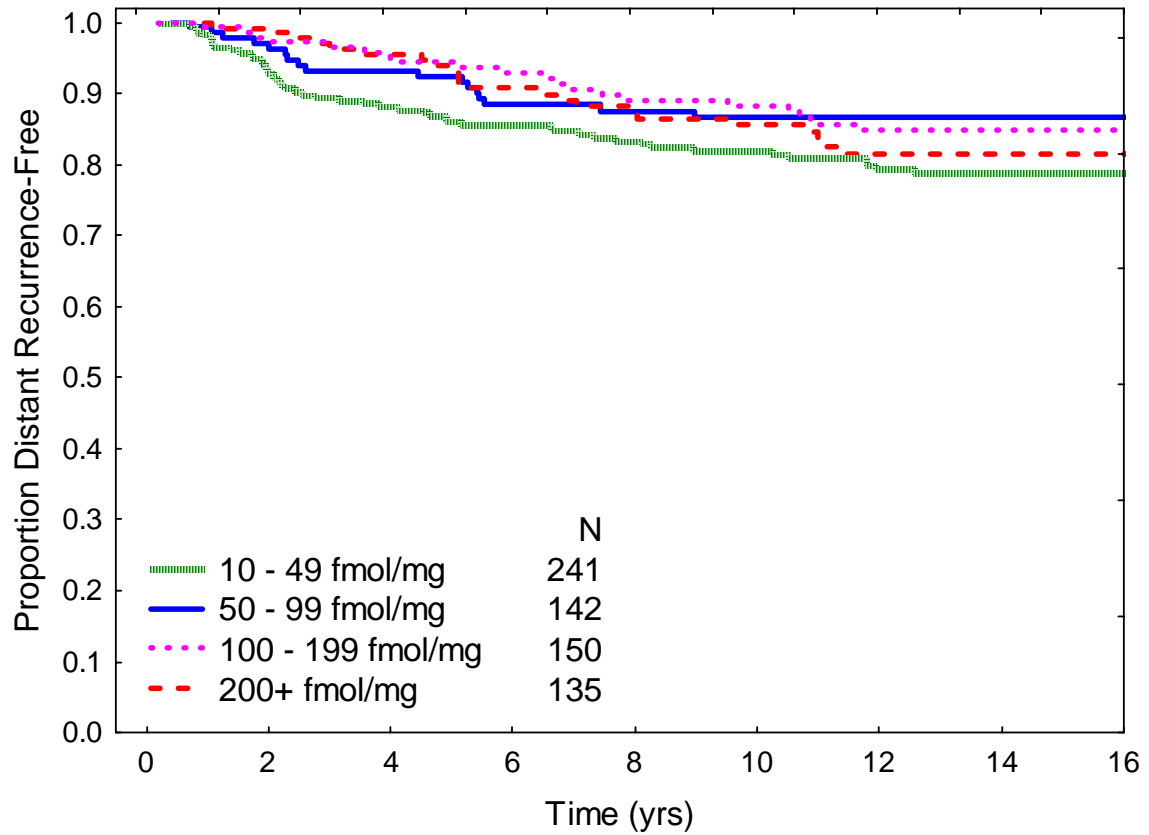
Supplementary Table 3B. Pairwise Comparison of Assessments of Tumor Grade					
Pathologist C vs Pathologist A					
		Pathologist A			
		Well	Moderate	Poor	Total
Pathologist C	Well	76	31	0	107
	Moderate	140	221	52	413
	Poor	8	44	96	148
	Total	224	296	148	668
Concordance = 59%; Kappa = 0.34 (0.28, 0.40)					

Supplementary Table 3C. Pairwise Comparison of Assessments of Tumor Grade					
Pathologist C vs Pathologist B					
		Pathologist B			
		Well	Moderate	Poor	Total
Pathologist C	Well	56	50	1	107
	Moderate	74	309	30	413
	Poor	2	78	68	148
	Total	132	437	99	668
Concordance = 65%; Kappa = 0.34 (0.27, 0.40)					

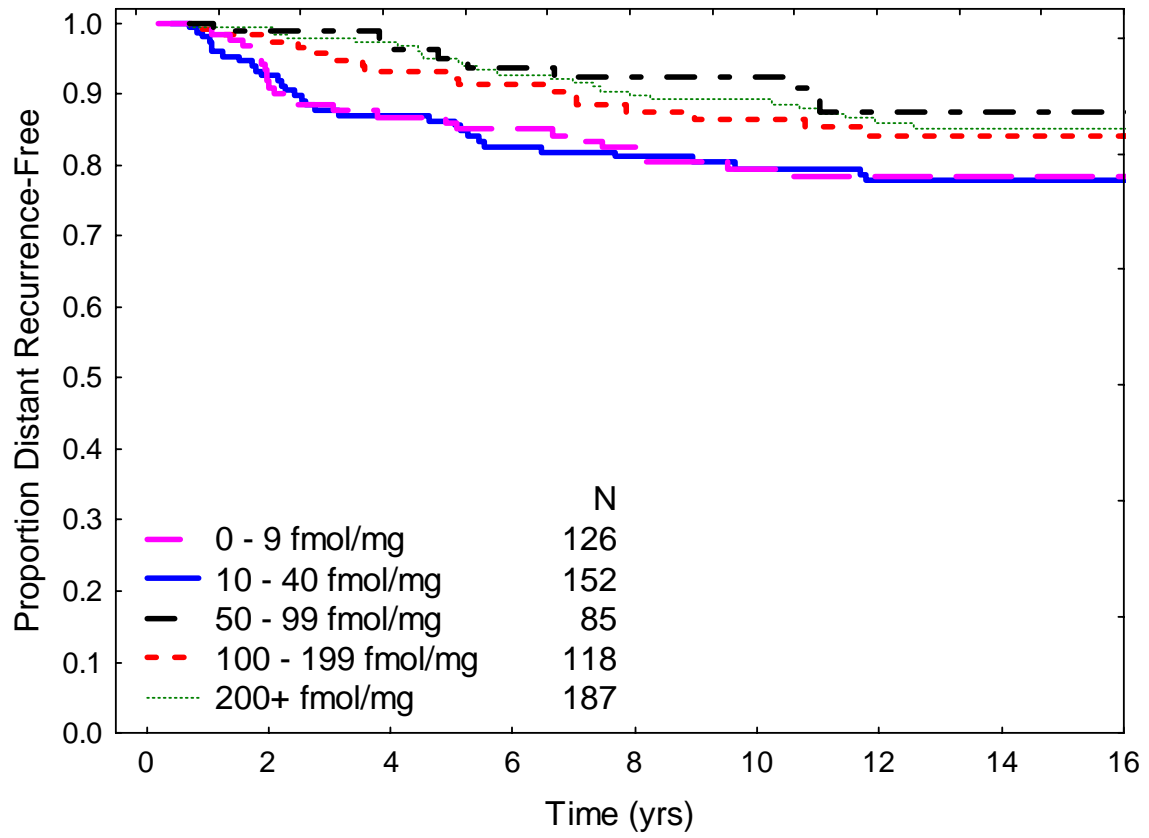
The overall agreement in tumor grade among the three pathologists is 43%

Supplementary Figure 1A and 1B. Kaplan-Meier Plots of DRFI for ER and PR by the Ligand Binding Assay.

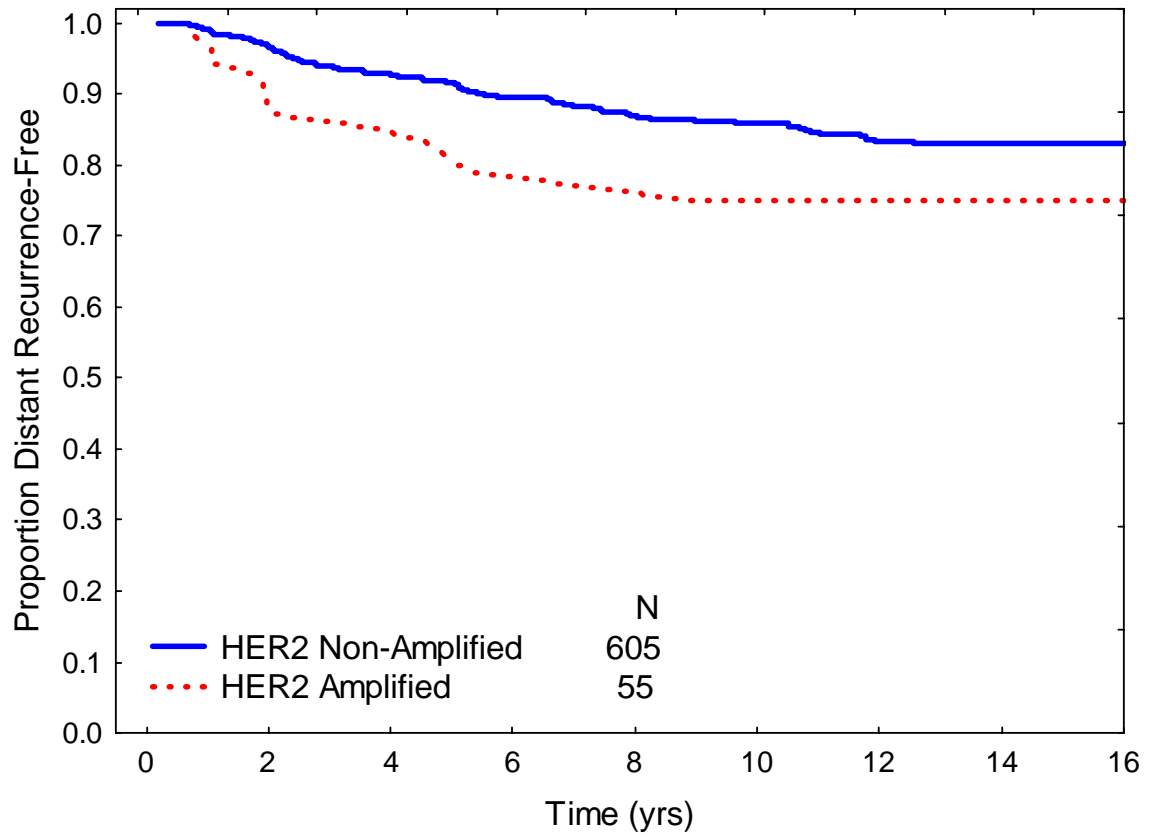
Supplementary Figure 1A. ER by the Ligand Binding Assay and DRFI



Supplementary Figure 1B. PR by the Ligand Binding Assay and DRFI

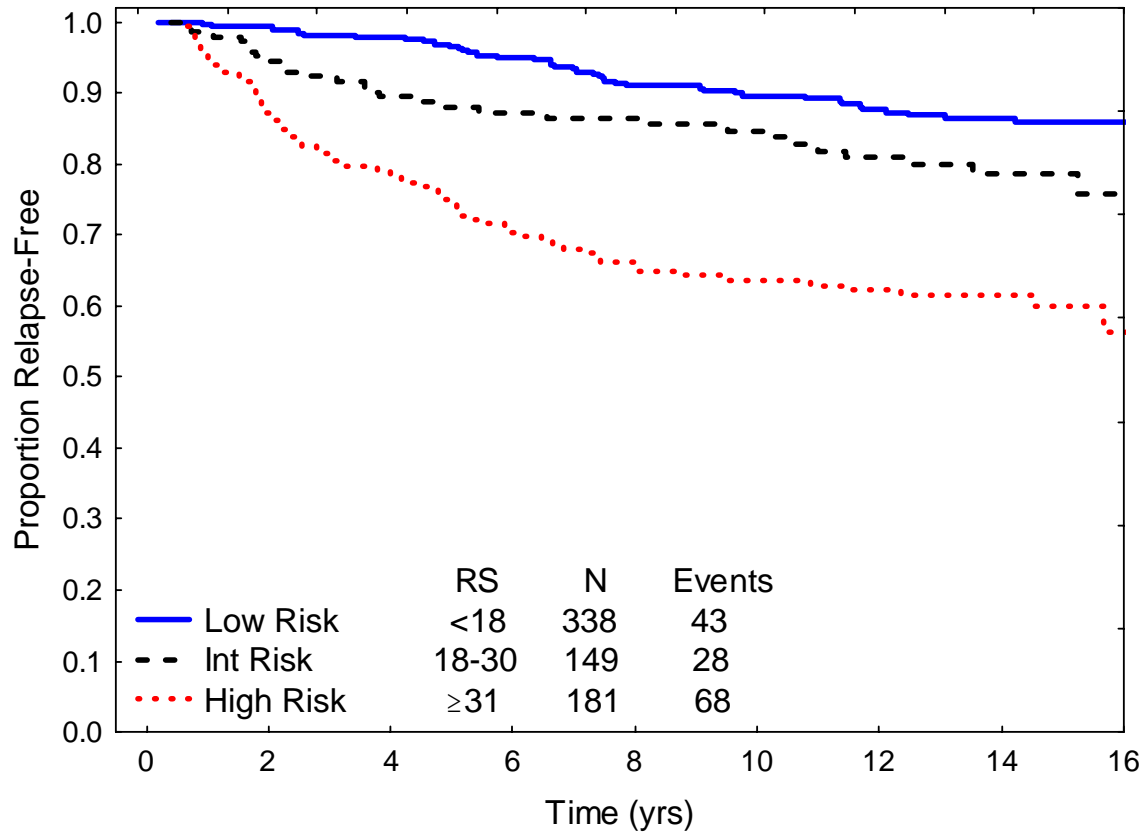


Supplementary Figure 2A. Kaplan-Meier Plot of DRFI for HER2 amplification by FISH.



Supplementary Figure 2B. Recurrence Score Categories and Relapse-Free Interval.

The difference between groups is significant ($P < 0.001$).

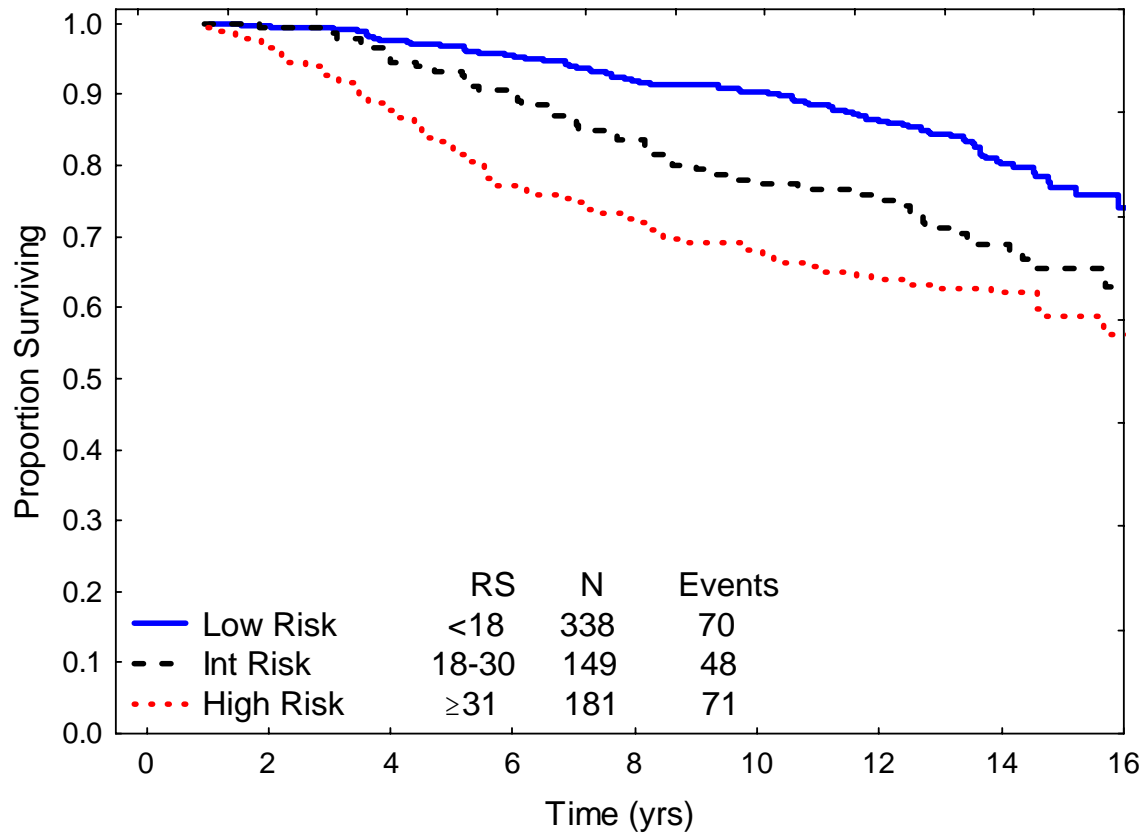


Number at Risk

Low	338	328	313	298	276	258	231	170	38
Int	149	139	128	116	104	96	80	66	16
High	181	154	137	119	105	91	83	63	13

Supplementary Figure 2C. Recurrence Score Categories and Overall Survival.

The difference between groups is significant ($P < 0.001$).

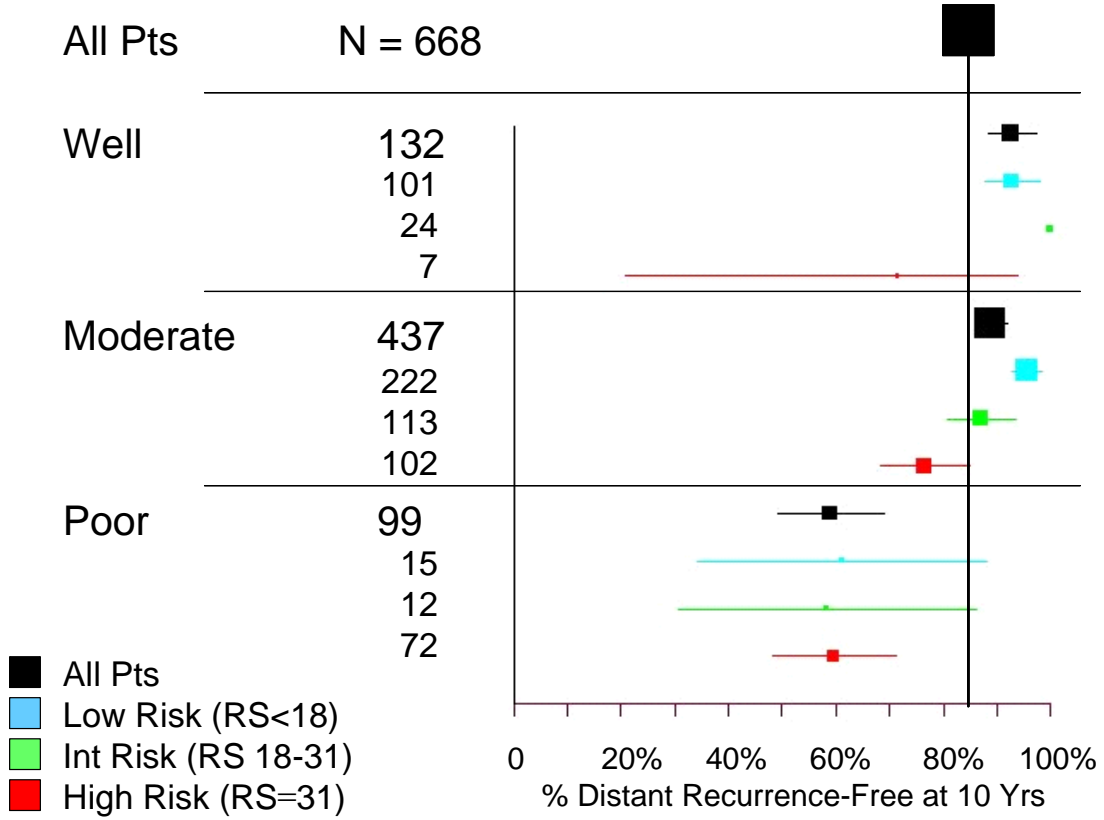


Number at Risk

Low	338	328	313	298	276	258	231	170	38
Int	149	139	128	116	104	96	80	66	16
High	181	154	137	119	105	91	83	63	13

Supplementary Figure 3. 10 year Kaplan-Meier estimates and 95% confidence intervals for the percent of patients distant recurrence-free for tumor grade subgroups.

Supplementary Figure 3A. Tumor Grade Subgroups for Pathologist B



Supplementary Figure 3B. Tumor Grade Subgroups for Pathologist C

