Diagnostic predictors of breast cancer derived from diffuse optical tomography image time series correlate with subject demographic and tumor phenotype data

H.L. Graber[1,2], R.L. Barbour[1,2], Y. Pei[2], Y. Xu[1,2], D.C. Lee[1], M.S. Katz[3], N. Patel[4], K. Jagarlamudi[5], O.R. Nwanguma[6], and W.B. Solomon[7]

[1]SUNY Downstate Medical Center; [2]NIRx Medical Technologies LLC; [3]Drexel University College of Medicine; [4]Kaiser Permanente; [5]The Brooklyn Hospital Center; [6]Pennsylvania State University; [7]Maimonides Medical Center

Introduction

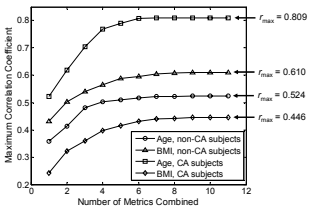
- Measures of microvascular dynamics of breast tissue:
  - Can distinguish between women with and without breast cancer (CA) with high diagnostic sensitivity and specificity (each >95%; NIH01-103).
  - An important explanatory factor is that a volume of tissue extending far from the borders of the tumor participates in the abnormal response
- Hypothesis: performance of the tumor-diagnosing optical metrics may be related to the impact of tumor-associated phenotypes on the microvasculature of the affected breast.
  - *e.g.*, angiogenesis, apoptosis, invasiveness, proliferation, stiffness

Data Analysis

1. Previous analysis (NIH01-103) of data from a 46-subject (28 with active breast CA) fNIRS breast imaging study yielded:
   - 22 multi-variate metrics that are highly sensitive and specific breast-CA predictors
   - 11 of them have no missing independent-variable values for any subject

Multivariate Predictor	Hb_{oxy}	Hb_{deoxy}	Hb_{total}	Sensitivity, Retr. (%)	Specificity, Retr. (%)	Sensitivity, Pro. (%)	Specificity, Pro. (%)
1		1-5		64.3	95	78.6	86.7
2	1-3	1-5	4,5	78.6	90	64.3	60
3	6_b	$6i,6_b$		50	90	85.7	93.3
5	2,3,6_i	6_i	1	85.7	95	57.1	86.7
6		1-5,6_i,6_b		71.4	90	57.1	93.3
7	1-3	5,6_i,6_b	1-3	85.7	95	78.6	100
8	1-3	5,6_i,6_b	1-3,5	100	90	78.6	100
14	3			50	80	64.3	93.3
15		6_i		57.1	95	57.1	100
18	3	6_i		78.6	80	85.7	100
19	3,5,6_i	6_i	4	100	95	78.6	93.3

2. Linear regression computations:
   - using the 11 predictors as regressors, and either age or body-mass index (BMI) as the dependent variables
   - yield statistically significant correlations for both age and body-mas index (BMI) when data for the cancer and non-cancer subgroups are evaluated separately (r = 0.45-0.81)



Variable	Age				BMI		
	Cancer subjects (N = 28)		Non-cancer subjects (N = 46)		Cancer subjects (N = 28)	Non-cancer subjects (N = 38)	
	n = 4[a]	n = 6[b]	n = 3[a]	n = 4[b]	n = 2[b]	n = 3[a]	n = 5[b]
Constant	40.1±5.8 (p = 5.0e-7)	44.2±6.0 (p = 3.1e-7)	40.2±2.5 (p = 2.9e-19)	41.1±2.7 (p = 1.4e-18)	29.3±3.1 (p = 8.4e-10)	32.2±1.5 (p = 2.6e-21)	31.5±1.8 (p = 1.0e-17)
MpM 1	—	-11.9±7.9 (p = 0.15)	12.5±6.9 (p = 0.075)	15.5±7.4 (p = 0.042)	—	11.2±5.6 (p = 0.054)	10.8±5.6 (p = 0.064)
MpM 2	—	—	—	—	-2.9±2.7 (p = 0.29)	—	—
MpM 3	—	—	16.8±5.8 (p = 0.006)	18.0±5.9 (p = 0.004)	3.2±3.2 (p = 0.33)	—	5.4±4.8 (p = 0.27)
MpM 5	-9.9±4.3 (p = 0.031)	-6.8±4.6 (p = 0.16)	—	—	—	—	-4.3±2.8 (p = 0.14)
MpM 6	—	—	—	—	—	-10.0±5.0 (p = 0.052)	-9.2±5.0 (p = 0.072)
MpM 7	—	—	—	—	—	—	—
MpM 8	—	-10.8±8.0 (p = 0.19)	—	—	—	—	—
MpM 14	36.2±7.5 (p = 6.6e-5)	45.9±9.1 (p = 5.0e-5)	—	-8.4±7.8 (p = 0.29)	—	-14.8±5.2 (p = 0.007)	-14.4±5.3 (p = 0.011)
MpM 15	27.1±7.6 (p = 0.002)	26.3±7.6 (p = 0.002)	—	—	—	—	—
MpM 18	-30.1±8.8 (p = 0.002)	-23.0±9.2 (p = 0.022)	—	—	—	—	—
MpM 19	—	—	-16.2±7.7 (p = 0.041)	-12.7±8.3 (p = 0.13)	—	—	—
Correlation Coefficient	0.769 (p = 2e-6)	0.809 (p < 1e-6)	0.481 (p = 7.2e-4)	0.502 (p = 3.8e-4)	0.308 (p = 0.11)	0.539 (p = 4.8e-4)	0.588 (p = 1.0e-4)

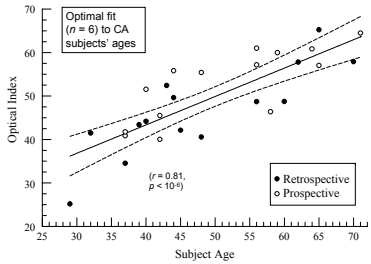
N: number of subjects, *n*: number of regressors considered in the least-squares fit. MpM = Multiparameter Metric.

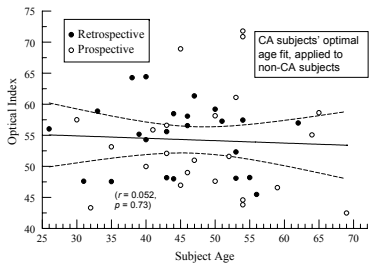
a) This number of metrics yields the optimal tradeoff between maximizing the correlation coefficient and maximizing the significance of all regressors included (*i.e.*, each regressor should have $p \le 0.1$). For the cancer subjects-BMI computation, this criterion is never satisfied.

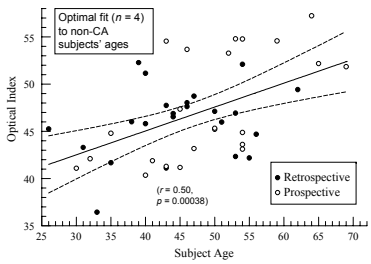
b) This number of metrics yields the optimal tradeoff between maximizing the value of the correlation coefficient estimate and minimizing its standard error.

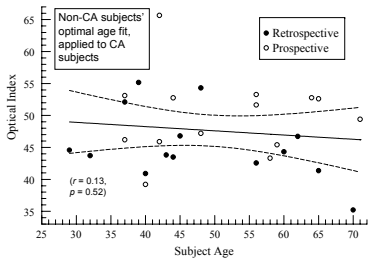
Acknowledgments: This research was supported by the National Institutes of Health (NIH) under grants R41CA096102 and R01CA066184, by the U.S. Army under grant DAMD017-03-C-0018, by the New York State Department of Health, and by Susan G. Komen for the Cure® under grant IMG0403022.

3. No linear combination is significantly correlated with both the active-CA and non-CA subgroups simultaneously, however
   - This suggests that the cancer-status predictors are influenced by differences between the tumor phenotypes of the subgroups





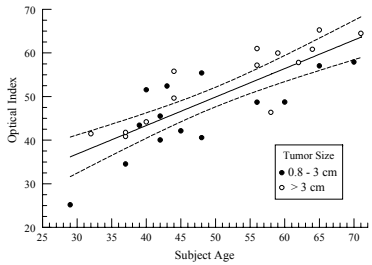




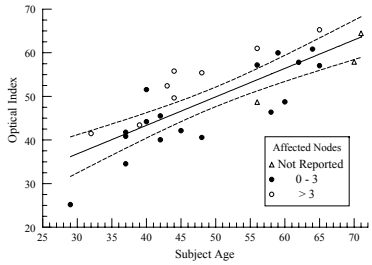
Conclusions

1. Predictors derived from diffuse optical tomography time series successfully discriminate among breast-CA and non-CA subjects because they consider dynamic features that are correlates of breast-tumor phenotypes
2. The mathematical correlations observed here suggest the feasibility of developing noninvasive optical techniques to estimate the properties of interest (*e.g.*, tumor grade, Her2 status, etc.)

4. Hypothesis: cancer-status predictors are influenced by differences between the tumor phenotypes of the subgroups
   - Additional evidence in favor of the hypothesis is the observed tendency for subjects with more advanced cancer to have higher values for the age-correlated combination of tumor predictors



Size	No. of Values Above Regression Line	No. of Values Below Regression Line	Fisher's Exact Test
≤ 3 cm	5	9	p = 0.024
> 3 cm	11	3	



# of Metastatic Lymph Nodes	No. of Values Above Regression Line	No. of Values Below Regression Line	Fisher's Exact Test
0-3	7	10	p = 0.006
> 3	8	0	

5. Follow-up computations on the cancer-subgroup data yielded combinations of tumor predictors having statistically significant correlations with tumor size, tumor grade, metastasis, and ER, PR and Her2 scores (r = 0.47-0.60)

Variable	Tumor Size (N = 27, n = 4)	Axillary Node Status[a] (N = 25, n = 4)	ER (N = 27, n = 3)	PR (N = 26, n = 4)	Her2 (N = 27, n = 4)	Grade (N = 27, n = 6)
Constant	2.2±1.7 (p = 0.21)	1.5±0.8 (p = 0.066)	3.8±1.0 (p = 0.001)	3.0±0.9 (p = 0.002)	2.6±0.9 (p = 0.008)	2.3±0.3 (p = 6.0e-7)
MpM 1	-3.6±2.5 (p = 0.16)	-2.8±2.0 (p = 0.18)	-2.0±0.9 (p = 0.049)	-1.8±0.8 (p = 0.027)	—	1.1±0.6 (p = 0.079)
MpM 2	—	1.9±1.3 (p = 0.17)	—	—	—	-0.43±0.37 (p = 0.25)
MpM 3	—	—	—	—	—	-0.91±0.44 (p = 0.052)
MpM 5	—	—	—	—	-0.93±0.72 (p = 0.22)	—
MpM 6	—	—	2.0±0.9 (p = 0.037)	—	0.86±0.83 (p = 0.31)	-0.49±0.29 (p = 0.10)
MpM 7	-8.7±4.9 (p = 0.09)	—	—	3.6±2.0 (p = 0.085)	1.3±1.3 (p = 0.32)	—
MpM 8	9.2±4.7 (p = 0.064)	—	—	-2.4±1.9 (p = 0.23)	—	1.8±0.6 (p = 0.006)
MpM 14	6.1±2.8 (p = 0.040)	2.9±1.5 (p = 0.15)	—	—	-2.5±1.1 (p = 0.030)	-0.91±0.57 (p = 0.12)
MpM 15	—	-1.3±0.9 (p = 0.14)	—	—	—	—
MpM 18	—	—	—	—	—	—
MpM 19	—	—	-2.4±1.0 (p = 0.030)	-1.5±0.8 (p = 0.073)	—	—
Correlation Coefficient	0.510 (p = 0.007)	0.483 (p = 0.014)	0.510 (p = 0.007)	0.540 (p = 0.004)	0.468 (p = 0.014)	0.804 (p = 8.5e-4)

N: number of subjects, *n*: number of regressors considered in the least-squares fit. MpM = Multiparameter Metric. The value of *n* presented here is the number of metrics that yields the optimal tradeoff between maximizing the value of the correlation coefficient estimate and minimizing its standard error.

a) The numerical value of this index is 0, 1, 2, or 3, depending on whether breast cancer cells are found in 0, 1, 2-3, or more than 3 lymph nodes, respectively.