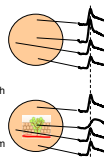
Derivation and validation of metrics for breast cancer screening from diffuse optical tomography imaging data

R.L. Barbour[1,2], H.L. Graber[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

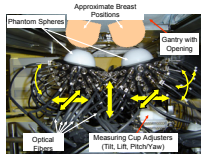
Introduction

- Dynamic Near Infrared Optical Tomography (DYNOT)
  - Provides measures of relative concentrations of hemoglobin (Hb)
    - Oxygenated, deoxygenated, total
  - Noninvasive functional imaging *in vivo*
    - Exogenous contrast agents not required
- Growth of solid tumors frequently accompanied by:
  - Marked changes in the vascular supply sustaining tumor growth
    - State of impaired perfusion
    - Relatively hypoxic environment
- To image the preceding, we developed:
  - A dual-breast diffuse optical tomography (DOT) imaging system
  - Capable of simultaneous bilateral measurements



- Healthy Breast
  - Coherent response to provocation
- Tumor Breast
  - Sluggish perfusion
  - Reduced response to natural effectors
  - Elevated Hb levels

Measuring Head (1st-Generation) for Simultaneous Dual-Breast Imaging



- DYNOT simultaneous dual-breast measuring head
  - Each optical fiber (31 per breast) is both source and detector for NIR light
  - Data collected from 124 channels (62 fibers × 2 wavelengths) in parallel
  - Sources location is time-multiplexed
  - Measurement rate: 2 complete image frames (> 8×10^3 data points) per second

Physiological Hypotheses

- There are mechanisms by which a cancerous tumor's "volume of influence" may be appreciably larger than the tumor itself
  - This is a consequence of well-characterized differences between vasculature in cancerous solid tumors, and in healthy tissue (or in non-cancer pathologies)
- A key to increasing diagnostic power is comparing detector and image data between the two simultaneously examined breasts
  - Vascular responses under autonomic control should be similar in the two breasts
  - Responses under local control (*e.g.*, autoregulation) also should be similar, if both breasts are healthy
- If one breast has a cancerous tumor and the other doesn't, what macroscopic differences could we expect?
  - 1: Increased amplitudes for vasomotor rhythms, in the tumor-bearing breast (TBB)
    - Owing to hypoxic environment of many solid tumors
  - 2: Greater temporal correlation across the breast volume, and greater spatial homogeneity, in the tumor-free breast (TFB)
  - 3: Abnormal response, in the TBB, to events that stress the microvasculature
- Therefore, three categories of diagnostic metrics will be considered in the clinical study
  - Each is devised to reveal one of the three types of expected difference between the TBB and TFB
  - **Group 1:** Indices of resting vasomotion amplitude
    - Computed from resting-state measurement data
    - Are sensitive to enhanced vasomotion, possibly associated with tumor angiogenesis
  - **Group 2:** Index of spatially coordinated dynamics
    - Computed from resting-state measurement data
    - Is sensitive to differences in timing of blood delivery within each breast
  - **Group 3:** Measures of pressure-induced blood volume and oxygenation shifts
    - Computed from data collected during Valsalva maneuver
    - Are sensitive to venous congestion and delayed reperfusion, which can lead to tumor hypoxia
    - In the TBB, compared to the contralateral TFB, one would expect to see:
      - Increased oxygen desaturation of Hb
      - Increased blood volume change
      - Increase in time needed for recovery

Clinical Study Design

1. Subject population
   - Retrospective – subjects whose data is used for derivation of breast-cancer diagnostic metrics
   - Prospective — subjects whose data is used for testing and validation

Subject Group	Breast Pathology Status	*N*	Age (yr) [mean ± SD]	BMI (kg·m^{-2}) [mean ± SD]	Tumor Size [largest dimension]	Clinical Description
Retrospective	Active CA	14	47.9 ± 12.3	28.7 ± 5.3	1 ≤ 3 cm 4 > 3 cm	10 ductal carcinoma 1 ductal & lobular carcinoma 1 mucinous carcinoma 1 metastatic CA 1 inflammatory CA 4 Grade 2, 9 Grade 3, 1 had no biopsy data
	Prior CA	3	50.7 ± 9.4	30.4 ± 0.5	—	All had lumpectomies 2-3 yr prior to NIRS study
	Pre-CA	0	—	—	—	—
	Non-CA Pathology	11	45.7 ± 5.6	28.7 ± 5.5 (*N* = 7)	—	3 fibrocystic disease 4 breast cyst 1 axillary cyst 2 benign breast lumps 1 breast reduction surgery
	No History of Breast Pathology	9	41.6 ± 10.0	30.3 ± 7.2	—	—
Prospective	Active CA	14	51.4 ± 10.9	30.4 ± 4.5	5 ≤ 3 cm 9 > 3 cm	13 ductal carcinoma 1 axillary adenocarcinoma with mammary duct ectasia and hyperplasia 3 Grade 2, 11 Grade 3
	Prior CA	4	60.8 ± 9.3	25.5 ± 1.7	—	3 prior ductal carcinoma 1 prior mucinous carcinoma All had lumpectomies 2-6 yr prior to NIRS study
	Pre-CA	4	53.5 ± 3.4	29.0 ± 4.1 (*N* = 3)	—	2 DCIS 1 atypical ductal hyperplasia 1 extremely dense breasts
	Non-CA Pathology	6	43.7 ± 8.4	26.6 ± 4.9 (*N* = 4)	—	1 cystic disease 2 fibrocystic changes 1 fibrosis 1 benign breast lump 1 breast reduction surgery
	No History of Breast Pathology	8	44.0 ± 6.8	30.5 ± 8.9	—	—

Subject Group		Age (years)				BMI (kg·m^{-2})			
		n_1	Mean	SD	Range	n_2	Mean	SD	Range
Active Breast Cancer (CA)	Retrospective / Training Group	14	47.9	12.3	29-70	14	28.7	5.3	21.6-43.9
	Prospective / Validation Group	14	51.4	10.9	37-71	14	30.4	4.5	22.7-38.1
Non-CA	Retrospective / Training Group	23	44.7	8.6	26-62	19	30.1	6.1	18.4-44.4
	Prospective / Validation Group	22	48.7	10.0	30-69	18	28.2	6.6	21.2-48.5

2. Analysis of dual-breast DOT image time series
   - 4-D (volume + time) data sets are reduced to the following scalar values:

Experimental Condition	Tumor-Associated Phenotype	Scalar Metric
Resting State	Angiogenesis	1. $SMTSD = \sum\sqrt{\sum x(r,t)^2/N_t}/N_v$
		2. $SSDTSD = \sqrt{\sum\left[\sqrt{\sum x(r,t)^2/N_t} - SMTSD\right]^2/N_v}$
		3. $TMSSD = \sum\sqrt{\sum\left[x(r,t)-\sum x(r,t)/N_t\right]^2/N_t}/N_v$
		4. $TSDSM = \sqrt{\sum\left[\sum x(r,t)/N_t\right]^2/N_v}$
		5. $TSDSSD = \sqrt{\sum\left[\sqrt{\sum\left[x(r,t)-\sum x(r,t)/N_t\right]^2/N_t} - TMSSD\right]^2/N_v}$
	Spatial Coordination	6. $SC(t_{ref}) = 100\sqrt{\sum\left[N_f u(f,t_{ref}) - \sum u(f,t_{ref})/N_v\right]^2/N_f}\Big/\sum u(f,t_{ref})$, where: $x(r,t) \to M(r,f,t)$ (Complex Morlet wavelet decomposition), $U(r,f,t_{ref}) = N_t \cdot M(r,f,t,t_{ref})/\sum M(r,f,t,t_{ref})$ (Normalize to unit mean value, in the frequency bands of interest) $u(f,t_{ref}) = \sum U(r,f,t_{ref})/N_v$ (Average over the image volume).
Evoked Response	Hypoxia	7. $A = \sum_{t_1}^{t_2}\left[y(t) - \left\{[y(t_2)-y(t_1)]/(t_2-t_1)\right\}t + \left\{[ty(t_1)-t_1y(t_2)]/(t_2-t_1)\right\}\right]/(t_2-t_1)$
		8. $R = \max_{t_1<t<t_2}[y(t)] - \min_{t_1<t<t_2}[y(t)]$

- Differences between metric values, for each subject's two breasts, are calculated as:
  - Tumor minus non-tumor for training-set cancer subjects
  - Left minus right for training-set non-cancer subjects, and for validation-set subjects
- Each metric is converted into six candidate diagnostic parameters, by normalizing the inter-breast difference in a variety of ways:
  - Difference divided by larger, smaller, or average value of the two individual-breast values
  - Difference multiplied by larger, smaller, or average of the individual-breast values

2. Analysis of dual-breast DOT image time series [continued]
   - Assessment of sensitivity, specificity, positive and negative predictive values
     - Univariate: parametric and nonparametric tests for difference between means of CA and non-CA subgroups of the training set
     - Multivariate: binary logistic regression (BLR)
       - Use leave-out-one cross-validation (LOOCV) to determine sensitivity to idiosyncrasies of the training-set subjects
       - Validate the predictors that are most successful with respect to the retrospective group, by applying them to the data for the prospective group
   - Aggregates (averages) for 1 or more multivariate predictors
     - The overall number of aggregates is ${}_{22}C_1 + {}_{22}C_2 + \ldots + {}_{22}C_{21} + {}_{22}C_{22} = 4{,}194{,}303$
     - For each subject, the aggregate breast-cancer probability is the average of the separate probabilities according to the multivariate predictors
     - For sensitivity and specificity calculations, a breast-cancer probability of 0.5 is taken as the diagnostic threshold

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.

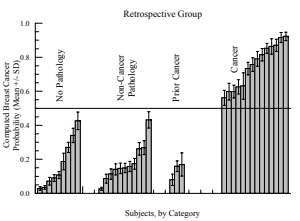
Results

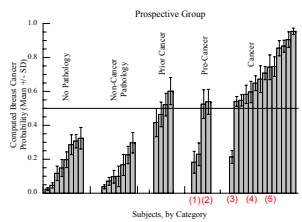
1. Diagnostic accuracy for multivariate predictors
   - As determined by BLR-LOOCV computations, these had statistically highly significant discriminatory ability and robustness
   - Predictors 1-13 are hypothesis-based: physiological premises (above) guided the selection of which univariate metrics to include
   - Predictors 14-22 are data-driven: a backward-elimination algorithm was used to determine which combinations of univariates yield the "best" multivariates, without regard to biological significance

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	$6/6_b$		50	90	85.7	93.3
4	7,8	7,8		80	100	71.4	92.3
5†	2,3,6_t	6_t	1	85.7	95	57.1	86.7
6		1-5,6_t,8_b		71.4	90	57.1	93.3
7	1-3	5,6,6_t	1-3	85.7	95	78.6	100
8	1-3	5,6,6_t	1-3,5	100	90	78.6	100
9	1,8		3	80	90.9	78.6	92.3
10		6_t,8		80	90.9	71.4	84.6
11	8	$6,6_t$,8		80	90.9	71.4	92.3
12†	6_t	8	1	90	81.8	57.1	84.6
13†	2,5,6,8	4,6_t,8	1,3	90	90.9	64.3	76.9
14	3			50	80	64.3	93.3
15		6_t		57.1	95	57.1	100
16	8			90	90.9	71.4	84.6
17	7,8			80	90.9	71.4	84.6
18	3	6_b		78.6	80	85.7	100
19	3,5,6_t	6_b	4	100	95	78.6	93.3
20	8	6_t		80	81.8	64.3	84.6
21		4	3	90	90.9	57.1	61.5
22		6_t	4	80	90.9	71.4	69.2

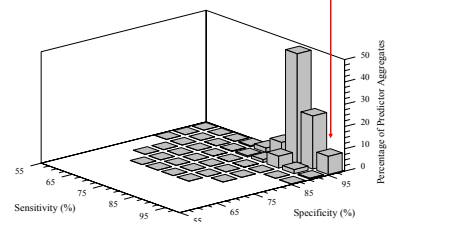
† The Difference*Maximum formulation is used for these predictors, and Difference/Maximum for all the others

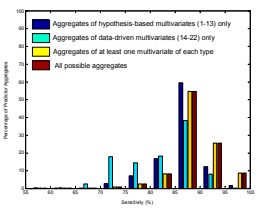
3. Average breast-CA probability, for each subject, across the 196,639 highly successful aggregates
   - (1) Low-grade DCIS, and previously diagnosed DCIS not confirmed upon re-exam
   - (2) High-grade DCIS, and dense, cystic breasts
   - (3) Small (< 1 cm) tumor, near the chest wall
   - (4) Adenocarcinoma in axillary lymph node
   - (5) Lumpectomy 8 yr. prior to NIRS exam, relapse 1 yr. after NIRS exam
4. Correlations between computed breast-CA probability and demographic variables: age and BMI
   - Only the No Pathology and Non-CA Pathology sub-groups were considered here
   - Results are suggestive of a trend, in all cases, but sample size is too small for statistical significance

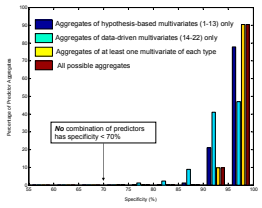




2. Diagnostic accuracy for **aggregates of** multivariate predictors
   - There are more than 3×10^5 aggregates that yield sensitivity and specificity both > 95%
   - Of these, there are 196,639 that correctly identify:
     - All Non-CA Pathology and No History of Breast Pathology subjects, in both the retrospective and prospective groups
     - All Active CA subjects in the retrospective group
     - All but one Active CA subject in the prospective group







Conclusions

1. Prospective-group results show that the derived multivariate metrics do operate on features of the image data that are related to breast-cancer status, and not to any idiosyncratic properties of the retrospective group.
2. The observation that the inter-breast differences have the predicted directionality implies that our assumptions, regarding the effects of tumor development and growth on the dynamic properties of the vasculature, are largely correct
3. The successful application of metrics derived by spatial integration of DOT image data implies that the vascular correlates of tumor growth and development are present over a volume that extends a substantial distance beyond the histological margins of the tumor.
4. The preceding conclusions, plus the suggestive results from preliminary examination of the trends between breast-CA probability and age or BMI, indicate that the relationships between the DOT metrics and other data types (demographics, histology, tumor-receptor findings) should be more fully explored (NIH01-104).