10.1109/ULTSYM.2014.0476

Evaluation of classification strategies using quantitative ultrasound markers and a thyroid cancer rodent model

Maria Luisa Montero[1], Omar Zenteno[2],Benjamin Castaneda[2], Michael Oelze[3] and Roberto Lavarello[2]
[1]Laboratorio de Estadística, Departamento de Ciencias
Pontificia Universidad Católica del Perú, San Miguel, Lima 32, Perú.
[2]Laboratorio de Imágenes Médicas, Departamento de Ingeniería
Pontificia Universidad Católica del Perú, San Miguel, Lima 32, Perú.
[3]Bioacoustics Research Laboratory, Department of Electrical and Computer Engineering
University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA.

Abstract— The incidence rate of diagnosed thyroid cancer has increased over the last decades. Although ultrasonic imaging has increased the malignancy detection rate, current ultrasonography markers do not provide a sufficient level of diagnostic accuracy to replace the fine needle aspiration biopsy. Recently, studies have reported that significant differences were observed in the values of quantitative ultrasound (QUS) parameters derived from a thyroid cancer rodent model between normal/benign and malignant tissues. In the present study, the performance of a multi-parametric classification for the differentiation of thyroid cancer in this rodent model has been evaluated. The experimental database consisted of 32 mice having different predispositions to developing thyroid abnormalities; 6 of them developed thyroid cancer papillary carcinoma (PTC), 5 follicular variant papillary thyroid carcinoma (FV-PTC), 6 developed benign tumors (c-cell adenoma) and 15 did not develop any thyroid abnormalities. Backscattered data was obtained from excised thyroid tissues using a 40 MHz, f/3 single element transducer. A total of five QUS parameters were derived from the ultrasound data: two from backscatter coefficients (i.e., the effective scatterer diameter (ESD) and effective acoustic concentration (EAC)), two from envelope statistics (i.e., the μ and k parameters), and one from ultrasound attenuation (i.e., attenuation coefficient slope). A two-class classification between normal/benign and malignant cases was performed using linear discriminant analysis with both one- and two-dimensional feature spaces. When using a two-dimensional feature space, it was found that the combination of EAC and 10/μ resulted in both a sensitivity and specificity of 100%.

Keywords—Quantitative ultrasound, tissue characterization, thyroid cancer, linear discriminant analysis.

I. INTRODUCTION

Thyroid cancer incidence rates have been steadily rising over the past decades. In the United States, between 4% and 7% of the general population have clinically palpable nodules [1]. The American Cancer Society estimates for 2014 about 62,980 new cases of thyroid cancer [2]. Recent studies suggest that the rise in the number of detected cases may not be fully explained by the use of improved screening technologies (i.e., high frequency ultrasound) [3]. Medical practice tends to prioritize patient health, leading clinicians to treat some detected asymptomatic benign tumors as malignant. This may lead to unnecessary surgery, which potentially has personal and societal consequences [4].

Recent studies suggest that quantitative ultrasound (QUS) parameters have the potential to discriminate between cancerous and normal thyroids. Lavarello *et al.,* [5] estimated parameters derived from backscatter coefficients (i.e., the effective scattering diameter (ESD) and the effective acoustic concentration (EAC)) from a rodent model of follicular variant papillary thyroid carcinoma (FV-PTC) *ex vivo*. Statistically significant differences were observed between the FV-PTC and normal thyroids when using both the ESD and EAC. This study was later expanded to include two more classes (i.e., papillary thyroid carcinoma (PTC) and c-cell adenoma) [6]. In addition, QUS spectral estimates were complemented with information extracted from the envelope of the backscattered signal, the k parameter (ratio of coherent to incoherent signal energy) and the μ parameter (number of scatterers per resolution cell). Significant differences were observed between cancerous and normal thyroids based on the ESD, EAC and μ parameters. More recently, Zenteno *et al.,* [7] expanded the results in [6] by estimating the attenuation coefficient slopes (ACSs) of the same samples. Statistically significant differences were found between the ACS values of c-cell adenoma and normal and malignant thyroids, and between the ACS values of PTC and FV-PTC thyroids.

In spite of the encouraging results listed above, none of the previous studies attempted to classify thyroid tissues based on the estimated QUS parameters. Therefore, the objective of this study is to complement the findings in [5-7] by performing a two-class classification between normal/benign and malignant tissues from a rodent thyroid cancer animal model using linear discriminant analysis in both one- and two-dimensional feature spaces.

II. METHODS

A. Animal modes and data adquisition:

The animal model consisted of mice that had introduced a dominant negative mutant thyroid nuclear receptor gene. As a

result, many developed variations of papillary metastatic thyroid tumors or presented a high concentration of parafollicular cells (c-Cells), which are known to potentially lead to medullary carcinoma. When the thyroid showed indications of lesions or signs of hyperplasia, both thyroid lobes were extracted along with a portion of the trachea and placed in a tank of degased 0.9% saline for ultrasonic scanning.

A sample set of 32 transgenic mice thyroids was scanned using a 40-MHz weakly-focused (f/3) single element transducer using a Olympus Panametrics 5900 pulser/receiver (Olympus NDT, Waltham, MA) for excitation. The radiofrequency (RF) data was captured by a PC using a 14-bit UF3-4121 A/D card with 250 MHz sampling frequency (Strategic Test Corporation, Woburn, MA).

Each thyroid sample was scanned using a micro-positioning system controlled by custom LabVIEW (National Instruments, Austin, TX) software. Several slices were taken at 0.2-0.4 mm apart across the thyroid sample from the top view. The slices which fully contained both thyroid lobes were selected resulting in an average of six usable ultrasound slices per rodent. Each scan line within a slice had a separation of 50 μm. After scanning, each sample was excised, formalin-fixed and paraffin-embedded for histologic evaluation by light microscopy.

The histopathological classification reported that the sample consisted of six papillary thyroid carcinoma, five follicular variant pattern papillary carcinoma, six c-Cell positives, and fifteen mice which did not develop thyroid nodules. The latter were labeled as normal and used as a control group.

B. Spectral based quantitative ultrasound parameters

Three spectral-based parameters were derived from the estimated backscatter coefficient [8] by applying a spherical Gaussian model to the data [9] and the attenuation coefficient estimated using the method in [6]:

1. The ESD in μm, which represents the typical size of unresolvable scatterers within the insonified region.

2. The EAC in dB(mm^{-3}), which is proportional to the product of the number of scatterers per unit of volume and the square of the fractional acoustic impedance mismatch between the scatterers and background.

3. The ACS in dB/cm-MHz, which is the slope of the linear fit of the attenuation coefficient vs. frequency.

C. Envelope based quantitative ultrasound parameters

Two envelope statistics parameters were estimated assuming a homodyned K distribution [10]:

1. The *k* parameter (unitless), which quantifies the ratio of the coherent scattering signal to the incoherent scattered signal. If scatterers are regularly spaced or large single scatterers are present, *k* will increase.

2. The μ parameter (unitless), which provides an estimate of the number of scatterers per resolution cell. If the resolution cell of the imaging system can be estimated, then an estimate of the number density of scatterers can be obtained.

D. Statistical analysis of independent variables

For classification of benign/normal and malignant cases, all QUS estimated parameters were statistically evaluated. The mean, standard deviation and standard error (i.e., the ratio between the standard deviation and the sample size) were calculated for the complete dataset and for each class to analyze their significance to the classification analysis.

Linear discriminant analysis (LDA) custom software using SPSS (IBM, Armonk, NY) was implemented. LDA seeks to find a linear combination of features which may differentiate between two or more classes of objects or events. In the particular case of two groups, a linear function of the independent variables is derived such that the variance of the function would be maximum between the two classes (i.e., benign/normal and malign) and minimum for each class [11]. LDA was applied to all variables independently to measure the classification performance of each single QUS parameter. In addition, linear discriminant analysis was performed with all pairwise combinations of the five parameters in order to evaluate the discriminant potential of combinations of the QUS markers. The performance of LDA classification was measured in terms of the sensitivity, specificity and accuracy using a leave-one-out (LOO) cross-validation approach.

III. EXPERIMENTAL RESULTS

Table I presents the mean, standard deviation and standard error values of the individual variables. The results in terms of sensitivity and specificity when using single QUS parameters are shown in Table II. In all cases no accurate classification was achieved. Specifically, both the ACS and *k* parameters presented a poor classifier performance between both considered classes.

TABLE I. **SUMMARY DESCRIPTORS OF VARIABLES**

		Mean	**Std. Dev.**	**Std. Err.**
ESD (μm)	Bening / Normal	21.24	1.77	0.39
	Malignant	17.02	1.37	0.41
	Total	19.79	2.61	0.46
EAC dB(mm^{-3})	Bening / Normal	54.93	2.79	0.61
	Malignant	60.92	2.36	0.71
	Total	56.99	3.90	0.69
ACS (dB/cm-MHz)	Bening / Normal	1.40	0.22	0.05
	Malignant	1.24	0.15	0.04
	Total	1.35	0.21	0.04
k	Bening / Normal	0.48	0.09	0.02
	Malignant	0.53	0.05	0.02
	Total	0.50	0.08	0.01
μ	Bening / Normal	1.79	1.59	0.35
	Malignant	2.92	0.52	0.16
	Total	2.18	1.42	0.25

TABLE II. PERFORMANCE MEASUREMENTS OF SINGLE VARIABLE CLASSIFICATION

	Sensitivity (%)	Specificity (%)	Accuracy (%)
ESD	100.0	81.0	87.5
EAC	81.8	85.7	84.4
ACS	63.6	61.9	62.5
k	81.8	71.4	75.0
μ	81.8	85.7	84.0

A two-variable classification with all possible paired combinations of the five QUS parameters was also performed. The sensitivity, specificity and accuracy of the test are presented in Tables III, IV, and V, respectively. Each table contains the four possible combinations for each parameter in a two entry format.

The sensitivity results in Table III reveal a good performance (i.e., higher than 90%) for some parameter combinations. In particular, 100% sensitivity was obtained with the ESD-ACS, EAC-ACS, ESD-*k* and ESD-*μ* combinations, and 90.9% sensitivity was achieved with the EAC-*μ* combination. On the other hand, the results presented in Table IV reveal that the combinations above did not necessarily allow obtaining the best specificity. In particular, only the combinations conformed by EAC-*μ* and EAC-*k* achieved a specificity higher than 90%. Finally, the classification accuracy results in Table V revealed that the ESD-ACS, ESD-*k*, EAC-ACS and EAC-*μ* combinations performed better on this particular metric.

TABLE III. SENSITIVITY OF A TWO-VARIABLE CLASSIFICATION

	ESD	EAC	ACS	k	μ
ESD	×	81.8	100.0	100.0	100.0
EAC	81.8	×	100.0	81.8	90.9
ACS	100.0	100.0	×	81.8	81.8
k	100.0	81.8	81.8	×	72.7
μ	100.0	90.9	81.8	72.7	×

TABLE IV. SPECIFICITY OF A TWO-VARIABLE CLASSIFICATION

	ESD	EAC	ACS	k	μ
ESD	×	85.7	85.7	85.7	81.0
EAC	85.7	×	85.7	90.5	90.5
ACS	85.7	85.7	×	57.1	76.2
k	85.7	90.5	57.1	×	85.7
μ	81.0	90.5	76.2	85.7	×

TABLE V. ACCURACY OF A TWO-VARIABLE CLASSIFICATION

	ESD	EAC	ACS	K	μ
ESD	×	84.4	90.6	90.6	87.5
EAC	84.4	×	90.6	87.5	90.6
ACS	90.6	90.6	×	65.6	78.1
k	90.6	87.5	65.6	×	81.3
μ	87.5	90.6	78.1	81.3	×

From the results in Tables II-V, the EAC-*μ* combination provided the best overall performance. Therefore, this combination was selected for further analysis. The scatter plot of the data points considering these parameters is shown in Fig. 1.

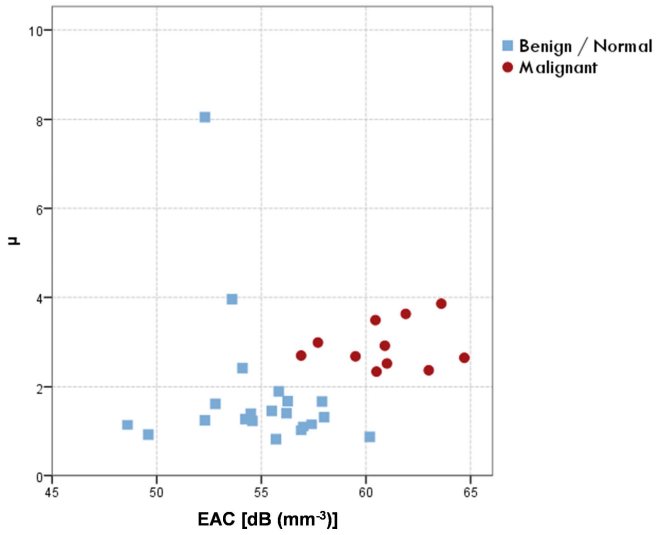

Fig. 1. Scatter plot of the effective acoustic concentration (EAC) versus *μ* parameter for the two classes under analysis.

As can be observed in Fig. 1, there were only two samples which did not adjust to the bivariate behavior of the rest of the Benign/Normal class. Therefore, a transformation of the feature space that could potentially move the point cloud closer to these values was applied. The selected transformation was EAC-10/*μ*. The corresponding scatter plot is presented in Fig 2. The LDA classification was applied again considering all pairwise combinations of 10/*μ* and the other QUS parameters. The results are shown in Table VI. A perfect classification was achieved for the EAC-10/*μ* combination.

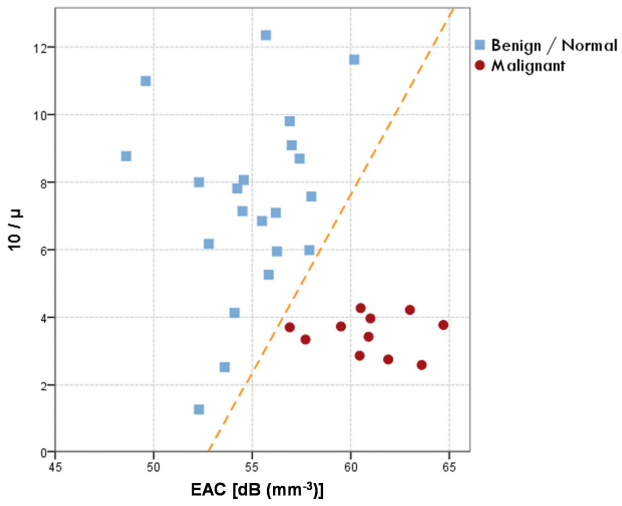

Fig. 2. Scatter plot of the effective acoustic concentration (EAC) versus 10/μ for the two classes under analysis.

TABLE VI. **PERFORMANCE MEASUREMENTS OF A TWO VARIABLE CLASSIFICATION WITH 10/μ**

	Sensitivity (%)	**Specificity (%)**	**Accuracy (%)**
ESD-10/μ	100.0	90.5	93.8
ACS-10/μ	100.0	81.0	87.5
k-10/μ	100.0	81.0	87.5
EAC-10/μ	1000	100.00	100.0

Additional evaluation of the classifier using the EAC-10/μ combination was performed using a non-exhaustive leave-four-out cross-validation. In particular, all the four groups (i.e., Normal, PTC, FV-PTC and c-cell) were considered as clusters and one element was randomly extracted from each cluster to construct a total of five disjoint validation sets. The discriminant function was estimated using each validation set as test data and the remaining cases as training data. This resulted in an accuracy of 100% for the training data and 95% for the test data.

IV. DISCUSSION

High-frequency QUS spectral- and envelope-based parameters provided information with the ability to differentiate cancerous thyroids from normal/benign thyroids in a rodent model. In all cases, the highest accuracy achieved using a single QUS parameter was 87.5%. In contrast, the performance of the two-dimensional feature space analysis was significantly better than one-dimensional feature space and allowed obtaining perfect classification accuracy.

The best performance was obtained when using EAC and 10/μ as a feature space. Both EAC and μ parameters were typically observed to increase from normal/benign thyroid to cancerous thyroids. This may be expected because both variables depend on the number density of scatterers present in the medium. Cancerous thyroids are characterized by cellular hyperplasia. Assuming that the cells contribute significantly to the backscattered signal, the increment in the number of cells result may be related to an increase in the estimated number density. This can be directly correlated with the increased proliferation of cells observed in histopathology analysis of the cancerous thyroids [7].

The main objective of the present work was to determine if it was possible to classify between normal/benign and malignant thyroids using QUS parameters as the input of the analysis. Therefore, the mice thyroid glands were scanned *ex vivo* to eliminate effects caused by intervening tissues (i.e., heterogeneous attenuation profiles, aberration, clutter, blood flow). Thyroid nodules would be examined *in vivo* in humans, and therefore the aforementioned effects on classification accuracy would need to be assessed.

Even though the linear classification approach is not as robust as other non-linear classifiers, a high level of accuracy was achieved. However, the use of a larger population sample size is required in order to understand the ability of QUS analysis to differentiate among different types of malignant tissues. The use of a larger number of labels may also require exploring different classification methods on feature spaces similar to the ones used in this preliminary study.

V. CONCLUSION

Although further studies in human subjects are required to assess the clinical implications of these observations, the results of this study suggest that multi-parametric QUS may prove valuable for the diagnosis of thyroid cancer.

ACKNOWLEDGMENT

The authors would like to thank Rita Miller, William Ridgway and Emily Hartman for their assistance in scanning and animal handling. This work was supported by NIH Grant R21-CA139095, PUCP Grants DGI 2012-0149 and 2013-0131, and funds from the PUCP Graduate School.

REFERENCES

[1] L. Morris, A. Sikora, T. Tosteson and L. Davies, "The increasing incidence of thyroid cancer: the influence of access to care," *Thyroid*, vol. 23, no. 7, pp.885-891, 2013.

[2] American Cancer Society. Cancer facts and figures 2014. Atlanta, GA: American Cancer Society; 2014.

[3] L. Davies, M.Ouellette, M.Hunter and G.Welch. "The increasing incidence of small thyroid cancers: Where are the cases coming from?," *Laryngoscope*, vol. 120, no.12 , pp.2446-2450, 2010.

[4] F. Yoo, I. Chaikhoutdinov, R. Mitzner, J. Liao, and D. Goldenberg, "Characteristics of incidentally discovered thyroid cancer," *JAMA Otolaryngology–Head & Neck Surgery*, vol. 139,no. 11,pp. 1181-1186, 2013

[5] R. Lavarello, W. Ridgway, S. Sarwate and M. Oelze, "Imaging of follicular variant papillary thyroid carcinoma in a rodent model using spectral-based quantitative ultrasound techniques," *in Proceedings of the IEEE International Symposium on Biomedical Imaging*, pp. 732-735, 2013.

[6] R. Lavarello, W. Ridgway, S. Sarwate and M. Oelze, "Characterization of thyroid cancer in mouse models using high-frequency quantitative ultrasound techniques," *Ultrasound in Medicine & Biology*, vol. 39, no 12 pp. 2333–2341, 2013

[7] O.Zenteno, W. Ridgway, S. Sarwate, M. Oelze and R.Lavarello, "Ultrasonic attenuation imaging in a rodent thyroid cancer model," in *Proceedings of the IEEE Ultrasonics Symposium*, pp. 88-91, 2013.

[8] R. Lavarello, G. Ghosaal and M.Oelze, "On the estimation of backscatter coefficients using single-element focused transducers," *Journal of the Acoustical Society of America*, vol 129, no. 5, pp. 2903-2911,2011.

[9] M. Oelze, J. F. Zachary and W. D. O'Brien, Jr., "Characterization of tissue microstructure using ultrasonic backscatter: Theory and technique for optimization using a Gaussian form factor," *Journal of the Acoustical Society of America*, vol. 112, no. 3, pp. 1202-1211, 2002.

[10] D. Hruska and M. Oelze, "Improved parameter estimates based on the homodyned K distribution," *IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control*, vol. 56,no. 11, pp. 2471-2481, 2009.

[11] R. A. Fisher,"The use of multiple measurements in taxonomic problems," *Annals of Eugenics*,vol. 7, no.2, pp.179–188, 1936.