Is Contrast Enhanced Ultrasonography a useful tool in a beginner’s hand? How much can a Computer Assisted Diagnosis prototype help in characterizing the malignancy of focal liver lesions?

Tudor Voicu Moga¹, Alina Popescu¹, Ioan Sporea¹, Mirela Danila¹, Ciprian David², Vasile Gu³, Nicoleta Iacob³, Gratian Miclaus³, Roxana Sirli¹

¹Department of Gastroenterology and Hepatology, “Victor Babes” University of Medicine and Pharmacy, ²Electronics and Telecommunications Faculty, “Politehnica” University, ³Department of Anatomy and Embryology, “Victor Babes” University of Medicine and Pharmacy, Timișoara, Romania

Abstract

Aim: Contrast enhanced ultrasound (CEUS) improved the characterization of focal liver lesions (FLLs), but is an operator-dependent method. The goal of this paper was to test a computer assisted diagnosis (CAD) prototype and to see its benefit in assisting a beginner in the evaluation of FLLs. Material and method: Our cohort included 97 good quality CEUS videos [34% hepatocellular carcinomas (HCC), 12.3% hypervascular metastases (HiperM), 11.3% hypovascular metastases (HipoM), 24.7% hemangiomas (HMG), 17.5% focal nodular hyperplasia (FNH)] that were used to develop a CAD prototype based on an algorithm that tested a binary decision based classifier. Two young medical doctors (1 year CEUS experience), two experts and the CAD prototype, reevaluated 50 FLLs CEUS videos (diagnosis of benign vs. malignant) first blinded to clinical data, in order to evaluate the diagnostic gap beginner vs. expert. Results: The CAD classifier managed a 75.2% overall (benign vs. malignant) correct classification rate. The overall classification rates for the evaluators, before and after clinical data were: first beginner-78%; 94%; second beginner-82%; 96%; first expert-94%; 100%; second expert-96%; 98%. For both beginners, the malignant vs. benign diagnosis significantly improved after knowing the clinical data (p=0.005; p=0.008). The expert was better than the beginner (p=0.04) and better than the CAD (p=0.001). CAD in addition to the beginner can reach the expert diagnosis. Conclusions: The most frequent lesions misdiagnosed at CEUS were FNH and HCC. The CAD prototype is a good comparing tool for a beginner operator that can be developed to assist the diagnosis. In order to increase the classification rate, the CAD system for FLL in CEUS must integrate the clinical data. Keywords: Contrast Enhanced Ultrasound; beginner; expert; focal liver lesions; computer assisted diagnosis.

Introduction

The widespread and easy access made the ultrasound examination ubiquitous among medical doctors. Ultrasound is the first line imaging method for detecting focal liver lesions (FLLs). Contrast Enhanced Ultrasound (CEUS) has become a validated easy-to-use, minimum invasive imaging method, for the quantitative and qualitative evaluation of FLLs. The efficacy and the utility of this technique are emphasized in the European Federation of Societies for Ultrasound in Medicine and Biology (EFSUMB) guidelines published in 2004 [1], updated in 2008 [2] and in October 2012 [3]. There are also some multicenter studies that demonstrate the accuracy of CEUS in characterizing FLLs [4,5]. Both methods are still operator-dependent and few studies have been made to overcome this drawback, having in mind the usefulness of a fast and reliable imaging diagnosis of a FLL [6]. It is quite hard to quantify the experience you need in order to be confident in using CEUS as a first line imaging method after a lesion was found in B-mode.

Computer assisted diagnosis (CAD) is still one of the major research topics in radiology and medicine,
having a pattern recognition software as a main feature, that calculates the likelihood of different lesions as being malignant or benign. This application is used in diagnostic radiology, for different imaging methods [computer tomography (CT), magnetic resonance imaging (MRI), conventional radiology (RX) and ultrasonography (USI)] in organs such as breast, lungs, brain, liver, kidney, colon, etc. There are only few studies regarding the utility of a CAD system for FLLs diagnosis by CEUS. Sugimoto et al [7] proposed two schemes for a CAD classifier of FLLs using CEUS, with encouraging results.

VueBox, from Bracco Italy, is a quantification toolbox for CEUS that can analyze General Imaging (GI) perfusion and the liver Dynamic Vascular Pattern (DVP) which moves one step forward the CAD in CEUS concept. Though, in order to use such a tool perfect examination conditions and good quality CEUS images and video clips are needed, thus returning to the initial operator dependence problem. The main purpose of a CAD system is to reduce the reading time and increase the diagnostic accuracy. We have to underline the differences in meaning of the words detecting/diagnosis. CAD in CEUS for FLLs will focus on the diagnosis of lesions characterized by CEUS (B-mode detection being a condition), detection software for the liver would engage other resources and in the same time other difficulties. Our study is oriented towards the utility of CEUS for an inexperienced user and to assess which tools can be used in order to improve the diagnosis accuracy of FLLs in daily practice.

The aim of our study was to create a CAD prototype for CEUS in evaluating FLLs, that can orientate the diagnosis to malignancy and which, together with the clinical data, can raise the diagnostic accuracy for the beginners. Through this, we wanted to highlight the most challenging lesions regarding FLLs interpretation, which a beginner might encounter during CEUS.

**Material and method**

A total of 97 CEUS video-cases of FLLs were used to develop the CAD prototype. From 97 FLLs, 33 (34%) were hepatocellular carcinomas (HCC), 12 (12.3%) hypervascular metastases (HiperM), 11 (11.3%) hypovascular metastases (HipoM), 24 (24.7%) hemangiommas (HMG) and 17 (17.5%) focal nodular hyperplasia (FNH). The video-cases were chosen from the database of our department according to quality parameters (obvious enhancing pattern, good ultrasound examination, good acoustic window). Contrast enhanced CT, MRI, or histology were available in each case to confirm the final diagnosis. The main focus was to analyze the five most frequent types of lesions: hepatocellular carcinoma, hypervascular metastases, hypovascular metastases, hemangioma and focal nodular hyperplasia. From the 97 FLLs, 33 of them were on cirrhotic liver (all the HCCs), the rest being detected on a normal liver on ultrasound examination. All selected video-cases were carefully chosen to have a typical CEUS pattern as described in EFSUMB guidelines [3].

The lesions were detected and assessed by standard ultrasound (B-mode) according to their number, size and location and then evaluated by CEUS with low mechanic index, convex probe (4C1), using SonoVue® (Bracco, Italy) as contrast agent, approx. 2.4 ml/lesion, followed by a 10-mL saline bolus. All examinations were made with a Siemens Acuson S2000 ultrasound machine with CEUS software and were performed by experienced operators (Level II – advanced and III – expert, according to the Romanian Society of Ultrasound in Medicine and Biology classification).

Each examination respected the standard EFSUMB protocol for CEUS [3]. The enhancement pattern of each lesion was followed up in 3 phases: arterial phase 10-30 sec, portal phase 30-120 sec and late phase >120 sec. All lesions were evaluated by four video-clips: one (10 second long) of conventional ultrasound study; and the other three, CEUS phases (arterial, portal and late phase), approx. 10-30 seconds for each phase. The lesions were characterized by comparing them with the normal enhancement of the surrounding liver, appearing as: homo-heterogeneous hyperenhanced – meaning the FLL’s enhancement pattern was homogeneous or presented irregularities; rim-like hyperenhancement – peripheral hyperenhancement, less than 1/4 of the tumor’s diameter; isoenhancement – similar enhancement pattern with the adjacent parenchyma at the same depth; hypoenhancement – the FLL enhances less than the adjacent parenchyma at the same depth; and wash-out – hyper or iso-enhancement in the arterial phase followed by hypoenhancement in the portal or late phases. The enhancement pattern of each lesion type, according to the EFSUMB 2012 guidelines, detailed in [3], was used as a template to characterize the FLLs.

**CAD prototype – Algorithm description**

Malignant/benign ruled based discrimination was obtained by computing the average relative intensity and by analyzing its evolution from the arterial to the late phase. These image attributes were considered in a user selected region of interest (ROI). The ROIs had been selected by one of the experts and included the entire lesion previously detected on standard ultrasound. By average relative intensity we understood the average color of the lesion area against the average color of the adjacent parenchyma. The region where the normal tissue was considered was taken.
adjacent to the ROI and had the same area. The equations related to the benign/malign decision are shown in Table I. In the left cell the average intensity is presented. We call relative average intensity the quantity obtained by dividing the average intensity of the lesion by the average intensity of the adjacent parenchyma (Table I- central cell). The criteria used to decide if we have the case of malignant FLL (Table I- right cell) are in concordance with the assumptions that malignant FLL presents wash out in the late phase. According to the average relative intensity (RAvgInt), the result was ≥1 in the arterial phase and ≤1 in the late phase for a malignant pattern, thus reducing the diagnosis to a binary decision (Fig 1 and Fig 2).

The spatial instability of the data, given by the patient’s respiratory movement and by the transducer movement, was surpassed by probabilistic filtering. Our hypothesis considers that the ROI was centered in most cases on the lesion, a fact that enables us to use the temporal median filter. The average relative intensity was measured for each frame. That value was the value considered for the final benign/malign decision.

**CEUS FLLs assessment**

In order to estimate or to quantify the diagnostic gap between a beginner vs. expert in establishing the diagnosis of malignancy in a FLL, two young medical doctors under fellowship in the Gastroenterology department, with CEUS experience less than one year, and two ultrasound experts, with more than 5 years’ experience in CEUS, were asked to reevaluate 50 CEUS video-cases (first 10 of each lesion type) from the same database. In the first phase, both beginners and seniors were asked to differentiate malignant vs. benign lesions blinded to clinical data. In the second phase, if misdiagnosed (failure to differentiate between malignant vs. benign), the clinical data were revealed (FLL detected on a normal liver, a cirrhotic liver or a patient with known neoplasia) and the video-cases were reassessed. Thirdly, the CAD prototype was used to randomly classify, the previously 50 FLLs seen by the experts and beginners, into malignant or benign, based on an algorithm that tested a binary decision based classifier.

**Statistical analysis**

Data are expressed as percentages of successful diagnosis for each examiner, before and after knowing relevant clinical data, and CAD, respectively. We used the Wilcoxon signed-ranks test to compare results within examiners and between examiners and CAD. The statistical analysis was performed for a 95% confidence interval, using IBM SPSS Statistics Package.

**Results**

The first beginner misdiagnosed 4 out of 10 FNHs, 3 out of 10 HCCs, 2 out of 10 HMGs, 1 out of 10 hypervascular metastases and 1 out of 10 hypovascular metastases (Fig 3). After knowing the clinical data, the first
beginner misdiagnosed 2/10 FNHs and 1/10 HCCs. The second beginner misdiagnosed 3/10 FNHs; 3/10 HCCs, 2/10 HMGs and 1/10 hypervascular metastases. After revealing the clinical data, the second beginner misdiagnosed 1/10 HCCs and 1/10 HMGs. The first expert misdiagnosed only 1 out of 10 FNHs, 1/10 HCCs and 1/10 HMGs. After revealing clinical data all lesions were correctly diagnosed. The second expert misdiagnosed 1 out of 10 FNHs and 1 out of 10 HCCs. After the clinical data were revealed, only 1 out of 10 FNH was still misdiagnosed.

The overall classification rates for the evaluators, before and after clinical data were revealed were: first beginner without knowing clinical data 78%; after knowing clinical data 94%, second beginner without clinical data 82%; with clinical data 96%. First expert: without clinical data 94%; with clinical data 100%. Second expert: without clinical data 96%; with clinical data 98%.

The CAD ruled based classifier managed an overall classification rate of 75.2%, malignant set classification rate-70.5% and benign set classification rate-80%.

All beginners’ malignant vs. benign diagnosis significantly improved after knowing the clinical data (table II). The expert was better than the beginner and better than CAD (table III). A comparison between the performance of experts and beginners, when CAD was favorable demonstrated that the beginner plus CAD can reach the expert diagnosis (table IV). The touchstones of the beginner’s evaluation were HCC and FNH.

### Discussions

The need for CAD software for the liver is necessary if we take into consideration the incidence of focal liver lesions and the impact upon the medical healthcare system [8-12]. CEUS in liver pathology has helped the characterization of focal liver lesions a lot, increasing the specificity and giving a boost to ultrasound users regarding the FLLs discrimination [13]. To our knowledge there is no published data regarding the learning curve in CEUS but we assume that according to EFSUMB recommendations [14], operators must have a competence level higher than level I according to minimal training requirements in ultrasound (EFSUMB 2006) to perform CEUS [15]. According to our local experience, a minimum of one year is required to gain experience in CEUS, by observing contrast studies done by experts.

Several multicenter studies regarding CEUS have confirmed the utility of this method: the German DEGUUM trial, with more than 1000 lesions, which showed a diagnostic accuracy of CEUS of 90.3% (in differentiating between malignant and benign lesions) [4]; the French STIC trial [5] with a diagnostic accuracy of 86% and the Romanian multicenter study [16] with a diagnostic accuracy of 89.3%. When comparing CEUS vs. CT/MRI for the diagnosis of FLLs [17,18], there were no statistical differences. CEUS for FLLs has already proved its cost

---

**Table II.** Diagnosis accuracy of contrast enhanced ultrasound in focal liver lesions after knowing the clinical data (p value of Wilcoxon test).

<table>
<thead>
<tr>
<th>Examiner/Success</th>
<th>Blinded to clinical data</th>
<th>Unblinded to clinical data</th>
<th>p*</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st Beginner</td>
<td>78%</td>
<td>94%</td>
<td>0.005</td>
</tr>
<tr>
<td>2nd Beginner</td>
<td>82%</td>
<td>96%</td>
<td>0.008</td>
</tr>
<tr>
<td>1st Expert</td>
<td>94%</td>
<td>100%</td>
<td>0.083</td>
</tr>
<tr>
<td>2nd Expert</td>
<td>96%</td>
<td>98%</td>
<td>0.564</td>
</tr>
</tbody>
</table>

**Table III.** The success rate of experts in comparison with beginners and computer assisted diagnosis (*p*, significance level of Wilcoxon test).

<table>
<thead>
<tr>
<th>Examiner</th>
<th>Percentage of successful diagnosis (%)</th>
<th>Beginner (p*)</th>
<th>Expert (p*)</th>
<th>CAD (p*)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beginner</td>
<td>80</td>
<td>–</td>
<td>0.04</td>
<td></td>
</tr>
<tr>
<td>Expert</td>
<td>95</td>
<td>0.04</td>
<td>–</td>
<td>0.01</td>
</tr>
<tr>
<td>CAD</td>
<td>75.2</td>
<td>0.01</td>
<td>–</td>
<td></td>
</tr>
</tbody>
</table>

**Table IV.** Comparison between the performance of experts and beginners, when computer assisted diagnosis was favorable (*p value of Wilcoxon test)

<table>
<thead>
<tr>
<th>Examiner (% correct diagnosis)</th>
<th>p value*</th>
</tr>
</thead>
<tbody>
<tr>
<td>1st Beginner + CAD (94%)</td>
<td>1</td>
</tr>
<tr>
<td>2nd Beginner + CAD (98%)</td>
<td>0.317</td>
</tr>
<tr>
<td>1st Expert (94%)</td>
<td>0.655</td>
</tr>
<tr>
<td>2nd Expert (96%)</td>
<td>0.564</td>
</tr>
</tbody>
</table>
Is Contrast Enhanced Ultrasonography a useful tool in a beginner’s hand?

Tudor Voicu Moga et al

efficiency as a first line imaging method [8,12] compared to CT or MRI, making it well worth considering in daily use. With all these features and advantages that CEUS offers, it is obvious that CAD software for CEUS still has potential that can be exploited.

In order to implement CAD for CEUS, the drawbacks of ultrasound must be overcome. First of all, a good and valid database is required so that the algorithm used for the software can assimilate the standard enhancing patterns of lesions. This implies a correct examination and a good acoustic window. One of the crucial parts of a CAD system is the software development. A basic knowledge in the field of engineering that a physician has, is not enough. A symbiosis between the two fields (medicine and engineering) is required, skills in image processing, data analysis and software development will be the necessary ingredients for reliable software. There are several examples in the literature that have proved the utility of a CAD system in ultrasound. The wide range of use has made it useful in breast lesion interpretation [19], quantifying liver steatosis [20] and even in assisting “killing” the cancer cells [21].

The ideal CAD will be one that can be of a real use for the young examiners. It must bring a diagnostic benefit in characterizing a FLL, especially in those lesions that can be misdiagnosed (atypical lesions, deep ones, small ones etc.). The goal is to create a tool that helps the young examiners in the diagnostic decision. In order to be confident in putting a diagnosis with the help of CEUS, you need a minimum experience in CEUS and ultrasound. There is not a validated and quantifiable ultrasound evaluation that can measure the experience you need in ultrasound/CEUS, this is why a CAD will bring a true benefit for the inexperienced users.

Unlike the study of Sugimoto et al [7], in which the authors developed two CAD schemes using CEUS, one based on physicians’ subjective classification and the other based on quantitative analysis that uses artificial neural networks and decision trees in order to differentiate FLLs, we focused mainly on the malignant/benign fast discrimination of a FLL. In Sugimoto et al study [7] 137 nodules were evaluated: 74 HCCs [23 well-differentiated (w-HCC), 36 moderately differentiated (m-HCC) and 15 poorly differentiated (p-HCC)], 33 liver metastasis and 30 hemangiomas, classifying them into eight patterns according to the enhancement: (1) absent; (2) dotted; (3) peripheral rim like; (4) peripheral nodular; (5) central with spoke wheel-shape; (6) diffuse heterogeneous; (7) diffuse homogeneous; and (8) others. Four artificial neuronal networks (ANN) were used afterwards as decisions models for the evaluators. The performance of the subjective classification for the 137 FLLs were 84.8% for metastasis, 93.3% for hemangioma, 65.2% for w-HCC, 41.7% for m-HCC, and 80.0% for p-HCC. The average classification accuracies for FLLs type three and five were 94.2% and 71.5%, respectively.

The second scheme that Sugimoto used [7] is based on a micro-flow imaging technique (MFI; Toshiba Medical Systems Co., Otawara, Japan) using high mechanical index disruptive flash frames [22]. Through this technique essential details about the lesions’ vessels had been obtained. CEUS with MFI has helped to differentiate the HCCs’ and other FLLs’. Different image features were used as input for the CAD. They used 103 nodules, from which 61 HCCs (24 w-HCC, 28 m-HCC and 9 p-HCC), 26 liver metastases and 16 liver hemangiomas. In this second computerized scheme Sugimoto et al [7] divided the process in three steps: 1. processing the images; 2. extraction of image features and 3. application of ANNs. The classification accuracy of the second CAD for the 103 lesions were 88.5% for metastasis, 93.8% for hemangioma, 79.2% for w-HCC, 50.0% for m-HCC and 77.8% for p-HCC. The classification rate for all HCCs was 86.9%. The average classification accuracies for FLLs type three and five were 88.3% and 75.7%.

Recently a new algorithm was proposed by Gastos [23] in which a quantification algorithm for the detection and evaluation of FLLs with the help of CEUS was tested. The study included 52 CEUS video sequences (30 benign, 22 malignant) in which lesion contour extraction and time intensity curve (TIC) were used that could be computed into the support vector machines (SVMs) classification algorithm in the design of the image analysis mode. The highest classification accuracy from the SVM model was 90.3%, with 93.1% sensitivity and 86.9% specificity.

Perhaps, more important than the correct discrimination of the FLLs is to exclude the malignancy. The malignant/benign status is influencing the therapeutic strategy and prognosis. In our study, after having the final results of the evaluators, we analyzed the weakest point, “Achilles heel”, in evaluating the FLLs. We noticed that the FNH and HCC were the most frequent FLLs that were misdiagnosed by the beginners (Table 1 and 2, Fig. 8 and 9). This might be due to the fact that according to the EFSUMB 2012 guidelines [3], both lesions are hyperenhancing in the arterial phase and FNH, as an additional feature, can be isoenhancing in the late phase like the HCC, making it worthy of expertise (Fig. 5). This problem might be solved by using Sonazoid (perfluorobutane; GE Healthcare, Oslo, Norway), as a contrast agent which allows a better late phase and parenchymal observation [24]. Both beginners increased significantly the diagnos-
tic accuracy for the HCC if the clinical data revealed a cirrhotic liver; thus, we can assert that the presence of cirrhosis influenced the final diagnosis.

Our CAD software had some good results on the malignant/benign fast discrimination but cannot compare with the expert results. Even though we used the same database to develop and test the CAD, if we take a close look at the lesions we notice that some of the misdiagnosed lesions by the beginners were correctly diagnosed by the CAD (Fig. 9); To summarize the algorithm that we propose, if a young investigator has doubts when evaluating a lesion through CEUS imaging, he can use the CAD software to orientate the diagnostic, thus achieving better results. We noticed that the clinical data had a major impact on the beginner’s results (being the major criterion of the results’ improvement), underlying an obvious fact that the imaging investigation should be made by clinicians and not by technicians. The experts’ results were not significantly affected after knowing the clinical data, maybe due to the fact that the CEUS movies might have been previously seen having in mind the retrospective approach.

In order to increase the classification rate, in future developments, we envisage introducing as many features as possible, like tracking algorithms for the ROI, more suitable probabilistic filters and a number of texture features that will increase the method’s robustness. Clinical data must be integrated into the software as it helps the diagnostic accuracy. Our prototype is more an assisting tool than a complex post processing one. However, the lack of research in this field makes this concept a good track to start. CAD for the liver is still at the beginning, a number of issues are to be improved and new, faster algorithms should be developed.

An important bias of our study is the fact the some of the lesions might have been previously seen by the experts during the last years and the beginners did not perform the CEUS investigations, they just reevaluated the video clips from the data base. Another limitation of the study is the lack of integration of clinical data into the CAD algorithm. We can mention the fact that all CEUS video clips were of good quality but even so the success rate of our CAD was less than the results of the beginners. The performance of CAD in the daily routine must be evaluated; however, improvements must be done in order to achieve a reliable software. Through this software we attempted the objective of developing a CAD for FLLs just by reducing the CEUS evaluation to a pattern hypen- enhancement or hypoenhancement decision, which might be a very easy task in “good quality videos”. Achieving this necessitates much more sophisticated software and much more hardware capacity.

Conclusions

The most frequent lesions misdiagnosed at CEUS were FNH and HCC. The CAD prototype is a good comparing tool for a beginner operator that can be developed to assist the diagnosis. In order to increase the classification rate, the CAD system for FLL in CEUS must integrate the clinical data.

Acknowledgement: The research published in this paper was made with support from the grant INOV-IAHEP awarded by the “Victor Babes” University of Medicine and Pharmacy Timisoara, in PROGRAMUL III – C2 – PCFI – 2015/2016.

Conflict of interests: none

References


