Ultrasound Obstet Gynecol 2009; 34: 699–710
Published online 18 November 2009 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/uog.7464

Intravenous contrast ultrasound examination using contrast-tuned imaging (CnTI™) and the contrast medium SonoVue® for discrimination between benign and malignant adnexal masses with solid components

A. C. TESTA*, D. TIMMERMAN†, V. VAN BELLE‡, E. FRUSCELLA*, C. VAN HOLSBEKE†, L. SAVELLI§, E. FERRAZZI¶, F. P. G. LEONE¶, H. MARRET**, F. TRANQUART**, C. EXACOUSTOS††, G. NAZZARO‡‡, D. BOKOR§§, F. MAGRI§§, S. VAN HUFFEL‡, G. FERRANDINA* and L. VALENTIN¶¶

*Department of Obstetrics and Gynecology, *Catholic University of Rome and ††University of 'Tor Vergata', Rome, §University of Bologna, Bologna, ¶DCS L. Sacco, University of Milan, Milan and ‡‡Federico II University of Naples, Naples and §§Bracco S.p.A., Milan, Italy, †Department of Obstetrics and Gynecology, University Hospitals KU and ‡Department of Electrical Engineering (ESAT), SCD, Katholieke Universiteit, Leuven, Belgium, **Department of Obstetrics and Gynecology, Unité INSERM, CHU Bretonneau, Tours, France and ¶¶Department of Obstetrics and Gynecology, Malmö University Hospital, Lund University, Malmö, Sweden*

KEYWORDS: contrast media; ovarian neoplasms; transvaginal ultrasound; ultrasonography

ABSTRACT

Objective *To determine whether intravenous contrast ultrasound examination is superior to gray-scale or power Doppler ultrasound for discrimination between benign and malignant adnexal masses with complex ultrasound morphology.*

Methods *In an international multicenter study, 134 patients with an ovarian mass with solid components or a multilocular cyst with more than 10 cyst locules, underwent a standardized transvaginal ultrasound examination followed by contrast examination using the contrast-tuned imaging technique and intravenous injection of the contrast medium SonoVue®. Time intensity curves were constructed, and peak intensity, area under the intensity curve, time to peak, sharpness and half wash-out time were calculated. The sensitivity and specificity with regard to malignancy were calculated and receiver–operating characteristics (ROC) curves were drawn for gray-scale, power Doppler and contrast variables and for pattern recognition (subjective assignment of a certainly benign, probably benign, uncertain or malignant diagnosis, using gray-scale and power Doppler ultrasound findings). The gold standard was the histological diagnosis of the surgically removed tumors.*

Results *After exclusions (surgical removal of the mass > 3 months after the ultrasound examination, technical problems), 72 adnexal masses with solid components were used in our statistical analyses. The values for peak contrast signal intensity and area under the contrast signal intensity curve in malignant tumors were significantly higher than those in borderline tumors and benign tumors, while those for the benign and borderline tumors were similar. The area under the ROC curve of the best contrast variable with regard to diagnosing borderline or invasive malignancy (0.84) was larger than that of the best gray-scale (0.75) and power Doppler ultrasound variable (0.79) but smaller than that of pattern recognition (0.93).*

Conclusion *Findings on ultrasound contrast examination differed between benign and malignant tumors but there was a substantial overlap in contrast findings between benign and borderline tumors. It appears that ultrasound contrast examination is not superior to conventional ultrasound techniques, which also have difficulty in distinguishing between benign and borderline tumors, but can easily differentiate invasive malignancies from other tumors. Copyright © 2009 ISUOG. Published by John Wiley & Sons, Ltd.*

INTRODUCTION

Malignant ovarian tumors are diagnosed at an advanced stage in 75% of cases, and they are associated with the

Correspondence to: Dr A. C. Testa, Università Cattolica del Sacro Cuore, L.go A. Gemelli 8, 00168 Rome, Italy (e-mail: atesta@rm.unicatt.it)

Accepted: 30 March 2009

Copyright © 2009 ISUOG. Published by John Wiley & Sons, Ltd.

highest mortality figures of all gynecological cancers[1]. It is sometimes difficult to determine preoperatively if an ovarian tumor is benign or malignant. However, this knowledge is essential for appropriate management[2].

Ultrasonography is a well-established imaging modality for the preoperative evaluation of pelvic masses. Subjective evaluation (pattern recognition) of the gray-scale ultrasound image by an expert is accurate with regard to malignancy in more than 90% of cases[3]. Color Doppler and power Doppler ultrasound can be used to detect neovascularization in malignant lesions[4], and Doppler examination may add information to the gray-scale ultrasound image. However, the contribution of Doppler ultrasonography seems to be limited in an ordinary population of adnexal masses[3,5,6].

Even though the sensitivity of color and power Doppler ultrasound has improved thanks to technical developments in recent years, vessels smaller than 100 microns in diameter cannot be detected by Doppler ultrasound. However, ultrasonography enhanced with intravascular contrast agents allows detection of signals from blood vessels with diameters of less than 40 microns[7]. Dedicated ultrasound technology has been developed to optimize the use of ultrasound contrast media in gynecology, e.g., contrast-tuned imaging (CnTI™) technology using the second-generation contrast agent SonoVue®[8]. In contrast to earlier-generation contrast agents, second-generation contrast agents provide a substantial harmonic response when insonated by ultrasound at low acoustic pressure[9]. When second-generation contrast bubbles are insonated with ultrasound with low acoustic pressure, they are not destroyed but remain in the blood circulation for several minutes.

The aim of this study was to determine whether intravenous contrast ultrasound characteristics obtained using CnTI–SonoVue are superior to gray-scale or power Doppler ultrasound characteristics, or to subjective pattern recognition using these imaging modalities, for discrimination between benign and malignant adnexal masses with complex ultrasound morphology, and whether contrast adds any information to conventional ultrasound.

METHODS

Patients were recruited from December 2004 to June 2005. Eight ultrasound centers participated: the Catholic University of Sacred Heart, Rome and Campobasso, Italy; Policlinico S. Orsola-Malpighi, University of Bologna, Italy; University Hospitals, Katholieke Universiteit Leuven, Belgium; DCS L. Sacco, University of Milan, Italy; Malmö University Hospital, Malmö, Sweden; Federico II University Hospital, Napoli, Italy; CHU Bretonneau, Tours, France; and University of 'Tor Vergata', Rome, Italy. The inclusion criteria were: ultrasound diagnosis of unilocular-solid, multilocular-solid or solid adnexal mass or multilocular adnexal cyst with more than 10 cyst locules, the whole tumor being accessible by transvaginal sonography; age 18 years or more; and written informed consent to participate in the study. Patients having an adnexal mass with ultrasound features compatible with dermoid cyst, hydrosalpinx or peritoneal cyst, pregnant or nursing patients, patients who had undergone previous chemotherapy, and patients with any contraindication to the use of SonoVue contrast medium were not eligible. Examples of contraindications to the use of SonoVue are cardiac insufficiency, severe lung disease, severe cardiac arrhythmia, recent myocardial infarction, unstable angina pectoris, acute endocarditis, artificial heart valves, acute systemic inflammation or sepsis, hypercoaguability, recent thrombo-embolic disease, and terminal renal or liver disease. Exclusion criteria were histological diagnosis obtained more than 3 months after the ultrasound examination and technical problems making the contrast ultrasound examination impossible to evaluate. The study protocol was approved by the local ethics committees of the participating centers.

Each patient was examined as described below. Before the ultrasound examination, a history was taken following a strict research protocol. This included number of first-degree relatives with ovarian cancer or breast cancer and use of hormone replacement therapy or contraceptive pills. A woman was considered to be postmenopausal if she reported a period of at least 12 months of amenorrhea after the age of 40 years, provided that medical therapy, pregnancy or disease did not explain the amenorrhea. Women 50 years or older who had undergone hysterectomy so that the time of menopause could not be determined, were also defined as postmenopausal.

All ultrasound examinations were performed using a high resolution (5.0–9.0 MHz) endovaginal probe connected to a Technos MPX ultrasound system (ESAOTE S.p.A.; Genova, Italy). A transvaginal gray-scale and power Doppler ultrasound examination was performed using a standardized examination technique, standardized definitions of ultrasound terms[9] and standardized power Doppler ultrasound settings (frequency 5 MHz, pulse repetition frequency 750 Hz, color gain just below the background noise level). The following parameters were assessed: location of the lesion, size of the lesion (three orthogonal diameters), unilateral or bilateral mass, presence of ascites and/or fluid in the pouch of Douglas, type of mass (unilocular-solid, multilocular, multilocular-solid, solid), presence of papillary projections (defined as any solid protrusion into a cyst cavity with a height of $\geq$3 mm[10]), number of papillary projections, irregularity of the surface of papillary projections, presence of solid tissue other than papillary projections and presence of septa. The color content of the papillary projections and of the solid tissue other than papillary projections at power Doppler examination was estimated subjectively by the ultrasound examiner using a color score as described by Timmerman *et al.* (1 = no vascularization; 2 = minimal vascularization; 3 = moderate vascularization, 4 = high vascularization)[10]. On the basis of subjective evaluation of the gray-scale and power Doppler ultrasound findings (pattern recognition) the examiner estimated the risk of malignancy (certainly benign, probably benign,

Copyright © 2009 ISUOG. Published by John Wiley & Sons, Ltd.

uncertain or malignant). Even when the examiner was uncertain about whether a tumor was benign or malignant, he or she was obliged to classify the mass as benign or malignant. Borderline tumors were classified as malignant.

Contrast examination was carried out after completion of the gray-scale and power Doppler ultrasound examination. The contrast-enhanced examination was performed using the CnTI technology applied to the transvaginal probe and the ultrasound contrast agent SonoVue (Bracco Imaging S.p.A, Milan, Italy). The same settings were used for all contrast examinations (pulse repetition frequency of 750 Hz; frequency = penetration; derated pressure 126 kPa; focus position immediately beneath the lesion; general gain 150). Each patient received two injections of 2.4 mL of SonoVue in bolus via an indwelling catheter (20 G or more) placed in an antecubital vein. The first injection of SonoVue was used to analyze time–intensity curves during the passage of contrast through the mass. This examination was performed on the section through the pelvic mass that was judged subjectively by the ultrasound examiner to contain the most vascularized solid part(s) of the tumor at power Doppler examination. A 3-min CnTI recording for time–intensity analysis was started immediately after completion of the injection of the first bolus dose of SonoVue. The examiner was not allowed to move the probe during the 3-min recording, because the quantitative contrast analysis had to be done on one and the same section through the tumor. If the image showed that the examiner had moved the transducer during the examination, the examination was excluded. After a 20-min interval (to allow disappearance of contrast from the tumor under study), another 2.4 mL intravenous bolus dose of SonoVue was injected. A 3-min scan through the whole tumor was started immediately after the injection and was stored electronically as a clip. This clip was used for qualitative analysis (i.e. determination of the presence or absence of contrast signals and assigning a contrast score). During the examination the intensity of the contrast signals in the papillary projections and in the solid tissue other than papillary projections was estimated subjectively by the ultrasound examiner using a contrast score: no vascularization, minimal vascularization, moderate vascularization, or high vascularization. In cases of multiple masses only the most complex mass – or, if all masses had similar ultrasound morphology, the largest one or the one most easily accessible by transvaginal ultrasound – was examined with CnTI–SonoVue and included in our statistical analysis. All contrast clips were stored electronically for later analysis using dedicated software. The patients were monitored for adverse events for 2 h after the last injection of contrast.

The time–intensity curves were analyzed by a single operator (F. M.), who had no knowledge of the histological diagnosis when performing the analysis. The software used was Qontraxt™ (Bracco Imaging S.p.A). This software processes the signal intensity changes over time induced in an organ or lesion by the intravenous injection of SonoVue. On CnTI examination the highest value of signal intensity (100%) corresponds to white in the gray-scale bar of the ultrasound equipment and the lowest value (0%) corresponds to black in the gray-scale bar. As a result of the processing of contrast signals the software provides pixel-by-pixel color-coded maps of the perfusion parameter for the lesion under investigation. The perfusion parameters that can be mapped in color are: TTP (time-to-peak, i.e. time from the arrival of contrast agent to its maximum signal intensity value), sharpness (wash-in rate), peak (maximum signal intensity during the transit of contrast), area under the signal intensity curve (AUC) (i.e. the integral of the signal intensity curve) and half wash-out (time from peak of signal intensity to half its value). High values of the perfusion parameter are coded in red (the darker the red the higher the intensity), low values are coded in blue (the darker the blue the lower the intensity), and intermediate intensities are coded in yellow (the darker the yellow the higher the intensity). Detailed information about the Qontraxt software has been published[11,12].

Using the color-coded maps, which spatially correspond to the gray-scale ultrasound image, the perfusion parameters (TTP, sharpness, peak, AUC, half wash-out) in papillary projections and in other solid components of the mass were calculated. The operator defined the region of interest (ROI) by drawing the contours of the solid component(s) excluding the cyst wall and any septa, so that solid areas and papillary projections were included in the ROI analysis. Two analyses of the color-coded images were performed: (1) the average of the time–intensity curves from all pixels included in the ROI, i.e. ROI analysis, and (2) the time–intensity curve from the pixel corresponding to the maximum peak value in the ROI, i.e. MAX analysis. The pixel with the maximum peak value in the ROI was identified by the operator by moving an indicator over the darkest red area to see the intensity values provided by the Qontraxt software. The method of analyzing the time–intensity curves is illustrated in Figures 1–3. If no perfusion of the solid components was detected after contrast injection, quantitative contrast analysis was not performed, but the pelvic mass was included for qualitative contrast analysis.

To determine intraobserver reproducibility of the time–intensity curve analyses, F. M. analyzed twice (3 years apart) the recordings of 22 tumors. These were selected by A. C. T. from the statistical data sheet without knowledge of any ultrasound or contrast results so as to include eight malignant tumors, two borderline tumors and 12 benign tumors.

The results of both unenhanced and contrast-enhanced examinations and those of pattern recognition were compared with those of the histological examination of the respective surgical specimens. Staging of the malignant tumors was done by the attending physician, in accordance with the classification system recommended by the International Federation of Gynecology and Obstetrics[13].

Copyright © 2009 ISUOG. Published by John Wiley & Sons, Ltd.

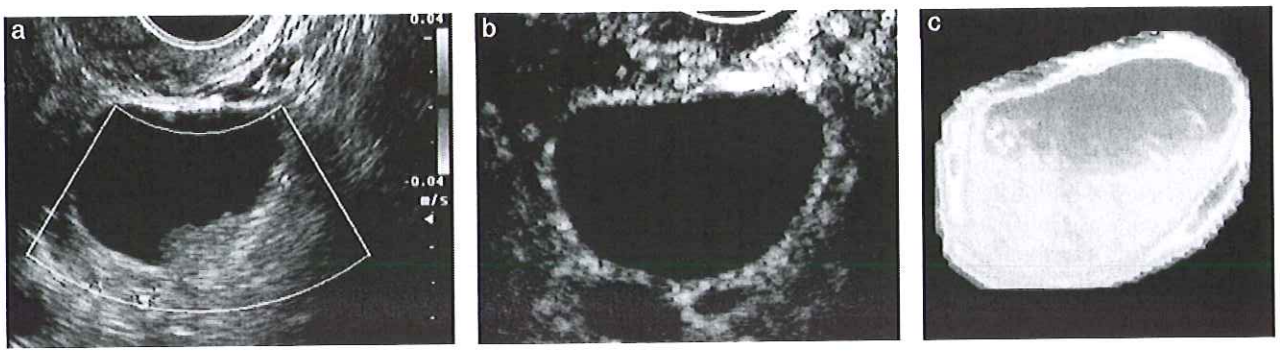

Figure 1 Color Doppler image (a), contrast-enhanced ultrasound image (b) and color-coded image (c) of an ovarian cyst. The peak value was coded in color using Qontraxt™ software. High values of the perfusion parameter are coded in red (the darker the red the higher the intensity) and low values are coded in blue (the darker the blue the lower the intensity). Note that the echogenic intracystic tissue in (a) manifests no signals at color Doppler examination.

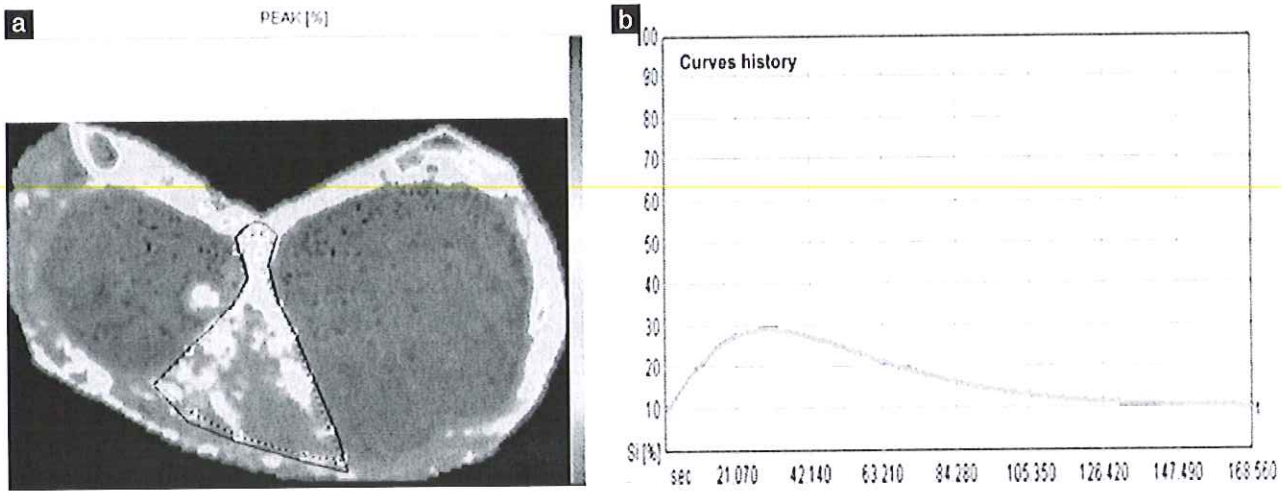

Figure 2 Illustration of calculation of time–intensity curves in the region of interest (ROI), i.e. ROI analysis. The ROI has been outlined on the color-coded image (created using Qontraxt™ software). (a) The variable that has been coded in color is the peak value. The ROI is selected by drawing the contours of the solid component(s) excluding the cyst wall and any septa, so that all solid areas and papillary projections are included in the ROI analysis. (b) The corresponding time–intensity curve, which is the mean of the time–intensity curves of all pixels included in the ROI.

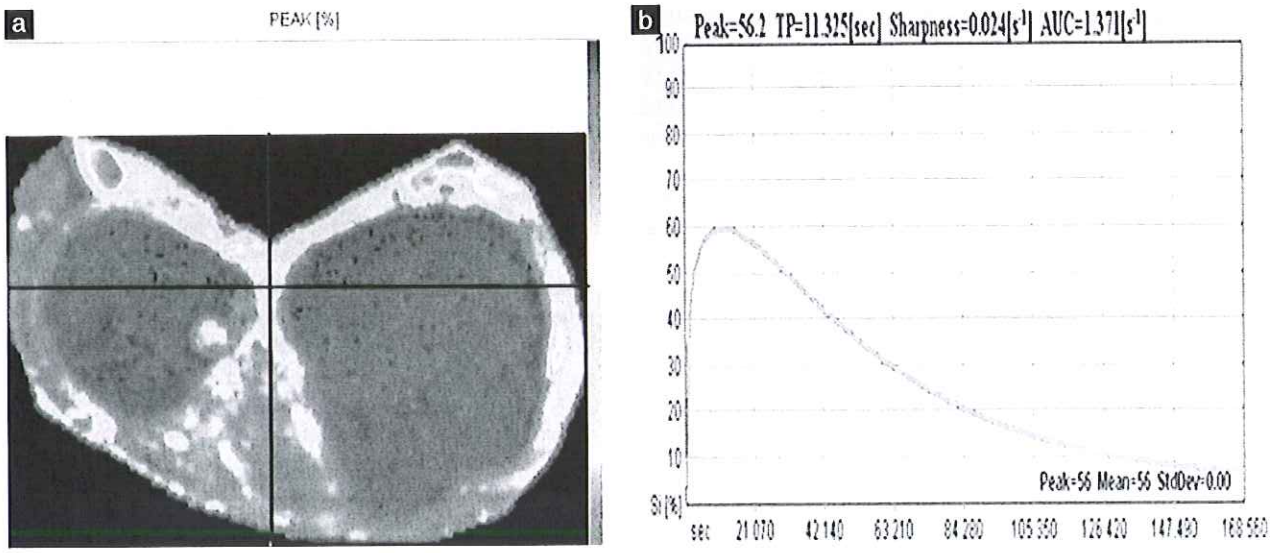

Figure 3 Illustration of calculation of the time–intensity curve from the pixel with the highest intensity, i.e. MAX analysis. (a) The cross has been placed in the pixel with the highest intensity in the color-coded image (created using Qontraxt™ software), the variable that has been coded in color being the peak value. (b) The corresponding time–intensity curve from the selected pixel.

Copyright © 2009 ISUOG. Published by John Wiley & Sons, Ltd.

Statistical analysis

For statistical analysis, primary invasive, borderline and metastatic invasive tumors were all classified as malignant. Statistical analysis was carried out using SAS 9.1.3 Service Pack 3 (SAS Institute Inc., Cary, NC, USA). The statistical significance of differences in categorical data with regard to the tumor class (benign, borderline or malignant) was determined using the χ^2-test or Fisher's exact test, or for ordinal data the Mantel–Haenszel χ^2-test[14] was used. For continuous data the Mann–Whitney *U*-test or the Kruskal–Wallis test was used as appropriate. Multivariate logistic regression analysis, including variables yielding $P < 0.05$ on the likelihood ratio test in univariate analysis, was carried out to build mathematical models to calculate the probability of malignancy for each patient, using forward selection to build the model. Selection of variables was based on *P*-values of the likelihood ratio test. We aimed at building two models, one including contrast variables and one not including contrast variables.

Receiver–operating characteristics (ROC) curves were calculated for single predicting variables as well as for the best multivariate logistic regression model to evaluate their diagnostic ability. The area under the ROC curve and the 95% CI of this area were calculated. If the lower limits of the CI for the area under the ROC curve was > 0.5, the diagnostic test was considered to have discriminatory potential. The ROC curves were also used to determine the mathematically best cut-off value to predict malignancy for each diagnostic test (single variables as well as logistic regression model), the mathematically best cut-off value being defined as that corresponding to the point on the ROC curve situated farthest from the reference line. The accuracy, sensitivity, specificity, and positive and negative likelihood ratios (LR+ and LR−) of the mathematically best cut-off value were also calculated. The statistical significance of a difference in sensitivity, specificity and accuracy between two tests when using the mathematically best cut-off to predict malignancy was determined using the McNemar test. We defined the best diagnostic test as the one with the largest area under the ROC curve, and $P < 0.05$ was considered statistically significant.

Intraobserver repeatability was expressed as the difference between the two measurement results obtained by the same observer. Limits of agreement (mean difference ± 2 SD) were calculated as described by Bland and Altman[15]. Systematic bias between the first and second analysis was determined by calculating the 95% CI for the mean difference (mean difference ± 2 SE). If zero lay within this interval, no bias was assumed to exist. Intraobserver repeatability was also expressed as the intraclass correlation coefficient (ICC), variance components being estimated from one-way analysis of variance[16].

Power calculation

It was not possible to make a proper power calculation, because at the time of planning the study there was no information available on results of quantitative analysis of the examination of adnexal masses using the CnTI–SonoVue technique. The sample size was estimated using preliminary results from the International Ovarian Tumor Analysis (IOTA) study[17], i.e. results that were unpublished at the time of planning the contrast study. These results showed that among the types of tumor that we wanted to include in our study (i.e. tumors with solid components and multilocular cysts with more than 10 locules), 43% were malignant, with 28% being primary invasive ovarian cancers, 6% metastatic invasive cancers, and 9% borderline tumors. If we were to include 160 tumors in our study, and if the results of the IOTA study applied to our contrast study, there would be 90 benign tumors, 45 primary invasive ovarian cancers, 14 borderline tumors, and 11 metastatic invasive tumors in our study. We thought that this would probably be a sufficient number to determine whether the CnTI–SonoVue technique would offer diagnostic information in adnexal tumors.

RESULTS

During the time limits set by the sponsors of the study, we recruited and examined 134 patients. No adverse effects of SonoVue occurred. After exclusions, 89 patients

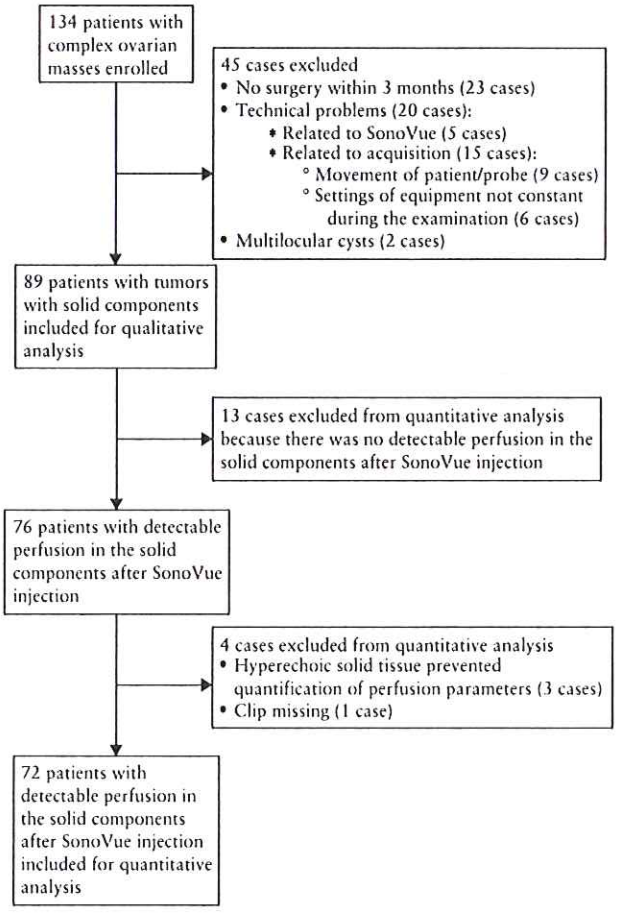

Figure 4 Flow chart illustrating the inclusion of patients in the study.

Copyright © 2009 ISUOG. Published by John Wiley & Sons, Ltd.

Table 1 Background data and histological diagnoses of included and excluded patients

Parameter	Patients included for qualitative analysis (n = 89)	Patients included for quantitative analysis (n = 72)	Patients excluded (n = 45)
Clinical data			
Age (years, mean (range))	50 (22–85)	52 (22–85)	52 (17–85)
Nulliparous	27 (30)	19 (26)	13 (29)
Postmenopausal	47 (53)	44 (61)	27 (60)
Hormonal therapy*	12 (13)	8 (11)	5 (11)
At least one first-degree relative with ovarian cancer	3 (3)	3 (4)	1 (2)
At least one first-degree relative with breast cancer	12 (13)	10 (14)	7 (16)
Personal history of ovarian cancer	2 (2)	2 (3)	2 (4)
Personal history of breast cancer	6 (7)	6 (8)	0
Results of pattern recognition			
Benign	8 (9)	6 (8)	12 (27)
Probably benign	23 (26)	14 (19)	20 (44)
Uncertain	29 (33)	25 (35)	6 (13)
Malignant	29 (33)	27 (38)	6 (13)
Not assessable	0	0	1 (2)†
Histology			
Benign tumors	52 (58)	37 (51)	21 (47)
Cystadenoma/cystadenofibroma	22 (25)	18 (25)	8 (18)
Endometrioma	11 (12)	6 (8)	0
Simple cyst	4 (4)	4 (6)	1 (2)
Dermoid	6 (7)	2 (3)	0
Ovarian fibroma	7 (8)	7 (10)	0
Functional cyst	2 (2)	0	6 (13)
Specific diagnosis not reported	0	0	6 (13)
Borderline tumors	10 (11)	9 (13)	2 (4)
Serous type, Stage I	7 (8)	6 (8)	2 (4)
Serous type, Stage II–III	3 (3)	3 (4)	0
Malignant tumors	27 (30)	26 (36)	4 (9)
Epithelial invasive ovarian carcinoma, Stage I	5 (6)	5 (7)	1 (2)
Epithelial invasive ovarian carcinoma, Stage III–IV	12 (13)	12 (17)	1 (2)
Non-epithelial ovarian malignant tumor	3 (3)	3 (4)	1 (2)
Tubal carcinoma	3 (3)	3 (4)	0
Metastatic ovarian carcinoma	4 (4)‡	3 (4)§	1 (2)
No surgery	0	0	18 (40)

Values are given as *n* (%) except where indicated. *Hormonal replacement therapy or contraceptive pill. †In one excluded patient the result of pattern recognition was not reported. ‡Two Krukenberg tumors, one ovarian metastasis from Burkitt's lymphoma, one ovarian metastasis from endometrial cancer. §Two Krukenberg tumors, one ovarian metastasis from Burkitt's lymphoma.

remained for qualitative analysis, and 72 remained for quantitative analysis (Figure 4). Because there were only two multilocular cysts these were excluded to obtain a homogeneous study group. The demographic background data, the results of pattern recognition and the histopathology for the patients included and excluded are shown in Table 1. The difference in results of pattern recognition between patients included ($n = 89$) and excluded ($n = 45$) for qualitative analysis was statistically significant, with fewer tumors being suspected to be malignant in the excluded group ($P = 0.0004$). The same was true of the difference in results of pattern recognition between patients included for quantitative analysis, i.e. time–intensity curve analysis ($n = 72$) and those excluded ($n = 45$) ($P < 0.0001$). There were no other statistically significant differences between the tumors included and excluded. Of the 89 tumors used for qualitative analysis, 27 (30%) were invasively malignant at pathological examination and 10 (11%) were borderline malignant. Of the 72 cases used for time–intensity curve analysis, 26 (36%) were invasively malignant and nine (13%) were borderline malignant.

The most important results of gray-scale, power Doppler analysis and qualitative contrast examination are shown in Table 2 (for detailed results see Table S1). Most gray-scale ultrasound features differed significantly between benign, borderline and malignant tumors. Papillary projections were more frequently detected in benign (69%) and borderline tumors (80%) than in invasively malignant tumors (26%), while a purely solid mass was more often observed in invasively malignant tumors (44%) than in benign (17%) and borderline tumors (10%). Power Doppler results differed too, the detection rate of Doppler signals in papillary projections being lower in benign (31%) than in borderline (75%) and malignant tumors (100%) and the color scores in papillary projections and in solid tissue other than papillary projections were higher in malignant tumors than in benign tumors and borderline tumors (see Table S1). The detection rate of power Doppler signals in any solid component (papillary projection or other) was also lower in benign (46%) than in borderline (80%) and malignant tumors (100%), and the color score for any solid component was higher in malignant than in benign

Copyright © 2009 ISUOG. Published by John Wiley & Sons, Ltd.

Table 2 Ultrasound characteristics of benign, borderline and malignant tumors in included patients: gray-scale ultrasound, Doppler ultrasound and qualitative contrast parameters

Parameter	*Benign (n = 52)*	*Borderline (n = 10)*	*Malignant (n = 27)*	P*
Gray-scale characteristics				
Bilateral tumor	5 (10)	3 (30)	8 (30)	0.0473
Ascites	1 (2)	1 (10)	11 (41)	< 0.0001
Fluid in the pouch of Douglas	9 (17)	3 (30)	12 (44)	0.0351
Maximum diameter of lesion (mm, median (range))	52 (18–148)	66 (22–105)	86 (41–145)	0.0007
Type of tumor				
Unilocular-solid	27 (52)	5 (50)	5 (19)	0.0178
Multilocular-solid	16 (31)	4 (40)	10 (37)	
Solid	9 (17)	1 (10)	12 (44)	
Solid component				
Papillary projections only	34 (65)	6 (60)	3 (11)	< 0.0001
Solid tissue but no papillary projections	7 (13)	1 (10)	8 (30)	
Papillary projections and other solid tissue	2 (4)	2 (20)	4 (15)	
Purely solid mass	9 (17)	1 (10)	12 (44)	
Presence of septa	16 (31)	4 (40)	10 (37)	0.7741
Power Doppler characteristics				
Power Doppler signals detected in papillation, if papillation present	11/36 (31)	6/8 (75)	7/7 (100)	0.0003
Power Doppler signals detected in septa, if septa present	10/16 (63)	4/4 (100)	10/10 (100)	0.0425
Power Doppler signals detected in any solid component	24/52 (46)	8/10 (80)	27/27 (100)	< 0.0001
Color score in any solid component				
Absent	28/52 (54)	2/10 (20)	0/27 (0)	< 0.0001
Minimal	15/52 (29)	6/10 (60)	3/27 (11)	
Moderate	8/52 (15)	2/10 (20)	13/27 (48)	
High	1/52 (2)	0/10 (0)	11/27 (41)	
Qualitative contrast characteristics				
Contrast signals detected in papillation, if papillation present	26/36 (72)	8/8 (100)	7/7 (100)	0.0879
Contrast signals detected in septa, if septa present	15/16 (94)	4/4 (100)	10/10 (100)	1.0
Contrast signals detected in any solid component	39/52 (75)	10/10 (100)	27/27 (100)	0.0044
Contrast score in any solid component				
Absent	13/52 (25)	0/10 (0)	0/27 (0)	< 0.0001
Minimal	16/52 (31)	1/10 (10)	0/27 (0)	
Moderate	15/52 (29)	6/10 (60)	4/27 (15)	
High	8/52 (15)	3/10 (30)	23/27 (85)	

Values are given as *n* (%) except where indicated. *Kruskal–Wallis test for continuous variables and χ^2-test, Fisher's exact test or Mantel–Haenszel χ^2-test as appropriate for other variables.

and borderline tumors (Table S1). Contrast scores also differed between the three groups of tumor, the detection rate of contrast signals in any solid component being higher in invasively malignant (100%) and borderline tumors (100%) than in benign tumors (58%), and the finding of a high contrast score in any solid component was more common in invasively malignant tumors (high contrast score in 85%) than in borderline tumors (high contrast score in 30%) and in benign tumors (high contrast score in 15%).

Among the 43 tumors containing papillary projections but no other solid components flow was detected in the papillary projections at power Doppler examination in eight (89%) of the nine malignancies and in 10 (29%) of the 34 benign tumors. On qualitative contrast examination, perfusion was detected in the papillary projections in all nine malignancies and in an additional 15 of the 34 benign tumors. Thus, the ability to discriminate between benign and malignant tumors was poorer for the presence of contrast signals than for the presence of power Doppler signals in papillary projections. Among the remaining 46 tumors with solid components, flow was detected in the solid components on power Doppler examination in 27 (96%) of the 28 malignancies and in 14 (78%) of the 18 benign tumors. On qualitative contrast examination, perfusion was detected in the solid components in all 28 malignancies and in the same 14 benign tumors that manifested power Doppler signals. Thus, the ability to discriminate between benign and malignant tumors was similar in the presence of contrast signals and in the presence of power Doppler signals in solid components.

The gray-scale and power Doppler ultrasound results in the benign, borderline and malignant masses used for time–intensity curve analysis ($n = 72$) were essentially the same as those described in Table 2. However, the solid components other than papillary projections in the benign tumors used for time–intensity curve analysis were larger (median largest diameter 51 (range, 12–95) mm than those in the benign tumors used for qualitative contrast examination (32 (range, 9–95) mm). The results of the time–intensity curve analysis are shown in Table 3, and the overlap in results between benign, borderline and malignant tumors is illustrated in Figure 5. Both in the ROI and MAX analyses, the peak values and AUC values

Copyright © 2009 ISUOG. Published by John Wiley & Sons, Ltd.

were highest and the half wash-out time was longest in the malignant tumors. The values for the malignant tumors were statistically significantly higher than those for the benign tumors (P varying from < 0.0001 to 0.0093) and also statistically significantly higher than those for the borderline tumors (P varying from 0.0008 to 0.0114). However, the difference in half wash-out time between the malignant tumors and the borderline tumors was not

Table 3 Results of quantitative contrast analysis using Qontraxt™ software

	Benign		*Borderline*		*Malignant*		
Parameter	n	*Value*	n	*Value*	n	*Value*	P*
ROI analysis							
Peak	37	8.81 (1.06–25.11)	9	13.18 (4.16–16.44)	26	19.44 (9.19–30.31)	< 0.0001
Sharpness	37	0.25 (0.02–0.78)	9	0.23 (0.13–0.30)	26	0.20 (0.03–0.55)	0.3402
TTP (s)	37	13.17 (6.49–90.42)	9	17.07 (7.46–26.67)	26	13.06 (8.21–40.75)	0.4015
AUC†	36	647.94 (123.90–3331.17)	7	807.06 (349.15–2608.48)	25	1769.05 (644.69–3418.09)	< 0.0001
Half wash-out (s)‡	35	38.00 (18.50–84.00)	7	35.00 (22.00–61.00)	24	51.00 (29.00–112.00)	0.0287
MAX analysis							
Peak	37	20.33 (6.27–61.32)	9	20.30 (9.71–30.93)	26	37.90 (18.82–53.22)	< .0001
Sharpness	37	0.20 (0.03–0.85)	9	0.22 (0.05–0.96)	26	0.13 (0.02–0.63)	0.1691
TTP (s)	37	13.84 (6.01–58.72)	9	20.22 (9.65–35.65)	26	12.27 (8.63–38.13)	0.2542
AUC†	36	1083.67 (189.40–3497.65)	7	1181.00 (757.95–3460.89)	25	2779.61 (927.60–8679.14)	< 0.0001
Half wash-out (s)§	34	31.50 (10.00–76.88)	6	32.00 (17.00–44.00)	24	53.50 (25.00–117.00)	< 0.0001

All data are given as median (range). *Kruskal-wallis test. †Data are missing in four cases because the duration of the clip was less than 3 min. ‡Data are missing in six cases because in four cases the duration of the clip was less than 3 min and in two cases the signal intensity was not reduced to half of its maximum value during the 3-min recording. §Data are missing in eight cases because in four cases the duration of the clip was less than 3 min and in four cases the signal intensity was not reduced to half of its maximum value during the 3-min recording. AUC, area under the curve; half wash-out, time from the peak signal intensity to half its value; MAX analysis, analysis using the time–intensity curve in the pixel corresponding to the maximum peak value (MAX) in the region of interest (ROI), defined by drawing the contours of the solid component(s) excluding the cyst wall and any septa; ROI analysis, analysis using the average of the time–intensity curves of all pixels included in the ROI; TTP, 'time-to-peak', i.e., the time from the arrival of contrast agent to the tumor to its maximum signal intensity value.

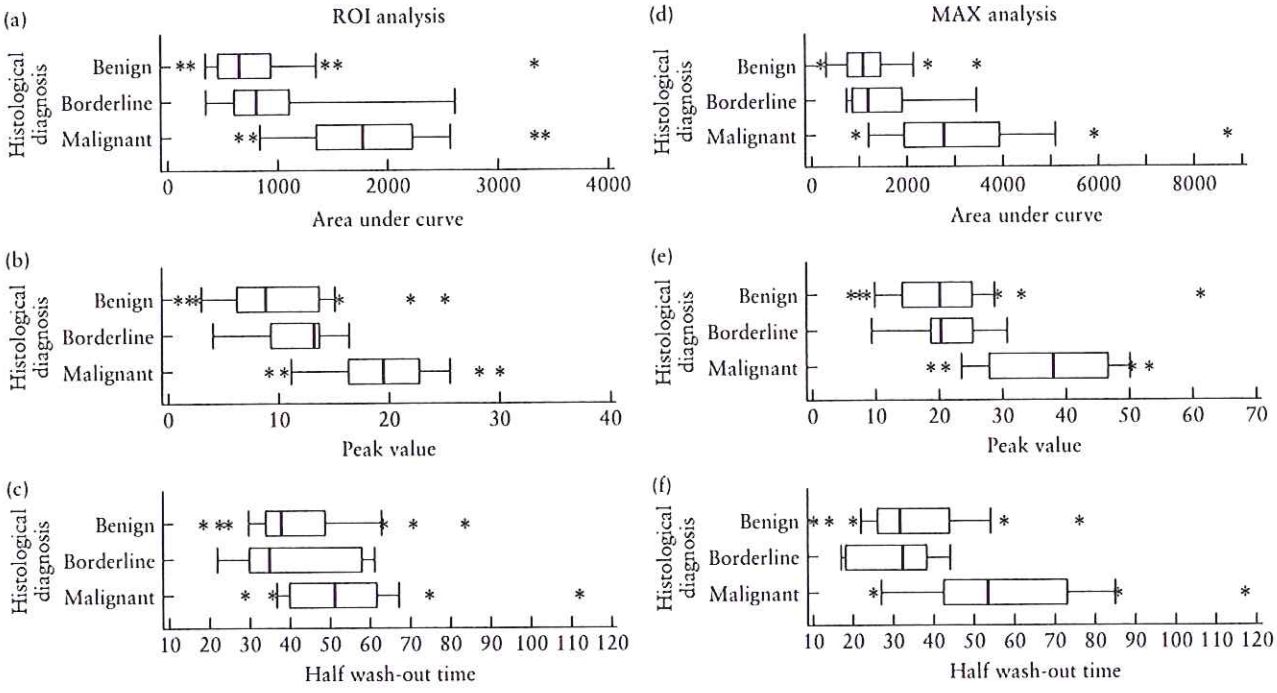

Figure 5 Results of time–intensity curve analysis in benign, borderline malignant and invasively malignant tumors. Results of region of interest (ROI) analysis are shown in (a), (b) and (c): area under the curve (a), peak value (b) and half wash-out time (c). Results of analysis in the pixel manifesting the highest intensity (MAX analysis) are shown in (d), (e) and (f): area under the curve (d), peak value (e) and half wash-out time (f). Box-and-whisker plots show the median (line), range of the middle 50% of the values (box) and the 10th and 90th percentiles (whiskers). *Data points that lie outside the 10th and 90th percentiles. Note the overlap in results, in particular between benign and borderline tumors.

Copyright © 2009 ISUOG. Published by John Wiley & Sons, Ltd.

Table 4 Diagnostic performance of the best gray-scale, power Doppler and contrast variables and of pattern recognition with regard to discriminating between benign and borderline or invasively malignant tumors

Parameter	*Area under ROC curve, estimate (95% CI)*	*Optimal cut-off**	*Sens. (%)*	*Spec. (%)*	*Acc. (%)*	*LR+*	*LR−*	*TP*	*TN*	*FN*	*FP*	*Total*
Gray-scale characteristics												
Largest solid component	0.75 (0.63–0.87)	16 mm	82.9	63.9	73.2	2.29	0.27	29	23	6	13	71†
Power Doppler results												
Highest color score in any solid component	0.79 (0.69–0.89)	Moderate	68.6	75.7	72.2	2.82	0.42	24	28	11	9	72
Qualitative contrast results												
Highest contrast score in any solid component	0.78 (0.69–0.88)	High	68.6	78.4	73.6	3.17	0.40	24	29	11	8	72
Time–intensity curve analysis												
ROI analysis												
Peak	0.83 (0.73–0.93)	10.55	88.6	67.6	77.8	2.73	0.17	31	25	4	12	72
AUC	0.84 (0.74–0.94)	1112.56	75.0	80.6	77.9	3.86	0.31	24	29	8	7	68‡
MAX analysis												
Peak	0.79 (0.68–0.90)	29.03	60.0	91.9	76.4	7.40	0.44	21	34	14	3	72
AUC	0.82 (0.72–0.93)	1760.93	75.0	83.3	79.4	4.50	0.30	24	30	8	6	68‡
Pattern recognition (four confidence levels)§	0.93 (0.88–0.98)	Malignant	74.3	97.3	86.1	27.50	0.26	26	36	9	1	72
LR without contrast¶	0.86 (0.77–0.94)	0.20	88.6	75.0	77.1	3.17	0.27	27	27	7	9	70
LR with contrast**	0.89 (0.81–0.96)	0.34	85.7	75.7	80.5	3.52	0.19	30	28	5	9	72

*Values equal to or above the cut-off indicate malignancy. †In one case data were missing. ‡Data were missing in four cases, because the duration of the clip was less than 3 min. §The false-positive case was a serous cystadenoma; the nine false-negative cases comprised four serous borderline tumors, two tubal carcinomas, one epithelial carcinoma, and two non-epithelial carcinomas. ¶LR without contrast (multivariate logistic regression model not including a contrast variable): Probability of malignancy (P) = $e^z/(1+e^z$ where $z = -4.6359 + (0.0331 \times$ largest diameter of the largest solid component in mm) $+ (1.5922 \times$ highest color score in any solid component) $+ (0.7456 \times$ number of papillations). **LR with contrast (multivariate logistic regression model including a contrast variable): $P = e^z/(1+e^z)$ where $z = -5.5367 + (0.2526 \times$ peak ROI) $+ (0.0325 \times$ largest diameter of the lesion in mm). AUC, area under the curve; FN, false negative; FP, false positive; LR−, negative likelihood ratio; LR+, positive likelihood ratio; MAX analysis, analysis using the time–intensity curve in the pixel corresponding to the maximum peak value (MAX) in the region of interest (ROI); ROC, receiver–operating characteristics curve; ROI analysis, analysis using the average of the time–intensity curves of all pixels included in the ROI; Acc., accuracy; Sens., sensitivity; Spec., specificity; TN, true negative; TP, true positive.

statistically significant in the ROI analysis ($P = 0.1428$). The values for the benign tumors and borderline tumors were very similar and not statistically significantly different. Sharpness and TTP values did not differ significantly between the three groups of tumor, but the TTP values were highest in the borderline tumors and low in the benign and invasively malignant tumors.

The diagnostic performance of the best gray-scale, power Doppler and contrast variables, as well as that of pattern recognition, are shown in Table 4. The diagnostic performance of the highest color score and highest contrast score in any solid component was very similar (area under ROC curve 0.79 vs. 0.78). The areas under the ROC curves of the best quantitative contrast variables were larger than those of the best gray-scale and power Doppler ultrasound variables but smaller than that of pattern recognition (Figure 6). The multivariate logistic regression model that best predicted malignancy included two variables – the maximum diameter of the lesion and the peak value in the ROI (probability of malignancy = $e^z/(1+e^z)$ where $z = -5.5367 + (0.2526 \times$ peak ROI) $+ (0.0325 \times$ largest diameter of the lesion in mm)). This model had an area under the ROC curve of 0.89. The best risk calculation model not including any contrast variable had an area under the ROC curve of 0.86 ($z = -4.6359 + (0.0331 \times$ largest diameter of the largest solid component in mm) $+ (1.5922 \times$ highest color score in any solid component) $+ (0.7456 \times$ number of papillations)). Only pattern recognition had an area under the ROC curve larger than that of the risk estimation models. With one exception (accuracy was significantly higher for pattern recognition than for the highest color score in any solid component; $P = 0.03$) there were no statistically significant differences in accuracy between any of the tests in Table 4.

Intraobserver reproducibility for the two contrast variables with the highest predictive performance (peak and AUC in the ROI and MAX analysis) is shown in Table 5. The ICC values were high, indicating that these measurements could discriminate between the individuals in the sample.

DISCUSSION

The tumor population in this study was deliberately a highly selected one. We wanted to study tumors that are usually considered to be difficult to classify as benign or malignant. Clinically it would not be reasonable to perform a contrast examination in cases of tumors where

Copyright © 2009 ISUOG. Published by John Wiley & Sons, Ltd.

the ultrasound examiner is certain about the nature of the mass. Hence masses in which the ultrasound morphology was clearly suggestive of dermoid cyst, hydrosalpinx or peritoneal cyst, unilocular cysts, and multilocular cysts with fewer than 10 locules were not eligible for inclusion. When using pattern recognition, multilocular cysts with more than 10 locules are often considered difficult to classify as benign or malignant[18]. However, we excluded the two multilocular cysts with more than 10 locules that were examined with contrast to obtain a more homogeneous study population consisting only of tumors with solid components.

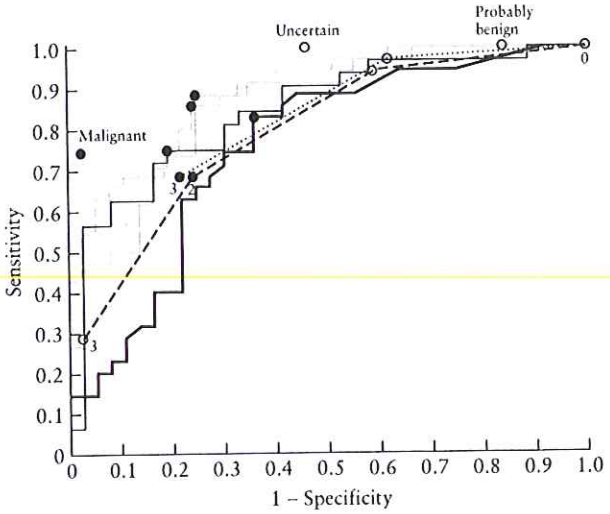

Figure 6 Receiver–operating characteristics (ROC) curves of the diagnostic methods with the largest area under the ROC curve (AUC). The AUCs of the following methods are shown: pattern recognition i.e. subjective assessment of the risk of malignancy using four levels of diagnostic confidence (benign, probably benign, uncertain, malignant) (AUC 0.93;); the best logistic regression model including a contrast variable (AUC 0.89;); the best logistic regression model without a contrast variable (AUC 0.86;); the size of the largest solid component (AUC 0.75; ——); color score in any solid component (AUC 0.79; – – –); contrast score in any solid component (AUC 0.78; .....); area under the time–intensity curve in the region of interest (AUC 0.84; ——). ●, Sensitivity and false-positive rate of the optimal cut-off. ○, Sensitivity and false-positive rate of the different levels for categorical data. Numbers 3, 2 and 0 indicate cut-offs in the color or contrast scoring systems.

In our selected tumor population comprising only tumors with solid components, many ovarian masses were considered difficult to characterize as benign or malignant when using pattern recognition. The examiners were uncertain about the nature of the tumor in about one third of the cases, while in an ordinary tumor population experienced examiners are uncertain in less than 10% of cases[18]. In this selected tumor population, the results of ultrasound contrast examination (CnTI–SonoVue) differed between benign and malignant tumors. However, there was substantial overlap between the benign and malignant tumors, and in particular between benign and borderline tumors. Even though the contrast variables seemed to have a better diagnostic performance with regard to discrimination between benign and malignant masses than the other ultrasound variables, their performance was not clearly superior, and it was not superior to that of pattern recognition. Our logistic regression model, including tumor size and peak intensity in the ROI, had the second largest area under the ROC curve (0.89) of all the diagnostic tests examined, only pattern recognition having a larger area (0.93). On the other hand, a logistic regression model not including any contrast variable performed almost as well (area under the ROC curve 0.86) as the model including a contrast variable. The logistic regression models are likely to manifest much poorer performance when tested prospectively in another population. No contrast variable – and no other variable – fulfilled the criteria of an excellent diagnostic test, i.e. none was associated with an LR+ > 10 and at the same time an LR− < 0.1[19,20]. This is perhaps not surprising given the high proportion of difficult tumors in our study sample. Qualitative contrast analysis did not perform any better than qualitative power Doppler analysis: the sensitivity with regard to malignancy of detectable flow in papillary projections, septa or solid parts after contrast injection was only slightly superior to that of detectable flow at power Doppler examination, but the false positive rate was much higher (Table 2).

To the best of our knowledge, there are only three published studies describing analysis of time–intensity curves in adnexal masses after injection of ultrasound contrast[21–23]. The study by Fleischer *et al.*[22], in which

Table 5 Intraobserver reproducibility of Qontrax^{TM} analysis

Parameter	*Measurement value (mean (SD))**	*Difference between two measurements*		
		Mean (95% CI)	*Limits of agreement*	*ICC*
ROI analysis				
Peak	13.03 (7.26)	−0.83 (−1.56 to −0.10)	−4.08 to 2.42	0.97
AUC	1150.01 (779.70)	−93.18 (−200.08 to 13.72)	−527.86 to 341.51	0.95
MAX analysis				
Peak	24.28 (11.59)	−0.01 (−1.36 to 1.34)	−5.94 to 5.92	0.97
AUC	1664.63 (1194.02)	50.93 (−45.67 to 147.50)	−341.98 to 443.83	0.99

*Mean and SD were calculated for all 44 values (22 cases). AUC, area under the curve; ICC, intraclass correlation coefficient; MAX analysis, analysis using the time–intensity curve from the pixel corresponding to the maximum peak value (MAX) in the region of interest (ROI); ROI analysis, analysis using the average of the time–intensity curves of all pixels included in the ROI.

Copyright © 2009 ISUOG. Published by John Wiley & Sons, Ltd.

a second-generation contrast agent was used and a technique similar but not identical to ours was applied, included only 17 patients with 23 tumors (14 benign and nine malignant tumors). In agreement with the findings in our study, the AUC and peak values were highest in the malignant tumors, and TTP values were similar in benign and malignant tumors. No borderline tumors were included in their study. The studies by Ordén *et al.*[21] and Marret *et al.*[23] differ from ours in several aspects: a first-generation contrast agent was used (Levovist®, Schering, Berlin, Germany), power Doppler signals were analyzed instead of signals generated by ultrasound contrast, and the method of analyzing the contrast results and the indices used to describe them were different. The mix of tumors was also dissimilar: 57% of the tumors in the study by Ordén *et al.* contained solid components vs. 100% in our study, 20% vs. 30% were invasively malignant (the tumor stage was not described in the study by Ordén *et el.*), 6% vs. 11% were borderline tumors, and while endometriomas, dermoid cysts, functional cysts and hydrosalpinx were common in the study by Ordén *et al.*, they were uncommon in ours, cystadenomas constituting the major part of our benign tumors. In the study by Marret *et al.*, which included 99 patients, there were only two borderline tumors. Of the primary invasive malignancies 77% (17/22) were stage III or IV vs. 69% in our study, and the histopathology of the benign tumors was not described. Despite these differences, the results of the studies of Ordén *et al.* and Marret *et al.* and those of our study point in the same direction. Both in our study and in that of Ordén *et al.*, the AUC and peak values were highest in the invasively malignant tumors, and in both studies TTP values (called rising time in the study by Ordén *et al.*) were highest in the borderline tumors and similarly low in the benign and invasively malignant tumors. The latter result is intriguing. However, the results for borderline tumors must be interpreted with great caution, because the number of borderline tumors was low in both studies (four vs. nine), and the differences in TTP between benign and borderline tumors were not statistically significant. Both in our study and in that of Marret *et al.* the AUC was one of the best discriminators between benign and malignant tumors, with higher values in the malignancies, but in the study of Marret *et al.*, wash-out time and half wash-out time seem to have been better discriminators than in our study. Indeed, the sensitivity and specificity with regard to malignancy of the contrast parameters were much higher in the studies by Marret *et al.* and Ordén *et al.* than in ours. This is likely to be explained – at least partly – by differences in tumor mix, e.g. by the cancers in the Marret *et al.* study being more advanced than those in ours.

The role of the CnTI–SonoVue technique in gynecological clinical practice is still uncertain. SonoVue is a safe drug that has been used for diagnostic purposes in thousands of patients[24], but patient acceptability of the CnTI–SonoVue technique is not known. The drug is rather expensive, the technique involves an intravenous injection, the acquisition of a clip for quantitative analysis is difficult (no movement of the probe or patient is allowed during the acquisition), and the quantitative analysis of the clips is cumbersome. Clearly, qualitative analysis is much less labor intensive. A diagnosis of a benign lesion is highly likely if no perfusion can be detected within presumably solid components of ovarian masses when using the CnTI–SonoVue technique. This is likely to be clinically useful. However, qualitative evaluation of contrast-enhanced ultrasound examination does not seem to improve discrimination between benign and malignant tumors with papillary projections[25].

The greatest challenge for an ultrasound examiner is to discriminate between benign and borderline tumors, in particular between benign cystadenomas/cystadenofibromas with papillary projections and borderline tumors with papillary projections[18]. In view of this, we believe that it would be important to conduct a study to determine the value of time–intensity curve analysis in cysts with papillary projections, where the ultrasound examiner is uncertain about the diagnosis. The problem is that such cysts are rare. Therefore, such a study would need to include a larger number of ultrasound centers than the current study. The issues of costs and patient acceptability of the CnTI–SonoVue technique also need to be evaluated.

ACKNOWLEDGMENTS

This study was supported by Bracco Imaging S.p.A. Via XXV Aprile 4, 20097 San Donato Milanese, Milan, Italy and by ESAOTE, Via Siffredi, 58r 16153, Genoa Italy, the Swedish Medical Research Council (grants nos. K2001-72X 11605-06A, K2002-72X-11605-07B, K2004-73X-11605-09A and K2006-73X-11605-11-3) by funds administered by Malmö University Hospital and by GOA, IAP DYSCO, FWO G.0302.07, IWT TBM project 070706.

The ESAOTE Company provided appropriate ultrasound equipment to the participating centers and the BRACCO company provided the contrast agent SonoVue free of charge. The ESAOTE Company played no role in the planning of the study, the analysis of data or the writing of the manuscript. The Bracco company was involved in the planning of the study but had no role in the statistical analysis of the data or in the writing of the manuscript.

REFERENCES

1. Jemal A, Thomas A, Murray T, Thun M. Cancer statistics 2000. *CA Cancer J Clin* 2002; 52: 23–47.
2. Osmers R. Sonographic evaluation of ovarian masses and its therapeutical implications. *Ultrasound Obstet Gynecol* 1996; 8: 217–222.
3. Valentin L. Prospective cross-validation of Doppler ultrasound examination and gray-scale ultrasound imaging for discrimination of benign and malignant pelvic masses. *Ultrasound Obstet Gynecol* 1999; 14: 273–283.
4. Tekay A, Jouppila P. Validity of pulsatility and resistance indices in classification of adnexal tumors with transvaginal color Doppler ultrasound. *Ultrasound Obstet Gynecol* 1992; 2: 338–344.

Copyright © 2009 ISUOG. Published by John Wiley & Sons, Ltd.

5. Valentin L, Sladkevicius P, Marsal K. Limited contribution of Doppler velocimetry to the differential diagnosis of extrauterine pelvic tumors. *Obstet Gynecol* 1994; 83: 425–433.
6. Valentin L. Pattern recognition of pelvic masses by gray-scale ultrasound imaging: the contribution of Doppler ultrasound. *Ultrasound Obstet Gynecol* 1999; 14: 338–347.
7. Becker H, Burns PN. *Handbook of contrast Echocardiography. LV function and Myocardial Perfusion.* Springer Verlag: Heidelberg, 2000; 25–26.
8. Testa AC, Ferrandina G, Fruscella E, Van Holsbeke C, Ferrazzi E, Leone FP, Arduini, Exacoustos C, Bokor D, Scambia G, Timmerman D. The use of contrasted transvaginal sonography in the diagnosis of gynecologic diseases: a preliminary study. *J Ultrasound Med* 2005; 24: 1267–1278.
9. Agati L, Funaro S, Veneroso G, Volponi C, De Maio F, Madonna MP, Fedele F. Non-invasive assessment of myocardial perfusion by intravenous contrast echocardiography: state of the art. *Ital Heart J* 2001; 2: 403–407.
10. Timmerman D, Valentin L, Bourne TH, Collins WP, Verrelst H, Vergote I; International Ovarian Tumor Analysis (IOTA) Group. Terms, definitions and measurements to describe the sonographic features of adnexal tumors: a consensus opinion from the International Ovarian Tumor Analysis (IOTA) Group. *Ultrasound Obstet Gynecol* 2000; 16: 500–505.
11. Rubaltelli L, Corradin S, Dorigo A, Tregnaghi A, Adami F, Rossi CR, Stramare R. Automated quantitative evaluation of lymph node perfusion on contrast-enhanced sonography. *AJR Am J Roentgenol* 2007; 188: 977–983.
12. Huang-Wei C, Bleuzen A, Olar M, Portalez D, Roumy J, Trillaud H, Tranquart F. Role of parametric imaging in contrast-enhanced sonography of hepatic focal nodular hyperplasia. *J Clin Ultrasound* 2006; 34: 367–373.
13. Sheperd JH. Revised FIGO staging for gynaecological cancer. *Br J Obstet Gynaecol* 1989; 96: 889–892.
14. Mantel N, Haenszel W. Statistical aspects of the analysis of data from retrospective studies of disease. *J Natl Cancer Inst* 1959; 22: 719–748.
15. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* 1986; 1: 307–310.
16. Li L, Nawar S. Reliability analysis: calculate and compare intra-class correlation coefficients (ICC) in SAS. Proceedings of the 20th North East SAS Users Group Congress (NESUG), Statistics and Data Analysis, 2007; 1–4.
17. Timmerman D, Testa AC, Bourne T, Ferrazzi E, Ameye L, Konstantinovic ML, Van Calster B, Collins WP, Vergote I, Van Huffel S, Valentin L. Logistic regression model to distinguish between the benign and malignant adnexal mass before surgery: a multicenter study by the International Ovarian Tumor Analysis Group. *J Clin Oncol* 2005; 23: 8794–8801.
18. Valentin L, Ameye L, Jurkovic D, Metzger U, Lécuru F, Van Huffel S, Timmerman D. Which extrauterine pelvic masses are difficult to correctly classify as benign or malignant on the basis of ultrasound findings and is there a way of making a correct diagnosis? *Ultrasound Obstet Gynecol* 2006; 27: 438–444.
19. Jaeschke R, Guyatt G, Sackett DL. Users' guides to the medical literature. III. How to use an article about a diagnostic test. A. Are the results of the study valid? Evidence-Based Medicine Working Group. *JAMA* 1994; 271: 389–391.
20. Jaeschke R, Guyatt GH, Sackett DL. Users' guides to the medical literature. III. How to use an article about a diagnostic test. B. What are the results and will they help me in caring for my patients? The Evidence-Based Medicine Working Group. *JAMA* 1994; 271: 703–707.
21. Ordén MR, Jurvelin JS, Kirkinen PP. Kinetics of a US contrast agent in benign and malignant adnexal tumors. *Radiology* 2003; 226: 405–410.
22. Fleischer AC, Lyshchik A, Jones HW Jr, Crispens M, Loveless M, Andreotti RF, Williams PK, Fishman DA. Contrast-enhanced transvaginal sonography of benign versus malignant ovarian masses: preliminary findings. *J Ultrasound Med* 2008; 27: 1011–1018.
23. Marret H, Sauget S, Giraudeau B, Brewer M, Ranger-Moore J, Body G, Tranquart F. Contrast-enhanced sonography helps in discrimination of benign from malignant adnexal masses. *J Ultrasound Med* 2004; 23: 1629–1639.
24. Piscaglia F, Bolondi L; Italian Society for Ultrasound in Medicine and Biology (SIUMB) Study Group on Ultrasound Contrast Agent. The safety of Sonovue in abdominal applications: retrospective analysis of 23 188 investigations. *Ultrasound Med Biol* 2006; 32: 1369–1375.
25. Testa AC, Timmerman D, Exacoustos C, Fruscella E, Van Holsbeke C, Bokor D, Arduini D, Scambia G, Ferrandina G. The role of CnTI-SonoVue in the diagnosis of ovarian masses with papillary projections: a preliminary study. *Ultrasound Obstet Gynecol* 2007; 29: 512–516.

SUPPORTING INFORMATION ON THE INTERNET

The following supporting information may be found in the online version of this article:

Table S1 Ultrasound characteristics of benign, borderline and malignant tumors in included patients: gray-scale ultrasound, Doppler ultrasound and qualitative contrast parameters.

Copyright © 2009 ISUOG. Published by John Wiley & Sons, Ltd.