

Simple ultrasound-based rules for the diagnosis of ovarian cancer

D. TIMMERMAN*, A. C. TESTA†, T. BOURNE*, L. AMEYE‡, D. JURKOVIC§, C. VAN HOLSBEKE*, D. PALADINI¶, B. VAN CALSTER‡, I. VERGOTE*, S. VAN HUFFEL‡ and L. VALENTIN**

*Department of Obstetrics and Gynecology, University Hospitals KU Leuven ‡Department of Electrical Engineering, ESAT-SCD, Katholieke Universiteit Leuven, Leuven, Belgium, †Istituto di Clinica Ostetrica e Ginecologica, Università Cattolica del Sacro Cuore, Rome, ¶Universita degli Studi di Napoli, Naples, Italy, §Department of Obstetrics and Gynaecology, King's College Hospital London, UK and **Department of Obstetrics and Gynecology, Malmö University Hospital, Lund University, Malmö, Sweden

KEYWORDS: color Doppler imaging; ovarian neoplasms; ultrasonography

ABSTRACT

Objective To derive simple and clinically useful ultrasound-based rules for discriminating between benign and malignant adnexal masses.

Methods In a multicenter study involving nine centers consecutive patients with persistent adnexal tumors underwent transvaginal gray-scale and Doppler ultrasound examination using a standardized examination technique and standardized terms and definitions. Information on 42 gray-scale ultrasound variables and six Doppler variables was collected and entered into a research protocol. When developing simple ultrasound-based rules to predict malignancy (M-rules) we chose the ultrasound variable or the combination of ultrasound variables that had the highest positive predictive value (PPV) with regard to malignancy; when developing simple rules to predict a benign tumor (B-rules) we chose the ultrasound variable or the combination of ultrasound variables that had the lowest PPV with regard to malignancy. We selected ten rules that were in agreement with our clinical experience and were applicable to at least 30 tumors and then tested them prospectively on 507 tumors examined in three of the nine centers.

Results 1066 patients with 1233 adnexal tumors were included. There were 903 benign tumors (73%) and 330 malignant tumors (27%). In 167 patients the tumors were bilateral. We selected five simple rules to predict malignancy (M-rules): (1) irregular solid tumor; (2) ascites; (3) at least four papillary structures; (4) irregular multilocular–solid tumor with a largest diameter of at least 100 mm; and (5) very high color

content on color Doppler examination. We chose five simple rules to suggest a benign tumor (B-rules): (1) unilocular cyst; (2) presence of solid components where the largest solid component is < 7 mm in largest diameter; (3) acoustic shadows; (4) smooth multilocular tumor less than 100 mm in largest diameter; and (5) no detectable blood flow on Doppler examination. These ten rules were applicable to 76% of all tumors, where they resulted in a sensitivity of 93%, specificity of 90%, positive likelihood ratio (LR+) of 9.45 and negative likelihood ratio (LR–) of 0.08. When prospectively tested the rules were applicable in 76% (386/507) of the tumors, where they had a sensitivity of 95% (106/112), a specificity of 91% (249/274), LR+ of 10.37, and LR– of 0.06.

Conclusion Most adnexal tumors in an ordinary tumor population can be correctly classified as benign or malignant using simple ultrasound-based rules. For tumors that cannot be classified using simple rules, ultrasound examination by an expert examiner might be useful. Copyright © 2008 ISUOG. Published by John Wiley & Sons, Ltd.

INTRODUCTION

Correct characterization of adnexal masses is important for optimal patient management. Masses felt to be benign can be managed expectantly or with minimal-access surgery. Malignant pathology will require referral to an appropriately trained gynecological oncologist. The morphological features of an adnexal mass can be used to

Correspondence to: Prof. D. Timmerman, Department of Obstetrics and Gynecology, University Hospitals, KU Leuven, Herestraat 49, B-3000 Leuven, Belgium (e-mail: dirk.timmerman@uzleuven.be)

Accepted: 21 March 2008

indicate the likelihood of it being benign or malignant¹. Granberg *et al.* demonstrated that malignancy is unlikely in the presence of a unilocular cyst with smooth walls, but that the presence of solid projections into the cyst cavity increases the risk of malignancy¹. Morphological scoring systems weighting the relative importance of a variety of structural features such as the presence or absence of papillary projections or septations have been developed^{2,3}. These scoring systems have been modified to exclude the false positive test results associated with benign teratoma^{4,5}. In the studies cited the ultrasound variables and their relative importance were arbitrarily defined. For example, there was no general agreement as to when a solid protrusion from a cyst wall projects sufficiently into the cyst cavity for it to be termed a solid papillary projection. As a result of this, there is a lack of consistency between ultrasound units, and it is difficult to compare results of published reports. Color Doppler ultrasonography has also been used in order to evaluate the likelihood of malignancy in adnexal masses. In 1989 an association between low-impedance blood flow and malignancy was described⁶. Later, significant overlap in Doppler parameters between benign and malignant tumors was reported⁷⁻⁹. More recently, the use of a color score – a semi-quantitative attempt to evaluate the amount of blood flow in any given tumor – has been suggested^{9,10}. There is a worrying lack of agreement regarding which methodology to use to obtain blood-flow information and regarding how to quantify and describe blood flow in adnexal tumors. An attempt was made to achieve some degree of consistency when the International Ovarian Tumor Analysis (IOTA) group published a consensus paper defining morphological features of ovarian masses and suggesting a standardized examination technique¹¹. In the IOTA study a large number of adnexal tumors were examined in a standardized manner in order to build a database that could be used to develop new approaches to the preoperative classification of adnexal masses¹².

The aim of the current study was to develop clinically useful simple ultrasound rules that can be used to classify most adnexal masses as benign or malignant.

PATIENTS AND METHODS

We used the data collected in the IOTA study, which has been described in detail elsewhere¹². Briefly, it is a multicenter study comprising nine centers: Malmö (Sweden), Leuven (Belgium), London (UK), Rome, Naples, Monza, Milan (Italy) and two centers in Paris (France). Women with at least one persisting adnexal mass underwent transvaginal gray-scale and color Doppler ultrasound examination by an experienced ultrasound examiner using a standardized examination technique and standardized terms and definitions¹¹. A transvaginal scan was performed in all cases. Transabdominal sonography was used to examine large masses that could not be seen in their entirety using a transvaginal probe. Gray-scale and color Doppler imaging was performed to obtain 42

gray-scale ultrasound variables and six Doppler variables to characterize each adnexal mass. These variables have been illustrated, described and defined previously¹¹.

The final outcome measures of the study were the histological diagnoses and in case of malignancy the surgical stage. Surgery was performed if a mass was present 6–12 weeks after the initial scan. In case of symptomatic masses, suspected malignancy, or at the patient's request, surgery was performed earlier. All excised tissues were sampled for histological examination at the local centers. Tumors were classified according to the criteria recommended by the International Federation of Gynecology and Obstetrics¹³. The degree of differentiation of malignant tumors was recorded. The pathological samples from about 10% of the patients were randomly selected for peer review by Professor Philippe Moerman, Katholieke Universiteit, Leuven. If there was disagreement (malignant or benign) between the original pathology report and that of the external reviewer the patient was excluded from the study.

Statistical analysis was carried out using the SAS system release 9.1 (SAS Institute Inc., Cary, NC, USA). In the statistical analyses borderline tumors were classified as malignant. Student's *t*-test and the Mann–Whitney *U*-test were used to test the statistical significance of differences in continuous data between benign and malignant tumors, and the Chi-square test and Fisher's exact test were used to test the statistical significance of differences in categorical data. Cut-off values to define the parameters used to predict malignancy for continuous variables were chosen using the minimum *P*-value approach¹⁴. Receiver–operating characteristics curves were also constructed in order to check whether the cut-offs chosen using the minimal *P*-value method were reasonable¹⁵. Tumors with values \geq the cut-off chosen were classified as malignant and tumors with values $<$ the cut-off were classified as benign, with the exception that tumors with pulsatility index (PI) and resistance index (RI) values \geq the cut-off were classified as benign. Numerous combinations of two and three ultrasound variables were tested for their ability to predict benignity or malignancy. For example, tumor type was combined with color score and regularity of cyst wall. When developing simple ultrasound-based rules to predict malignancy (M-rules) we selected the ultrasound variable or the combination of ultrasound variables that had the highest positive predictive value (PPV) with regard to malignancy; when developing simple rules to predict a benign tumor (B-rules) we selected the ultrasound variable or the combination of ultrasound variables that had the lowest PPV with regard to malignancy. In practice we manually chose rules that were in agreement with our clinical experience and were applicable to a large number of tumors. The characteristics tested were not common to all tumors, for example, the B-rule 'unilocular cyst with irregular walls measuring ≥ 100 mm' had a PPV of 0%, but this rule was applicable to only two masses in the entire database. A rule based on such a small number of tumors is unreliable. We made the arbitrary decision

to only consider rules that were applicable to at least 30 tumors. We investigated the diagnostic performance of all variables, i.e. sensitivity, specificity, positive likelihood ratio (LR+) and negative likelihood ratio (LR-). We also determined the diagnostic performance of numerous combinations of variables. Finally ten simple rules that had high sensitivity and specificity and were applicable to a large number of tumors were chosen. The ten simple rules were then tested prospectively on 507 tumors that had been examined in three of the nine centers after the conclusion of the first phase of the IOTA study but using the IOTA study protocol described above.

RESULTS

Between June 1999 and June 2002 we enrolled 1066 patients with 1233 adnexal tumors in the study. In 167 patients the tumors were bilateral. In ten (1%) patients a malignant tumor on one side and a benign tumor on the other side were found. Nine hundred and three (73%) tumors were benign and 330 (27%) were malignant. The prospective test group consisted of 507 new tumors collected in Leuven, Malmö and Rome between July 2002 and December 2005. The tumor histology and stage of

malignant tumors are shown in Table 1. Tables 2, 3, and 4 summarize demographic background information and ultrasound findings in women with benign and malignant tumors. Cut-off values to predict malignancy for continuous ultrasound variables are shown in Table 5. The likelihood ratios presented in Tables 3, 4 and 5 show that no single demographic or ultrasound variable could reliably discriminate between benign and malignant adnexal tumors. In the case of continuous variables, this was true both when the cut-off chosen was tested in tumors where the variable was present (Table 5) and when it was tested in the whole study population (results not shown), e.g. it was true both when the cut-off for height of the largest papillary projection was used to discriminate between benign and malignant tumors with papillary projections (LR+ 2.21, LR- 0.39) and when it was used to discriminate between benign and malignant tumors in the whole tumor population (LR+ 5.22, LR- 0.73). The presence of ascites, a solid tumor or a high color content using color Doppler increased the risk of malignancy (LR+ 14.52, 5.09 and 6.17, respectively), while the presence of a unilocular cyst, acoustic shadowing, and the absence of detectable tumor blood flow decreased the risk (LR+ 0.04, 0.12 and 0.16, respectively). A depth of more than

Table 1 Histology of tumors in study group ($n = 1233$) and prospective test group ($n = 507$)

Histological result	Development set				Prospective test set			
	n	Not classifiable using the simple rules (n (%))	Classifiable using the simple rules (n (%))		n	Not classifiable using the simple rules (n (%))	Classifiable using the simple rules (n (%))	
			Correctly classified	Incorrectly classified			Correctly classified	Incorrectly classified
Benign pathology								
Endometrioma	242	50 (21)	182 (75)	10 (4)	101	14 (14)	84 (83)	3 (3)
Dermoid/teratoma	131	21 (16)	107 (82)	3 (2)	55	7 (13)	48 (87)	—
Simple cyst*	161	42 (26)	117 (73)	2 (1)	59	8 (14)	50 (85)	1 (2)
Hydrosalpinx†	41	16 (39)	22 (54)	3 (7)	16	7 (44)	8 (50)	1 (6)
Peritoneal (pseudo)cyst	7	4 (57)	1 (14)	2 (29)	5	2 (40)	3 (60)	—
Abscess	14	4 (29)	2 (14)	8 (57)	4	2 (50)	1 (25)	1 (25)
Fibroma‡	40	15 (38)	11 (28)	14 (35)	29	11 (38)	9 (31)	9 (31)
Serous cystadenoma	151	47 (31)	93 (62)	11 (7)	48	19 (40)	24 (50)	5 (10)
Mucinous cystadenoma	99	36 (36)	54 (55)	9 (9)	38	17 (45)	18 (47)	3 (8)
Rare benign§	17	9 (53)	5 (29)	3 (18)	9	3 (33)	4 (44)	2 (22)
Malignant pathology								
Primary invasive								
Stage I	47	11 (23)	32 (68)	4 (9)	24	8 (33)	14 (58)	2 (8)
Stage II	15	1 (7)	14 (93)	—	5	1 (20)	4 (80)	—
Stage III	105	11 (10)	91 (87)	3 (3)	59	5 (8)	52 (88)	2 (3)
Stage IV	23	1 (4)	22 (96)	—	11	5 (45)	6 (55)	—
Rare malignant¶	27	4 (15)	23 (85)	—	4	1 (25)	2 (50)	1 (25)
Borderline								
Stage I	56	16 (29)	31 (55)	9 (16)	15	4 (27)	10 (67)	1 (7)
Stage II	7	2 (29)	5 (71)	—	—	—	—	—
Stage III	—	—	—	—	5	1 (20)	4 (80)	—
Metastatic	50	6 (12)	41 (82)	3 (6)	20	6 (30)	14 (70)	—
Total	1233	296	853	84	507	121	355	31

*Including parasalpingeal cyst and inclusion cyst, and normal ovary. †Including salpingitis. ‡Including leiomyoma. §Including struma ovarii, Brenner tumor, Sertoli cell tumor, stromal cell tumor, Schwannoma, and lymphangioma. ¶Including granulosa cell tumor, Leydig cell tumor, dysgerminoma, gynandroblastoma, leiomyosarcoma, immature teratoma, malignant mixed Müllerian tumor, small cell cancer, Brenner cancer, carcinosarcoma, choriocarcinoma, and yolk sac tumor.

Table 2 Characteristics of benign and malignant tumors (continuous variables)

Parameter	Malignant		Benign		P
	n	Median (range)	n	Median (range)	
Demographic data					
Age (years)	266	56 (17–94)	800	42 (17–90)	< 0.01
Years past menopause	145	12 (0–44)	229	10 (1–40)	0.14
Gray-scale ultrasound findings					
Maximum diameter of the lesion (mm)	330	85 (8–410)	903	59 (8–320)	< 0.01
Volume of the lesion (mL)*	330	200 (0.1–11 829)	903	63 (0.2–7781)	< 0.01
Fluid in POD (mm)	144	24 (3–100)	140	12 (2–61)	< 0.01
Septum (mm)	167	4 (1–20)	372	2 (1–20)	< 0.01
Papillary projection height (mm)	139	14 (3–62)	161	7 (2–62)	< 0.01
Max papillary projection D (mm)	139	18 (4–110)	161	10 (3–90)	< 0.01
Volume papillary projection (mL)*	139	2 (0.008–226)	161	0.2 (0.006–77)	< 0.01
Ratio papillary projection : lesion	139	0.006 (0–0.420)	161	0.003 (0–0.456)	< 0.01
Papillary projection number	139	4 (1–≥ 4)	161	1 (1–≥ 4)	< 0.01
Locule number	330	2 (0–> 10)	903	1 (0–> 10)	< 0.01
Max solid diameter (mm)	303	46 (4–214)	334	21 (3–230)	< 0.01
Solid volume (mL)*	303	26 (0.008–2291)	334	2 (0.006–1978)	< 0.01
Ratio solid : lesion	303	0.265 (0–1)	334	0.035 (0–1)	< 0.01
Doppler results†					
Pulsatility index	300	0.76 (0.23–3.31)	556	0.95 (0.13–7.30)	< 0.01
Resistance index	300	0.52 (0.17–1.0)	556	0.59 (0.12–1.0)	< 0.01
Peak systolic velocity (cm/s)	300	23.9 (2.3–202)	556	11.2 (2.0–85.5)	< 0.01
TAMXV (cm/s)	293	16.0 (1.7–137)	546	6.6 (1.0–60.0)	< 0.01

*Volumes were calculated using the following formula: diameter 1 × diameter 2 × diameter 3 × π/6, the diameters being three orthogonal diameters measured using calipers on the frozen ultrasound image. †Results are those for the tumor vessel with the highest TAMXV. Fluid in POD, fluid in anteroposterior plane of pouch of Douglas; Septum, thickness of the thickest septum where it appeared to be at its thickest; Papillary projection height, height of largest papillary projection; Max papillary projection D, maximal diameter of the largest papillary projection; Volume papillary projection, volume of the largest papillary projection; Ratio papillary projection : lesion, ratio between the volume of the largest papillary projection and the volume of the lesion; Papillary projection number, number of separate papillary projections (1, 2, 3 or ≥ 4); Locule number, number of cyst locules (0, 1, 2, 3, 4, 5–10 or > 10); Max solid diameter, maximal diameter of the largest solid component; Solid volume, volume of the largest solid component; Ratio solid : lesion, ratio between the volume of the largest solid component and the volume of the lesion; TAMXV, time-averaged maximum velocity.

15 mm of fluid in the pouch of Douglas increased the risk of malignancy at most moderately (LR+ 5.60 when tested in the whole population, LR+ 1.81 when tested only in women with fluid in the pouch of Douglas), whilst absence of a solid component ≥ 2 mL decreased the risk slightly (LR+ 0.22 when used in tumors with solid components, LR+ 0.23 when used in the whole tumor population).

The best combinations of tumor type with one or two additional ultrasound variables to predict malignancy or benignity are presented in Tables 6, 7 and 8. Five of 377 unilocular cysts were malignant. Neither wall irregularity, size nor degree of vascularization seemed to change the risk of malignancy substantially in unilocular cysts. All 50 unilocular cysts with wall irregularities of < 3 mm in height were benign. Three of the five malignant unilocular cysts were found in patients with a partly solid mass in the contralateral adnexa suspected to be malignant by the ultrasound examiner, and both the unilocular cyst and the contralateral complex mass proved to be malignant (carcinosarcoma, Stage IIIc; mixed clear cell and endometrioid adenocarcinoma, Stage III; Krukenberg tumor). Two patients had a unilateral unilocular cyst that proved to be malignant: one simple cyst (140 × 115 × 105 mm) with color score 2 in a 26-year-old patient with a serum CA-125 level of 13 U/mL was a Stage Ia borderline malignant serous papillary

cystadenoma, and one unilocular cyst (55 × 38 × 35 mm) with color score 1 in a 42-year-old patient with a serum CA-125 level of 7 U/mL was a Stage Ia borderline malignant mucinous cystadenoma (Figure 1). Smooth-walled multilocular cysts were also rarely malignant, particularly if they had a volume < 100 mL (LR+ 0.11), and so were smooth-walled and poorly vascularized unilocular–solid cysts (LR+ 0.18). In multilocular cysts, wall irregularities of < 3 mm slightly increased the odds of malignancy: nine of 41 (22%) multilocular cysts (eight containing at least five locules and one containing two locules) with wall irregularities < 3 mm were malignant (three borderline tumors and six primary invasive tumors) vs. 13/172 (8%) multilocular cysts with smooth internal cyst walls ($P = 0.01$). For all other types of tumor the risk of malignancy was substantial. Purely solid irregular tumors were virtually always malignant irrespective of their size and irrespective of whether they were poorly or strongly vascularized. The risk of malignancy was also high in irregular unilocular–solid tumors and irregular multilocular–solid tumors, especially if they were well vascularized or large (≥ 100 mL). Ultrasound images of malignant tumors of various types are shown in Figures 1–5.

The ten rules that we finally selected to characterize ovarian masses as benign or malignant are presented in Table 9. In all, 937 (76%) of the 1233 tumors could

Table 3 Characteristics of benign and malignant tumors (binary variables)

Variable	Positive predictive value of the variable (% (n))	Presence of the variable in malignant tumors (sensitivity) (% (n))	Presence of the variable in benign tumors (false-positive rate, i.e. 1 - specificity) (% (n))	P*	LR+	LR-
Demographic data						
Personal history of ovarian cancer	57.1 (8/14)	3.0 (8/266)	0.7 (6/800)	0.01	4.29	0.98
Postmenopausal bleeding	43.3 (26/60)	17.9 (26/145)	14.8 (34/229)	0.43	1.21	0.96
Personal history of breast cancer	39.5 (15/38)	5.6 (15/266)	2.9 (23/800)	0.04	1.93	0.97
Family history of ovarian cancer	39.4 (13/33)	4.9 (13/266)	2.5 (20/800)	0.06	1.96	0.98
Postmenopausal	39.1 (169/432)	63.5 (169/266)	32.9 (263/800)	< 0.01	1.93	0.54
Hysterectomy	35.9 (28/78)	10.5 (28/266)	6.2 (50/800)	0.02	1.69	0.95
Family history of breast cancer	27.7 (33/119)	12.4 (33/266)	10.7 (86/800)	0.46	1.16	0.98
Hormonal therapy	20.0 (47/235)	17.7 (47/266)	23.5 (188/800)	0.05	0.75	1.08
Pelvic pain	18.4 (52/282)	19.6 (52/266)	28.7 (230/800)	< 0.01	0.68	1.13
Nullipara	16.7 (67/402)	25.2 (67/266)	41.9 (335/800)	< 0.01	0.60	1.29
Gray-scale ultrasound findings						
Ascites†	83.0 (112/135)	42.1 (112/266)	2.9 (23/800)	< 0.01	14.52	0.60
≥ 1 Irregular papillary projection	59.0 (115/195)	82.7 (115/139)	49.7 (80/161)	< 0.01	1.66	0.34
Irregular internal cyst wall	48.6 (263/541)	79.7 (263/330)	30.8 (278/903)	< 0.01	2.59	0.29
Papillary projection present	46.3 (139/300)	42.1 (139/330)	17.8 (161/903)	< 0.01	2.37	0.70
Bilateral masses	38.4 (83/216)	31.2 (83/266)	16.6 (133/800)	< 0.01	1.88	0.82
Incomplete septum	14.7 (14/95)	4.2 (14/330)	9.0 (81/903)	< 0.01	0.47	1.05
Acoustic shadows	4.2 (5/120)	1.5 (5/330)	12.7 (115/903)	< 0.01	0.12	1.13
Doppler results						
Blood flow in papillary projection‡	67.1 (114/170)	82.0 (114/139)	34.8 (56/161)	< 0.01	2.36	0.28
Venous blood flow only§	15.2 (14/92)	4.2 (14/330)	8.6 (78/903)	< 0.01	0.49	1.05

*Statistical significance of the difference between benign and malignant tumors with regard to the presence of the variable. †Ascites, fluid outside the pouch of Douglas. ‡Presence of blood flow within at least one papillary projection. §No arterial blood flow. LR+, positive likelihood ratio; LR-, negative likelihood ratio.

Table 4 Characteristics of benign and malignant tumors (categorical variables)

Variable	Positive predictive value of the variable (% (n))	Presence of the variable in malignant tumors (sensitivity) (% (n))	Presence of the variable in benign tumors (false-positive rate, i.e. 1 - specificity) (% (n))	P*	LR+	LR-
Type of tumor†						
Solid	65.3 (111/170)	33.6 (111/330)	6.6 (59/903)	< 0.01	5.09	0.71
Multilocular-solid	43.0 (139/323)	42.1 (139/330)	20.4 (184/903)	< 0.01	2.06	0.73
Unilocular-solid	37.1 (53/143)	16.1 (53/330)	10.0 (90/903)	0.02	1.61	0.93
Multilocular	10.3 (22/213)	6.7 (22/330)	21.2 (191/903)	< 0.01	0.32	1.18
Unilocular	1.3 (5/377)	1.5 (5/330)	41.2 (372/903)	< 0.01	0.04	1.68
Not classifiable	0.0 (0/7)	0 (0/330)	0.8 (7/903)	0.67	0.00	1.01
Presence of solid tissue	47.6 (303/637)	91.8 (303/330)	37.0 (334/903)	< 0.01	2.48	0.13
Echogenicity of cyst fluid						
No cyst fluid	63.0 (63/100)	19.1 (63/330)	4.1 (37/903)	< 0.01	4.66	0.84
Low level	28.8 (65/226)	19.7 (65/330)	17.9 (161/903)	0.94	1.10	0.98
Anechogenic	28.3 (138/488)	41.8 (138/330)	38.8 (350/903)	0.88	1.08	0.95
Hemorrhagic	25.0 (3/12)	0.9 (3/330)	1.0 (9/903)	1	0.90	1.00
Mixed echogenicity	15.5 (23/148)	7.0 (23/330)	13.8 (125/903)	< 0.01	0.51	1.08
Ground glass	14.7 (38/259)	11.5 (38/330)	24.5 (221/903)	< 0.01	0.47	1.17
Color score						
Very strong flow (Score 4)	69.4 (120/173)	36.4 (120/330)	5.9 (53/903)	< 0.01	6.17	0.68
Moderately strong flow (Score 3)	35.8 (136/380)	41.2 (136/330)	27.0 (244/903)	< 0.01	1.53	0.81
Minimal flow (Score 2)	14.7 (58/395)	17.6 (58/330)	37.3 (337/903)	< 0.01	0.47	1.31
No flow (Score 1)	5.6 (16/285)	4.8 (16/330)	29.8 (269/903)	< 0.01	0.16	1.36

*Statistical significance of the difference between benign and malignant tumors with regard to the presence of the variable. †Tumor classification according to Granberg *et al.*¹, slightly modified²¹. LR+, positive likelihood ratio; LR-, negative likelihood ratio.

Table 5 Cut-off values to predict malignancy for continuous ultrasound variables

Variable	Cut-off	Positive predictive value (% (n))	Sensitivity (% (n))	False-positive rate (1 – specificity) (% (n))	LR+	LR–
Gray-scale ultrasound						
Fluid in POD	15 mm	65.1 (108/166)	75.0 (108/144)	41.4 (58/140)	1.81	0.43
Locule number	> 10	61.0 (75/123)	22.7 (75/330)	5.3 (48/903)	4.28	0.76
Max lesion D	100 mm	51.4 (142/276)	43.0 (142/330)	14.8 (134/903)	2.90	0.67
Volume of the lesion*	215 mL	47.9 (160/334)	48.5 (160/330)	19.3 (174/903)	2.51	0.64
Septum	3 mm	47.8 (117/245)	70.1 (117/167)	34.4 (128/372)	2.04	0.46
Max solid D	35 mm	69.2 (211/305)	69.6 (211/303)	28.1 (94/334)	2.47	0.42
Max papillary	14 mm	67.4 (93/138)	66.9 (93/139)	28.0 (45/161)	2.40	0.50
Papillary number	2	65.8 (108/164)	77.7 (108/139)	34.8 (56/161)	2.23	0.34
Papillary height	10 mm	65.6 (103/157)	74.1 (103/139)	33.5 (54/161)	2.21	0.39
Solid volume	2 mL	62.4 (269/431)	88.8 (269/303)	48.5 (162/334)	1.83	0.22
Ratio papillary : lesion	0.006	58.6 (75/128)	54.0 (75/139)	32.9 (53/161)	1.64	0.69
Ratio solid : lesion	0.020	58.8 (260/442)	85.8 (260/303)	54.5 (182/334)	1.57	0.31
Doppler variables†						
Pulsatility index	0.6	64.8 (92/142)	30.7 (92/300)	9.0 (50/556)	3.41	0.76
TAMXV	10 cm/s	56.6 (215/380)	71.7 (215/300)	29.5 (164/556)	2.41	0.40
Resistance index	0.5	55.0 (138/251)	46.0 (138/300)	20.3 (113/556)	2.26	0.68
Peak systolic velocity	15 cm/s	54.1 (222/410)	74.0 (222/300)	33.8 (188/556)	2.19	0.39

Cut-off chosen using the minimum *P*-value approach and cross-checked with receiver–operating characteristic curves to confirm that it was reasonable. *Volumes were calculated using the following formula: diameter 1 × diameter 2 × diameter 3 × $\pi/6$, the diameters being three orthogonal diameters measured using calipers on the frozen ultrasound image. †Results are those for the tumor vessel with the highest TAMXV. LR+, positive likelihood ratio; LR–, negative likelihood ratio; POD, pouch of Douglas; Locule number, number of cyst locules; Max lesion D, maximal diameter of the lesion; Septum, thickness of thickest septum where it appeared to be at its thickest; Max solid D, maximal diameter of the largest solid component; Max papillary, maximal diameter of the largest papillary projection; Papillary number, number of papillary projections; Papillary height, height of largest papillary projection; Solid volume, volume of largest solid component; Ratio papillary : lesion, ratio between the volume of the largest papillary projection and the volume of the lesion; Ratio solid : lesion, ratio between the volume of the largest solid component and the volume of the lesion; TAMXV, time-averaged maximum velocity.

Table 6 Diagnostic performance of a sonographic classification system using tumor type and wall regularity

Type of tumor*	Wall regularity	Positive predictive value (% (n))	Sensitivity (% (n))	False-positive rate (1 – specificity) (% (n))	LR+	LR–
Unilocular	Smooth	1.5 (5/327)	1.5 (5/330)	35.7 (322/903)	0.04	1.53
	Irregular	0.0 (0/50)	0.0 (0/330)	5.5 (50/903)	0.00	1.06
Multilocular	Smooth	7.6 (13/172)	3.9 (13/330)	17.6 (159/903)	0.22	1.17
	Irregular	22.0 (9/41)	2.7 (9/330)	3.5 (32/903)	0.77	1.01
Unilocular–solid	Smooth	22.7 (5/22)	1.5 (5/330)	1.9 (17/903)	0.81	1.00
	Irregular	39.7 (48/121)	14.6 (48/330)	8.1 (73/903)	1.80	0.93
Multilocular–solid	Smooth	17.1 (14/82)	4.2 (14/330)	7.5 (68/903)	0.56	1.04
	Irregular	51.9 (125/241)	37.9 (125/330)	12.8 (116/903)	2.95	0.71
Solid	Smooth	36.1 (30/83)	9.1 (30/330)	5.9 (53/903)	1.55	0.97
	Irregular	93.1 (81/87)	24.6 (81/330)	0.7 (6/903)	37.2	0.76
Unclassifiable		0 (0/7)	0 (0/330)	0.8 (7/903)	0.00	1.01
Solid component or irregular wall		42.8 (312/729)	94.5 (312/330)	46.2 (417/903)	2.05	0.10
Solid component and irregular wall		56.6 (254/449)	77.0 (254/330)	21.6 (195/903)	3.56	0.29

*Modified²¹ after Granberg *et al.*¹. LR+, positive likelihood ratio; LR–, negative likelihood ratio.

be classified as benign or malignant using these rules with a sensitivity of 93% (259/278), a specificity of 90% (594/659), LR+ of 9.45 and LR– of 0.08. The positive and negative predictive values were 80% (259/324) and 97% (594/613). A total of 296 (24%) tumors could not be classified using the simple rules. In 229 of these 296 tumors (i.e. in 19% of all the tumors) no rule was applicable (186 benign, 17 borderline, 21 primary

invasive and five metastatic tumors), and in 67 tumors (i.e. in 5% of all the tumors) both M-rules and B-rules were applicable (58 benign tumors, one borderline, seven primary invasive and one metastatic tumor). Benign tumors were more common among unclassifiable masses than among classifiable ones (82% vs. 70%, i.e. 244/296 vs. 659/937). When prospectively tested, the ten simple rules were applicable in 76% (386/507) of the

Table 7 Diagnostic performance of a sonographic classification system using tumor type, wall regularity and color Doppler score

Type of tumor*	Wall regularity	Color score	Positive predictive value (% (n))	Sensitivity (% (n))	False-positive rate (1 - specificity) (% (n))	LR+	LR-
Unilocular	Smooth	1-2	1.5 (4/268)	1.2 (4/330)	29.2 (264/903)	0.04	1.40
		3-4	1.7 (1/59)	0.3 (1/330)	6.4 (58/903)	0.05	1.07
	Irregular	1-2	0.0 (0/41)	0.0 (0/330)	4.5 (41/903)	0.00	1.05
		3-4	0.0 (0/9)	0.0 (0/330)	1.0 (9/903)	0.00	1.01
Multilocular	Smooth	1-2	7.5 (8/107)	2.4 (8/330)	11.0 (99/903)	0.22	1.10
		3-4	7.7 (5/65)	1.5 (5/330)	6.6 (60/903)	0.23	1.05
	Irregular	1-2	19.2 (5/26)	1.5 (5/330)	2.3 (21/903)	0.65	1.01
		3-4	26.7 (4/15)	1.2 (4/330)	1.2 (11/903)	0.99	1.00
Unilocular-solid	Smooth	1-2	6.3 (1/16)	0.3 (1/330)	1.7 (15/903)	0.18	1.01
		3-4	66.7 (4/6)	1.2 (4/330)	0.2 (2/903)	5.50	0.99
	Irregular	1-2	22.4 (15/67)	4.6 (15/330)	5.8 (52/903)	0.79	1.01
		3-4	61.1 (33/54)	10.0 (33/330)	2.3 (21/903)	4.29	0.92
Multilocular-solid	Smooth	1-2	13.8 (4/29)	1.2 (4/330)	2.8 (25/903)	0.44	1.02
		3-4	18.9 (10/53)	3.0 (10/330)	4.8 (43/903)	0.64	1.02
	Irregular	1-2	27.5 (19/69)	5.8 (19/330)	5.5 (50/903)	1.04	1.00
		3-4	61.6 (106/172)	32.1 (106/330)	7.3 (66/903)	4.39	0.73
Solid	Smooth	1-2	20.0 (8/40)	2.4 (8/330)	3.5 (32/903)	0.68	1.01
		3-4	51.2 (22/43)	6.7 (22/330)	2.3 (21/903)	2.86	0.96
	Irregular	1-2	90.9 (10/11)	3.0 (10/330)	0.1 (1/903)	27.5	0.97
		3-4	93.4 (71/76)	21.5 (71/330)	0.6 (5/903)	39.1	0.79

*Modified²¹ after Granberg *et al.*¹. LR+, positive likelihood ratio; LR-, negative likelihood ratio.

Table 8 Diagnostic performance of a sonographic classification system using tumor type, wall regularity and largest lesion diameter

Type of tumor*	Wall regularity	Maximum lesion diameter (mm)	Positive predictive value (% (n))	Sensitivity (% (n))	False-positive rate (1 - specificity) (% (n))	LR+	LR-
Unilocular	Smooth	< 100	1.4 (4/291)	1.2 (4/330)	31.8 (287/903)	0.04	1.45
		≥ 100	2.8 (1/36)	0.3 (1/330)	3.9 (35/903)	0.08	1.04
	Irregular	< 100	0.0 (0/48)	0.0 (0/330)	5.3 (48/903)	0.00	1.06
		≥ 100	0.0 (0/2)	0.0 (0/330)	0.2 (2/903)	0.00	1.00
Multilocular	Smooth	< 100	3.8 (5/133)	1.5 (5/330)	14.2 (128/903)	0.11	1.15
		≥ 100	20.5 (8/39)	2.4 (8/330)	3.4 (31/903)	0.71	1.01
	Irregular	< 100	10.3 (3/29)	0.9 (3/330)	2.9 (26/903)	0.32	1.02
		≥ 100	50.0 (6/12)	1.8 (6/330)	0.7 (6/903)	2.76	0.99
Unilocular-solid	Smooth	< 100	16.7 (3/18)	0.9 (3/330)	1.7 (15/903)	0.55	1.01
		≥ 100	50.0 (2/4)	0.6 (2/330)	0.2 (2/903)	2.77	1.00
	Irregular	< 100	34.7 (34/98)	10.3 (34/330)	7.1 (64/903)	1.45	0.97
		≥ 100	60.9 (14/23)	4.2 (14/330)	1.0 (9/903)	4.24	0.97
Multilocular-solid	Smooth	< 100	11.5 (7/61)	2.1 (7/330)	6.0 (54/903)	0.35	1.04
		≥ 100	33.3 (7/21)	2.1 (7/330)	1.6 (14/903)	1.37	0.99
	Irregular	< 100	34.8 (48/138)	14.6 (48/330)	10.0 (90/903)	1.46	0.95
		≥ 100	74.8 (77/103)	23.3 (77/330)	2.9 (26/903)	8.10	0.79
Solid	Smooth	< 100	31.9 (22/69)	6.7 (22/330)	5.2 (47/903)	1.28	0.98
		≥ 100	57.1 (8/14)	2.4 (8/330)	0.7 (6/903)	3.67	0.98
	Irregular	< 100	92.5 (62/67)	18.8 (62/330)	0.6 (5/903)	34.2	0.82
		≥ 100	95.0 (19/20)	5.8 (19/330)	0.1 (1/903)	52.4	0.94

*Modified²¹ after Granberg *et al.*¹. LR+, positive likelihood ratio; LR-, negative likelihood ratio.

tumors, where they had a sensitivity of 95% (106/112), a specificity of 91% (249/274), LR+ of 10.37 and LR- of 0.06.

The histology of the classifiable and unclassifiable masses and that of the correctly and incorrectly classified masses are shown in Table 1. The results are shown separately for the tumors used to develop the rules and

for the tumors in the prospective test set. Histology was similar in the two groups.

DISCUSSION

This is the largest study to date analyzing the ultrasound features of benign and malignant adnexal masses. The

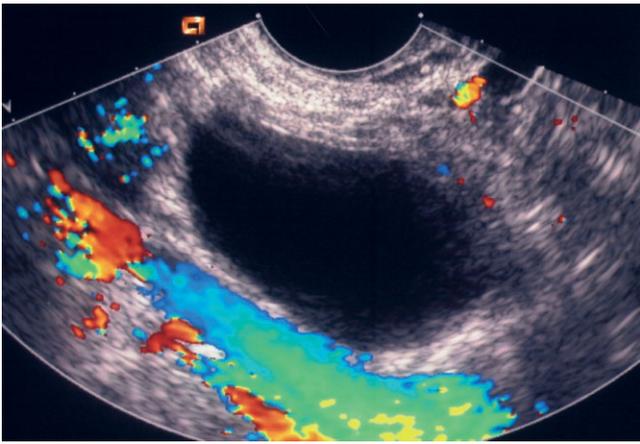


Figure 1 Color Doppler ultrasound image of a unilocular cyst that proved to be a borderline malignant mucinous cystadenoma of the endocervical type, Stage Ia. At macroscopic examination no papillary structures were visible but microscopic examination showed pseudostratification and atypia.

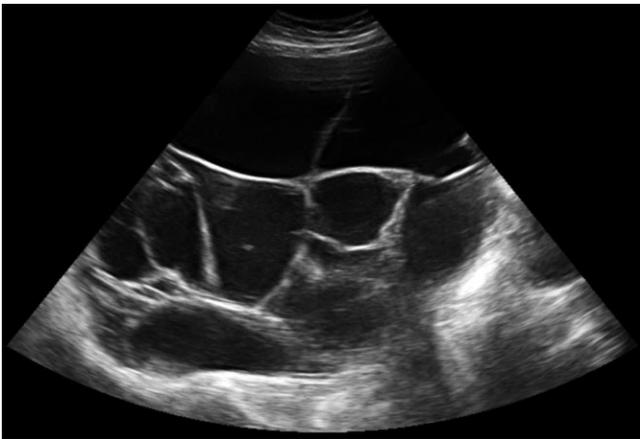


Figure 2 Ultrasound image of a multilocular cyst in a 52-year-old patient with a mucinous cystadenoma. 10% of multilocular cysts proved to be malignant.

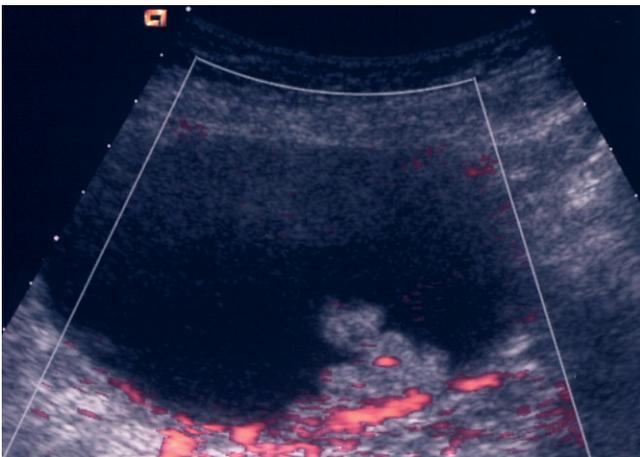


Figure 3 Color Doppler ultrasound image of a unilocular solid cyst in a 59-year-old patient with a borderline malignant serous papillary cystadenoma. Malignancy was found at surgery in 37% of unilocular solid masses.



Figure 4 Ultrasound image of a multilocular solid cyst in a 60-year-old patient with a clear cell carcinoma. Malignancy was found in 43% of multilocular solid masses.

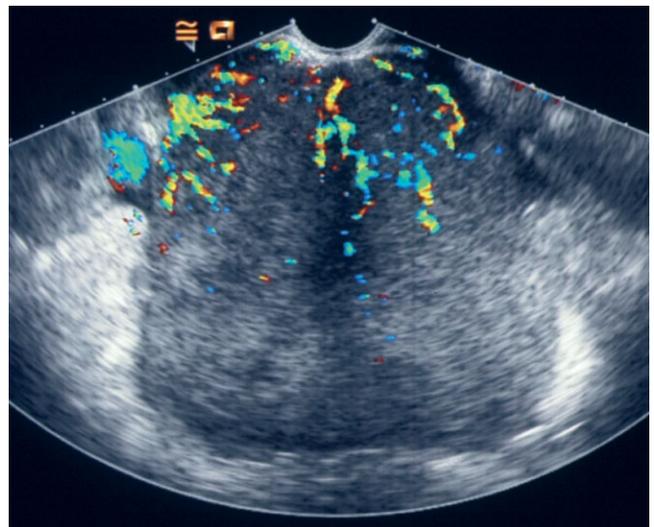


Figure 5 Color Doppler-flow ultrasound image of a solid mass in a 28-year-old patient with a dysgerminoma. 65% of solid tumors proved to be malignant.

subjective impression of the ultrasound morphology of an adnexal mass can be used to accurately determine its nature^{16–19}. Univariate analysis of the ultrasound data in this study showed that almost all ultrasound variables differed significantly between benign and malignant lesions, but that no single ultrasound finding could reliably discriminate between benign and malignant tumors. For example, the optimal threshold for maximum lesion diameter (above which a mass was more likely to be malignant) was 100 mm. However, the likelihood ratios show that in isolation this test had virtually no discriminatory power (LR+ 2.90, LR– 0.67)²⁰. The cut-off to predict malignancy for height of the largest papillary projection was 10 mm, but again this ultrasound feature did not discriminate well between benign and malignant tumors either in the total material (LR+ 5.22, LR– 0.73)

Table 9 Ten simple rules for identifying a benign or malignant tumor

Rules for predicting a malignant tumor (M-rules)		Rules for predicting a benign tumor (B-rules)	
M1	Irregular solid tumor	<input type="checkbox"/>	B1 Unilocular <input type="checkbox"/>
M2	Presence of ascites	<input type="checkbox"/>	B2 Presence of solid components where the largest solid component has a largest diameter < 7 mm <input type="checkbox"/>
M3	At least four papillary structures	<input type="checkbox"/>	B3 Presence of acoustic shadows <input type="checkbox"/>
M4	Irregular multilocular solid tumor with largest diameter \geq 100 mm	<input type="checkbox"/>	B4 Smooth multilocular tumor with largest diameter < 100 mm <input type="checkbox"/>
M5	Very strong blood flow (color score 4)	<input type="checkbox"/>	B5 No blood flow (color score 1) <input type="checkbox"/>

If one or more M-rules apply in the absence of a B-rule, the mass is classified as malignant. If one or more B-rules apply in the absence of an M-rule, the mass is classified as benign. If both M-rules and B-rules apply, the mass cannot be classified. If no rule applies, the mass cannot be classified.

or among tumors with papillary projections (LR+ 2.21, LR– 0.39).

Previous studies have demonstrated a low risk of malignancy in unilocular ovarian cysts^{21–23}. The results of this study support the idea that there is a very low risk of malignancy associated with unilocular adnexal cysts, but they also show that this is true irrespective of size, wall regularity and vascularization. Any morphological appearance other than that of a unilocular cyst is associated with a variably increased risk of malignancy.

Experienced ultrasound examiners take demographic, clinical and ultrasound information into account when they estimate the risk of malignancy in an adnexal mass, and they subconsciously apply a set of rules – based on their previous observations – when evaluating a mass. This skill is not easily transferable to less experienced ultrasound examiners. A simple form using tick boxes that might be easily used in clinical practice to help less experienced operators is shown in Table 9. A significant limitation of this study in clinical practice is the high percentage of tumors in which these rules cannot be applied (almost 25% of the tumors, both in the development and test groups), because not all masses will demonstrate the features clearly predictive of benignity or malignancy. The rules worked rather well for endometriomas, dermoid cysts, simple cysts and advanced invasive malignancies, but they worked less well for hydrosalpinx, peritoneal cysts, abscesses, fibromas, rare benign tumors, Stage I borderline tumors and Stage I primary invasive malignancies. This means that the rules worked well in tumors that are usually easily classifiable using pattern recognition but less well in tumors that tend to be more difficult to classify using pattern recognition, with the exception that hydrosalpinx is relatively easy to classify using pattern recognition^{17,19}, while the rules did not work well for hydrosalpinx. For masses where the simple rules cannot be applied, a less experienced examiner might find referral to an expert operator helpful.

The diagnostic performance of the rules was as good in the series of tumors where they were tested prospectively as in the series where they were created. However, they were tested in centers that had taken part in the first phase of the IOTA study. Their performance when used in new centers as well as the proportion of tumors that cannot be

classified using simple rules – and in particular when used by less experienced ultrasound examiners – remain to be seen.

ACKNOWLEDGMENTS

This research was supported by GOA-AMBioRICS, CoE EF/05/006, Belgian network DYSCO, BIOPATTERN (FP6-2002-IST 508803), ETUMOUR (FP6-2002-LIFE-SCIHEALTH 503094); by 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); funds administered by Malmö University Hospital; Allmänna Sjukhusets i Malmö Stiftelse för bekämpande av cancer (the Malmö General Hospital Foundation for fighting against cancer); ALF-medel and Landstingsfinansierad regional forskning (i.e. two Swedish governmental grants).

APPENDIX

IOTA Steering Committee

Dirk Timmerman, Lil Valentin, Tom Bourne, Antonia C. Testa, Sabine Van Huffel, Ignace Vergote

IOTA principal investigators (alphabetical order)

Jean-Pierre Bernard, Maurepas, France
Tom Bourne, London, UK
Enrico Ferrazzi, Milan, Italy
Davor Jurkovic, London, UK
Fabrice Lécuru, Paris, France
Andrea Lissoni, Monza, Italy
Ulrike Metzger, Paris, France
Dario Paladini, Naples, Italy
Antonia Testa, Rome, Italy
Dirk Timmerman, Leuven, Belgium
Lil Valentin, Malmö, Sweden
Caroline Van Holsbeke, Leuven, Belgium
Sabine Van Huffel, Leuven, Belgium
Ignace Vergote, Leuven, Belgium
Gerardo Zanetta (deceased), Monza, Italy

REFERENCES

1. Granberg S, Wikland M, Jansson I. Macroscopic characterization of ovarian tumors and the relation to the histological

- diagnosis: Criteria to be used for ultrasound evaluation. *Gynecol Oncol* 1989; 35: 139–144.
2. Sassone AM, Timor-Tritsch IE, Artner A, Westhoff C, Warren WB. Transvaginal sonographic characterization of ovarian disease: evaluation of a new scoring system to predict ovarian malignancy. *Obstet Gynecol* 1991; 78: 70–76.
 3. Bourne TH, Campbell S, Reynolds KM, Whitehead MI, Hampson J, Royston P, Crayford TJ, Collins WP. Screening for early familial ovarian cancer with transvaginal ultrasonography and colour blood flow imaging. *BMJ* 1993; 306: 1025–1029.
 4. Lerner JP, Timor-Tritsch IE, Federman A, Abramovich G. Transvaginal ultrasonographic characterization of ovarian masses with an improved, weighted scoring system. *Am J Obstet Gynecol* 1994; 170: 81–85.
 5. Ferrazzi E, Zanetta G, Dordoni D, Berlanda N, Mezzopane R, Lissoni AA. Transvaginal ultrasonographic characterization of ovarian masses: comparison of five scoring systems in a multicenter study. *Ultrasound Obstet Gynecol* 1997; 10: 192–197.
 6. Bourne TH, Campbell S, Steer C, Whitehead MI, Collins WP. Transvaginal colour flow imaging: a possible new screening technique for ovarian cancer. *BMJ* 1989; 299: 1367–1370.
 7. Tekay A, Joupilla P. Validity of pulsatility and resistance indices in classification of adnexal tumors with transvaginal color Doppler ultrasound. *Ultrasound Obstet Gynecol* 1992; 2: 338–344.
 8. Valentin L, Sladkevicius P, Marsál K. Limited contribution of Doppler velocimetry to the differential diagnosis of extrauterine pelvic tumors. *Obstet Gynecol* 1994; 83: 425–433.
 9. Valentin L. Gray scale sonography, subjective evaluation of the color Doppler image and measurement of blood flow velocity for distinguishing benign and malignant tumors of suspected adnexal origin. *Eur J Obstet Gynaecol* 1997; 72: 63–72.
 10. Timmerman D, Bourne TH, Taylor A, Collins WP, Verrelst H, Vandenberghe K, Vergote I. A comparison of methods for the pre-operative discrimination between benign and malignant adnexal masses: the development of a new logistic regression model. *Am J Obstet Gynecol* 1999; 181: 57–65.
 11. 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 ultrasonographic features of adnexal tumors: a consensus opinion from the international ovarian tumor analysis (IOTA) group. *Ultrasound Obstet Gynecol* 2000; 16: 500–505.
 12. Timmerman D, Testa AC, Bourne TH, 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.
 13. Heintz AP, Odcino F, Maisonneuve P, Beller U, Benedet JL, Creasman WT, Ngan HY, Pecorelli S. Carcinoma of the Ovary, in 25th Annual Report on the Results of Treatment in Gynecological Cancer. *Int J Gynecol Obstet* 2003; 83: (Suppl 1): S135–S137.
 14. Mazumdar M, Glassman Jr. Tutorial in biostatistics. Categorizing a prognostic variable: review of methods, code for easy implementation and applications to decision-making about cancer treatments. *Stat Med* 2000; 19: 113–132.
 15. Hanley JA, McNeil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology* 1982; 143: 29–36.
 16. Timmerman D, Schwärzler P, Collins WP, Claerhout F, Coenen M, Amant F, Vergote I, Bourne TH. Subjective assessment of adnexal masses using ultrasonography: an analysis of interobserver variability and experience. *Ultrasound Obstet Gynecol* 1999; 13: 11–16.
 17. Valentin L. Pattern recognition of pelvic masses by gray-scale ultrasound imaging: the contribution of Doppler ultrasound. *Ultrasound Obstet Gynecol* 1999; 14: 338–347.
 18. Jermy K, Luise C, Bourne TH. The characterization of common ovarian cysts in premenopausal women. *Ultrasound Obstet Gynecol* 2001; 17: 140–144.
 19. Valentin L, Ameye L, Jurkovic D, Metzger U, Lécuru F, Sabine 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.
 20. Jaeschke R, Guyatt GH, Sackett DL. 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. Timmerman D. Lack of standardization in gynecological ultrasonography (Editorial). *Ultrasound Obstet Gynecol* 2000; 16: 395–398.
 22. Roman LD. Small cystic pelvic masses in older women: is surgical removal necessary? *Gynecol Oncol* 1998; 69: 1–2.
 23. Ekerhovd E, Wienerroith H, Staudach A, Granberg S. Preoperative assessment of unilocular adnexal cysts by transvaginal sonography: A comparison between sonographic morphological imaging and histopathologic diagnosis. *Am J Obstet Gynecol* 2001; 184: 48–54.