Critical and conscious approach to IOTA models

Prof. Juan Luis Alcázar

Department of Obstetrics and Gynecology

Clínica Universidad de Navarra

Pamplona, Spain

IOTA Models

- Logistic models
  - LR1
  - LR2
  - ADNEX model
  - Simple rules risk assessment
- Classification systems
  - Simple rules
  - Easy descriptors
  - O-RADS

- Critical approach
  - How they were developed?
  - What are they designed for?
  - Performance?
  - Reproducible?

Ultrasound Obstet Gynecol 2000; 16: 500–505.

Terms, definitions and measurements to describe the sonographic features of adnexal tumors: a consensus opinion from the International Ovarian Tumor Analysis (IOTA) group

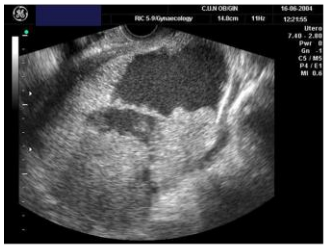
D. TIMMERMAN, L. VALENTIN*, T. H. BOURNE†, W. P. COLLINS‡, H. VERRELST§ and I. VERGOTE

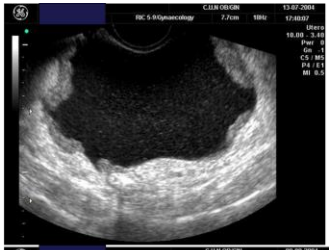
*Department of Obstetrics and Gynaecology, University Hospitals KU Leuven, Leuven, Belgium, *Department of Obstetrics and Gynaecology, University Hospital, Malmö, Sweden, †Department of Obstetrics and Gynaecology, St. George's Hospital Medical School, University of London, London, UK, ‡King's College, University of London, UK and §Department of Electrical Engineering, ESAT-SISTA, Katholieke Universiteit Leuven, Belgium*

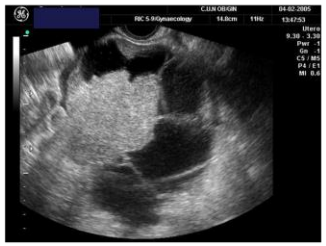
- Tumor size
- Cyst Wall thickness
  - Thin < 3 mm
  - Thick ≥ 3 mm
- Cyst Wall internal contour
  - Smooth
  - Irregular
- Papillary projections
  - Solid excrecences ≥ 3mm
  - Contour
    - Smooth
    - Irregular
- Septations
  - Thin < 3mm
  - Thick ≥ 3mm
  - Complete / incomplete
  - Number of locules
- Echogenicity
  - Anechoic
  - Low level
  - Hemorrhagic
  - Ground glass
  - Mixed
  - Solid
- External contour
  - Smooth
  - Irregular

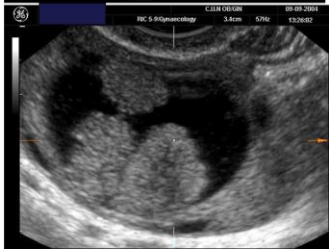
- Classification of masses
  - Unilocular
  - Unilocular solid
  - Multilocular
  - Multilocular solid
  - Solid
  - Non classifiable

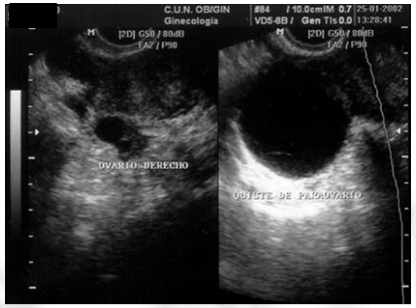
Pattern recognition

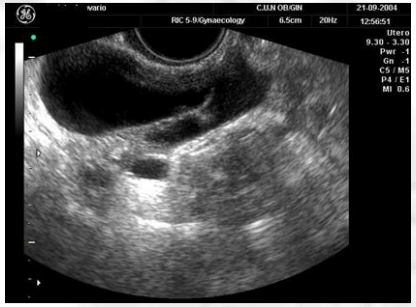


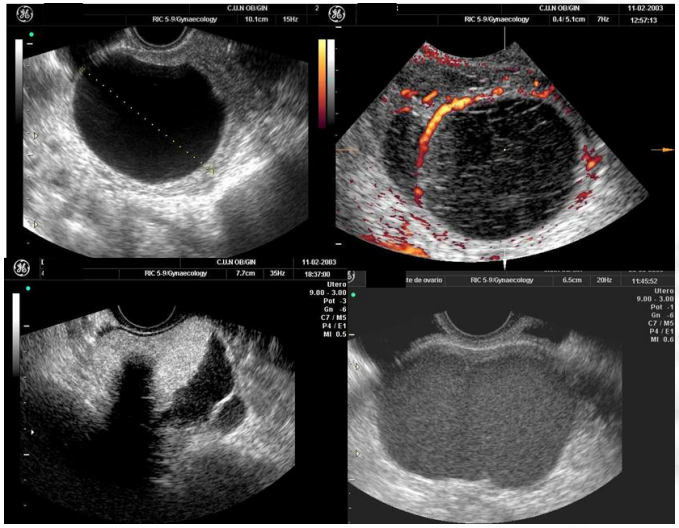


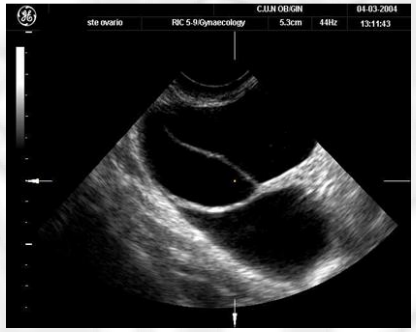












VOLUME 23 · NUMBER 34 · DECEMBER 1 2005
JOURNAL OF CLINICAL ONCOLOGY ORIGINAL REPORT

Logistic Regression Model to Distinguish Between the Benign and Malignant Adnexal Mass Before Surgery: A Multicenter Study by the International Ovarian Tumor Analysis Group

Dirk Timmerman, Antonia C. Testa, Tom Bourne, Enrico Ferrazzi, Lieveke Ameye, Maja L. Konstantinovic, Ben Van Calster, William P. Collins, Ignace Vergote, Sabine Van Huffel, and Lil Valentin

Table 7. Comparison of the Performance of Different Models in the Test Data Set With CA-125 Results Available (n = 236)

	Area Under ROC Curve	SE	Cutoff	Sensitivity (%)	Specificity (%)
M1*	0.936	0.020	0.10	92.7	74.3
M2*	0.916	0.021	0.10	89.9	70.7
RMI*	0.870	0.028	100	78.3	79.6
Tailor et al[8]†	0.869	0.025	0.25	63.2	88.2
Timmerman et al[9]*	0.903	0.023	0.25	79.7	80.8

Abbreviations: CA-125, serum level of the tumor marker; ROC, receiver operating characteristic.
*Applied on the cases in the test set with CA-125 (236 cases).
†Applied only on the cases in the test set with time-averaged maximum velocity (220 cases).

- LR1: 12 variables
  - Past history ovarian cancer
  - HRT
  - Age
  - Maximum diameter of the lesion
  - Ascites
  - Doppler flow in papillary projections
  - Solid echogenicity
  - Maximum diameter of solid component
  - Irregular internal wall
  - Acoustic shadowing
  - Color score
  - Pain during US examination

- LR2: 6 variables
  - Age
  - Ascites
  - Doppler flow in papillary projections
  - Maximum diameter of solid component
  - Irregular internal wall
  - Acoustic shadowing

N = 1066 cases

Ultrasound Obstet Gynecol 2010; 36: 226–234
Published online 19 July 2010 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/uog.7636

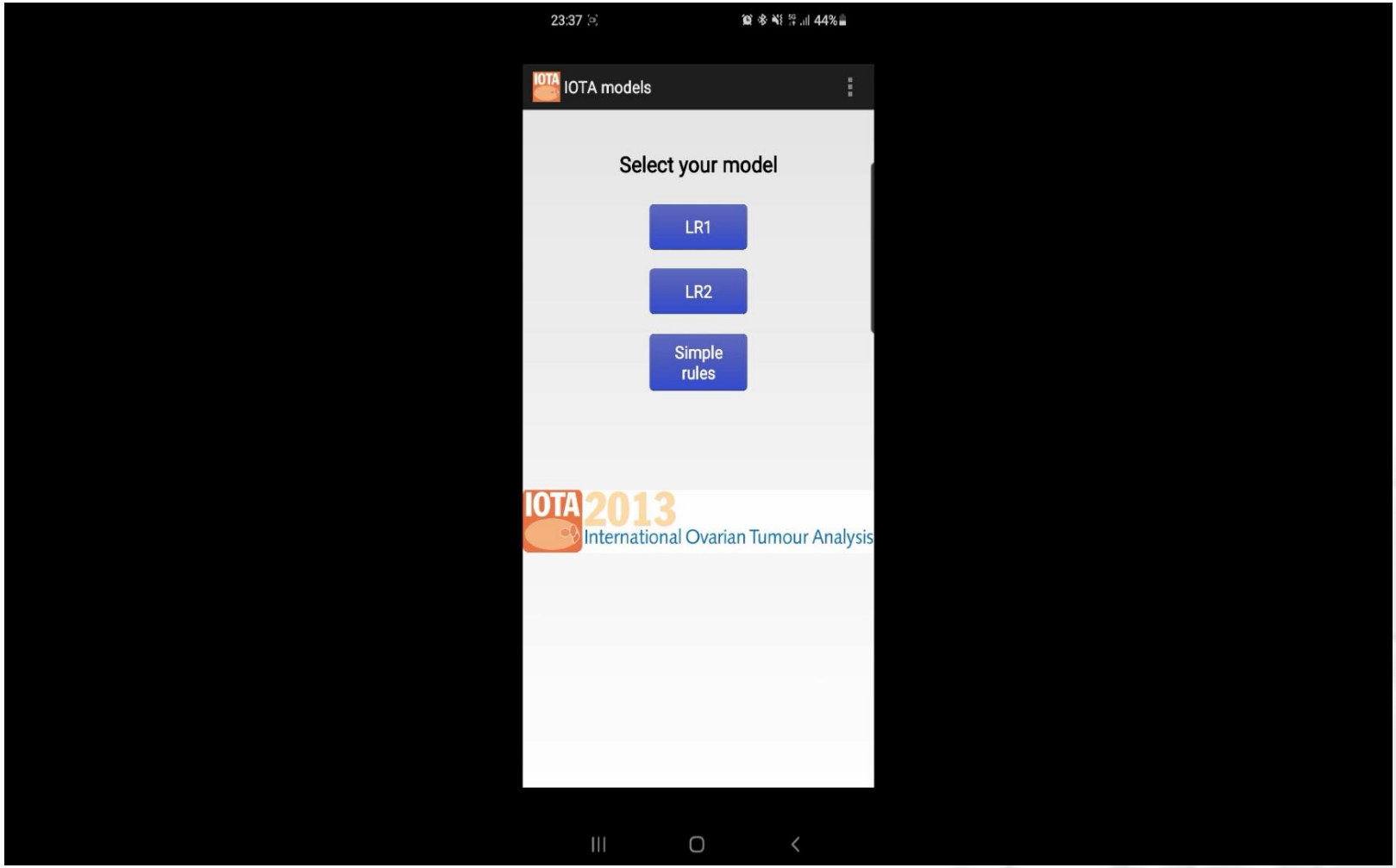
Ovarian cancer prediction in adnexal masses using ultrasound-based logistic regression models: a temporal and external validation study by the IOTA group

D. TIMMERMAN*, B. VAN CALSTER†, A. C. TESTA‡, S. GUERRIERO§, D. FISCHEROVA¶, A. A. LISSONI**, C. VAN HOLSBEKE*††, R. FRUSCIO**, A. CZEKIERDOWSKI‡‡, D. JURKOVIC§§, L. SAVELLI¶¶, I. VERGOTE*, T. BOURNE*#, S. VAN HUFFEL† and L. VALENTIN***

Table 4 Results of prospective validation of the diagnostic performance of two logistic regression models used to calculate the risk of malignancy in adnexal masses (LR1 and LR2) and results of the subjective interpretation of ultrasound findings ('subjective assessment')

Validation	*Statistic*	*LR1*	*LR2*	*Difference LR1–LR2*	*Subjective assessment*	*Difference LR1– subjective assessment*	*Difference LR2– subjective assessment*
External	AUC (95% CI)	0.956 (0.940, 0.968)	0.949 (0.931, 0.964)	0.007 (−0.001, 0.015)	0.949 (0.930, 0.964)	0.007 (−0.010, 0.023)	0.000 (−0.019, 0.018)
($n = 997$,	LR+ (95% CI)	6.84 (5.69, 8.25)	6.36 (5.33, 7.63)	0.48 (−0.36, 1.50)	11.0 (8.60, 14.1)	−4.2 (−7.5, −2.1)	−4.6 (−7.9, −2.4)
12 centers)	LR− (95% CI)	0.09 (0.06, 0.14)	0.10 (0.06, 0.14)	−0.01 (−0.03, 0.02)	0.14 (0.10, 0.19)	−0.05 (−0.09, 0.00)	−0.04 (−0.09, 0.01)
	Sensitivity (%)	92.2	91.8	0.4 (−2.1, 2.9)	87.5	4.7 (0.6, 9.1)	4.3 (0.0, 8.8)
	Invasive tumors only (%)	93.9	94.4	−0.5 (−3.3, 2.2)	89.2	4.7 (0.5, 9.2)	5.2 (0.7, 10.0)
	Specificity (%)	86.5	85.6	0.9 (−1.0, 2.9)	92.1	−5.6 (−8.0, −3.2)	−6.5 (−9.0, −4.0)
	Predicted vs. observed risk*	0.83	0.78				
Temporal	AUC (95% CI)	0.945 (0.930, 0.958)	0.918 (0.896, 0.936)	0.027 (0.017, 0.038)	0.959 (0.944, 0.973)	−0.014 (−0.029, 0.001)	−0.041 (−0.063, −0.022)
($n = 941$,	LR+ (95% CI)	4.77 (4.08, 5.61)	4.42 (3.78, 5.19)	0.35 (−0.23, 0.84)	14.1 (10.6, 19.0)	−9.4 (−13.9, −6.3)	−9.7 (−14.6, −6.7)
7 centers)	LR− (95% CI)	0.09 (0.06, 0.14)	0.14 (0.10, 0.19)	−0.05 (−0.08, −0.01)	0.07 (0.05, 0.11)	0.02 (−0.01, 0.05)	0.07 (0.02, 0.11)
	Sensitivity (%)	92.7	89.2	3.5 (0.7, 6.6)	93.0	−0.3 (−3.5, 2.7)	−3.8 (−7.8, −0.1)
	Invasive tumors only (%)	96.8	95.4	1.4 (−1.4, 4.5)	96.3	0.5 (−2.2, 3.2)	−0.9 (−4.4, 2.4)
	Specificity (%)	80.6	79.8	0.8 (−1.2, 2.8)	93.4	−12.8 (−15.8, −10.1)	−13.6 (−16.7, −10.7)
	Predicted vs. observed risk*	0.84	0.81				

95% CIs for area under the receiver–operating characteristics curve (AUC), differences in AUC and differences in LR+ and LR− were based on the bias-corrected bootstrap method using 1000 bootstrap samples; 95% CIs for LR+ and LR− were based on the Cox–Hinkley–Miettinen–Nurminen method[18]; 95% CIs for sensitivity and specificity differences were obtained using a method based on continuity-corrected Wilson score intervals (method ten from Justice *et al.*[17]). *Ratio between average predicted probability of malignancy and observed prevalence of malignancy.


23:37
44%
IOTA models
Select your model
LR1
LR2
Simple rules
IOTA 2013
International Ovarian Tumour Analysis

Imaging, Diagnosis, Prognosis

Clinical Cancer Research

External Validation of Diagnostic Models to Estimate the Risk of Malignancy in Adnexal Masses

Caroline Van Holsbeke[1,4], Ben Van Calster[1,2], Tom Bourne[1,5], Silvia Ajossa[6], Antonia C. Testa[7], Stefano Guerriero[6], Robert Fruscio[8], Andrea Alberto Lissoni[8], Artur Czekierdowski[10], Luca Savelli[9], Sabine Van Huffel[2,3], Lil Valentin[11], and Dirk Timmerman[1]

Table 5. Diagnostic performance of the different models when the cutoff recommended in the original article is used

Model	Cutoff	Sensitivity, %	Specificity, %	LR^+	LR^-	DOR (LR^+/LR^-; 95% CI)	Sensitivity for primary invasive stage I tumors,[a] %
RMI variants							
RMI Jacobs	200	67	95	12.7	0.34	37 (24–58)	51
RMI2 Tingulstad 1996	200	72	90	7.3	0.31	23 (16–34)	58
RMI3 Tingulstad 1999	200	67	93	9.8	0.35	28 (18–43)	51
RMI4 Yamamoto	450	67	94	11.8	0.35	34 (22–53)	52
Non-IOTA logistic regression models							
LR Timmerman a	0.25	75	91	8.2	0.28	29 (20–43)	63
LR Lu	0.20	83	85	5.4	0.20	27 (18–39)	73
LR Jokubkiene	0.12	77	88	6.6	0.26	26 (18–37)	81
LR Timmerman b	0.60	79	84	4.9	0.26	19 (13–27)	65
LR Minaretzis							
LR Tailor	0.50	20	98	9.7	0.82	12 (6.5–21)	19
Non-IOTA ANNs							
ANN Timmerman 1	0.45	75	89	6.9	0.28	25 (17–37)	65
ANN Timmerman 2	0.60	97	41	1.6	0.08	20 (10–41)	94
IOTA models							
LR1	0.10	92	87	6.8	0.09	75 (46–125)	92
LS-SVM rbf	0.12	89	90	8.8	0.12	75 (47–120)	90
BPER11	0.15	91	89	8.2	0.11	77 (48–125)	88
BMLP11-2b	0.15	91	87	6.8	0.10	65 (40–104)	90
RVM rbf	0.15	91	88	7.4	0.11	69 (43–111)	92
LS-SVM lin	0.15	87	91	9.9	0.14	70 (45–109)	85
LR2	0.10	92	86	6.4	0.10	66 (40–108)	92
RVM lin	0.20	91	89	8.0	0.11	75 (47–122)	90
BMLP11-2a	0.15	86	88	7.1	0.16	46 (30–69)	88
RVM add rbf	0.15	89	89	7.8	0.13	61 (39–96)	90
LS-SVM add rbf	0.12	87	89	7.7	0.15	53 (34–81)	92

Abbreviations: add rbf, additive radial basis function kernel; lin, linear kernel; LR^+, positive likelihood ratio; LR^-, negative likelihood ratio; rbf, radial basis function.

[a]Primary invasive stage I tumors includes rare primary invasive stage I tumors.

Interobserver Agreement in Describing the Ultrasound Appearance of Adnexal Masses and in Calculating the Risk of Malignancy Using Logistic Regression Models

Povilas Sladkevicius and Lil Valentin

Clin Cancer Res; 21(3) February 1, 2015

Table 5. Interobserver reproducibility of the risk of malignancy calculated using the risk calculation models LR1 and LR2

	Calculated risk of malignancy (both sonologists)		Difference between the risk calculated by sonologists 1 and 2			ICC
Parameter	Median	Range	Mean	95% CI	Limits of agreement	Point estimate (95% CI)
Risk of malignancy calculated using LR1	7.85 ($n = 234$)	0.10–99.10	−0.53 ($n = 117$)	−3.07–2.01	−28.05–26.99	0.911 (0.874–0.937)
Risk of malignancy calculated using LR2	6.65 ($n = 234$)	0.10–98.40	0.02 ($n = 117$)	−3.06–3.10	−33.22–33.26	0.832 (0.766–0.880)

Ultrasound Obstet Gynecol 2008; **31**: 681–690
Published online in Wiley InterScience (www.interscience.wiley.com). **DOI:** 10.1002/uog.5365

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*

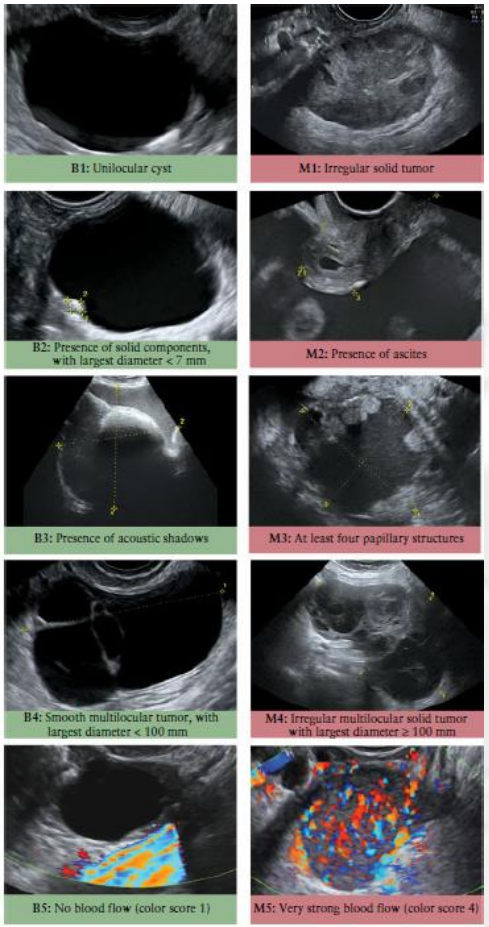
Table 9 Ten simple rules for identifying a benign or malignant tumor

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

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.

Highest PPV for malignancy

Highest PPV for benignity



At least one M without B → Malignant
At least one B without M → Benign
No M and B or at least one B and one M → inconclusive

Ultrasound Obstet Gynecol 2014; **44**: 95–99
Published online 21 May 2014 in Wiley Online Library (wileyonlinelibrary.com). **DOI:** 10.1002/uog.13254

Interobserver agreement in describing adnexal masses using the International Ovarian Tumor Analysis simple rules in a real-time setting and using three-dimensional ultrasound volumes and digital clips

B. RUIZ DE GAUNA*, P. SANCHEZ†, L. PINEDA‡, J. UTRILLA-LAYNA‡, L. JUEZ‡ and J. L. ALCÁZAR‡

**Barcelona Center for Maternal Fetal and Neonatal Medicine, Hospital Sant Joan de Deu and Hospital Clinic, Universitat de Barcelona, Barcelona, Spain; †Department of Obstetrics and Gynecology, Hospital Fundación Jimenez Diaz, Madrid, Spain; ‡Department of Obstetrics and Gynecology, Clinica Universidad de Navarra, Pamplona, Spain*

Reproducible

Ultrasound Obstet Gynecol 2014; **44**: 100–108
Published online 1 June 2014 in Wiley Online Library (wileyonlinelibrary.com). **DOI:** 10.1002/uog.13273

Intra- and interobserver agreement with regard to describing adnexal masses using International Ovarian Tumor Analysis terminology: reproducibility study involving seven observers

L. ZANNONI*†, L. SAVELLI*, L. JOKUBKIENE†, A. DI LEGGE‡, G. CONDOUS§, A. C. TESTA‡, P. SLADKEVICIUS† and L. VALENTIN†

**Department of Obstetrics and Gynecology, S. Orsola Malpighi Hospital, University of Bologna, Bologna, Italy; †Department of Obstetrics and Gynecology, Skåne University Hospital Malmö, Lund University, Malmö, Sweden; ‡Department of Obstetrics and Gynecology, Catholic University of the Sacred Heart, Rome, Italy; §Acute Gynaecology, Early Pregnancy and Advanced Endosurgery Unit, Sydney Medical School Nepean, University of Sydney, Sydney, Australia*

Personalized Medicine and Imaging

Clinical Cancer Research

Interobserver Agreement in Describing the Ultrasound Appearance of Adnexal Masses and in Calculating the Risk of Malignancy Using Logistic Regression Models

Clin Cancer Res; 21(3) February 1, 2015

Povilas Sladkevicius and Lil Valentin

PROGRESOS DE Obstetricia y Ginecología
Revista Oficial de la Sociedad Española de Ginecología y Obstetricia
Prog Obstet Ginecol 2018;61(4):390-400

Revisión Sistemática

IOTA Simple Rules for the differential diagnosis of ovarian adnexal masses: Systematic review and meta-analysis

Reglas Simples ("Simple Rules") de IOTA en el diagnóstico diferencial de las masas anexiales de ovario: revisión sistemática y meta-análisis

Andrea Ilundain López de Munain[1], Ana Salas[2], Enrique Chacón[2], Nabil Mazour[2] and Juan Luis Alcázar[1,2]

[1]Facultad de Medicina. Universidad de Navarra, Pamplona. [2]Departamento de Ginecología y Obstetricia. Clínica Universidad de Navarra, Pamplona

Inconclusive cases: 15-25%

Table II.
Diagnostic yield of the methods analyzed

	Diagnostic methods	
	SR+IM	SR+SA
Sensitivity	0.94 (95%CI, 0.92-0.96)	0.89 (95%CI, 0.86-0.91)
Specificity	0.78 (95%CI, 0.74-0.82)	0.90 (95%CI, 0.86-0.93)
LR+	4.3 (95%CI, 3.7-5.1)	9.0 (95%CI, 6.4-12.7)
LR–	0.07 (95%CI, 0.06-0.10)	0.12 (95%CI, 0.10-0.15)

SR+IM: Simple rules + inconclusive results considered malignant. SR+SA: Simple rules + inconclusive results evaluated based on the subjective assessment of an expert examiner.

Human Reproduction Update, Vol.20, No.3 pp. 449–462, 2014
Advanced Access publication on December 9, 2013 doi:10.1093/humupd/dmt059

human reproduction update

Presurgical diagnosis of adnexal tumours using mathematical models and scoring systems: a systematic review and meta-analysis

Jeroen Kaijser[1,2†], Ahmad Sayasneh[3†], Kirsten Van Hoorde[4], Sadaf Ghaem-Maghami[3], Tom Bourne[1,2,3], Dirk Timmerman[1,2,*] and Ben Van Calster[1]

Table II Pooled summary estimates of the expected operating point (sensitivity and specificity) and corresponding 95% confidence interval.

Model	Cut-off	Studies (*n*)	Centres[a] (*n*)	Sensitivity	Specificity
Morphologic scoring systems					
Sassone	>9	19	19	0.85 [0.77;0.90]	0.80 [0.73;0.86]
Lerner	>3	9	17	0.80 [0.70;0.86]	0.61 [0.53;0.68]
Depriest	>5	8	8	0.90 [0.81;0.95]	0.68 [0.57;0.77]
Ferrazzi	>9	7	7	0.86 [0.77;0.91]	0.80 [0.66;0.89]
Ultrasound rules					
Simple Rules	*n/a*[b]	5	17	0.93 [0.89;0.95]	0.81 [0.76;0.85]
Risk of Malignancy Indexes(RMI)					
RMI I	200	23	41	0.72 [0.67;0.76]	0.92 [0.89;0.93]
RMI II	200	15	32	0.75 [0.69;0.80]	0.87 [0.84;0.90]
RMI III	200	9	19	0.70 [0.60;0.78]	0.91 [0.88;0.93]
RMI IV	450	3	13	0.68 [0.59;0.76]	0.94 [0.91;0.96]
Logistic regression models					
Tailor	50%	6	24	0.35 [0.24;0.49]	0.96 [0.94;0.98]
LRa	25%	3	20	0.76 [0.70;0.81]	0.87 [0.82;0.90]
LRb	60%	4	21	0.82 [0.77;0.86]	0.78 [0.73;0.83]
Prömpeler	10%	2	10	0.61 [0.46;0.74]	0.81 [0.70;0.89]
Jokubkiene	12%	2	20	0.77 [0.71;0.82]	0.87 [0.83;0.89]
IOTA LR2	10%	3	13	0.92 [0.88;0.95]	0.83 [0.77;0.88]
Artificial neural networks					
ANN1	45%	3	20	0.77 [0.71;0.82]	0.86 [0.80;0.90]
ANN2	60%	4	21	0.97 [0.95;0.98]	0.37 [0.31;0.44]

[a]For the centre-specific results, only centres of multicentre studies that contributed at least 3 benign and 3 malignant cases are considered.
[b]For simple rules, the inconclusive cases are considered as malignant.

European Journal of Cancer 58 (2016) 17–29

Available online at www.sciencedirect.com

ScienceDirect

journal homepage: www.ejcancer.com

Review

Subjective assessment versus ultrasound models to diagnose ovarian cancer: A systematic review and meta-analysis

E.M.J. Meys [a], J. Kaijser [b,1], R.F.P.M. Kruitwagen [a], B.F.M. Slangen [a], B. Van Calster [c], B. Aertgeerts [d], J.Y. Verbakel [d,e], D. Timmerman [b], T. Van Gorp [a,*]

Table 2
Pooled summary point estimates for the subgroups of pre- and postmenopausal patients for all methods included in this review at their original cut-off point together with their 95% confidence interval (95%CI).

	Premenopausal					Postmenopausal				
	Sens. (95%CI)	Spec. (95%CI)	DOR (95%CI)	LR+ (95%CI)	LR− (95%CI)	Sens. (95%CI)	Spec. (95%CI)	DOR (95%CI)	LR+ (95%CI)	LR− (95%CI)
RMI-I	0.63 (0.51–0.73)	0.93 (0.89–0.95)	21 (16–28)	8.51 (6.38–11.37)	0.40 (0.30–0.54)	0.79 (0.77–0.82)	0.86 (0.79–0.91)	23 (15–36)	5.60 (3.80–8.30)	0.24 (0.21–0.27)
RMI-II	0.58 (0.46–0.68)	0.91 (0.87–0.94)	13 (8–22)	6.3 (4.4–9.0)	0.47 (0.36–0.60)	0.84 (0.80–0.87)	0.80 (0.76–0.83)	20 (14–27)	4.10 (3.40–4.90)	0.21 (0.17–0.25)
RMI-III	0.57 (0.45–0.68)	0.90 (0.85–0.93)	11 (7–18)	5.44 (3.95–7.49)	0.48 (0.37–0.62)	0.79 (0.76–0.82)	0.89 (0.87–0.92)	32 (23–45)	7.50 (5.80–9.60)	0.23 (0.20–0.27)
SA	0.90 (0.88–0.91)	0.94 (0.93–0.95)	151 (123–185)	15.89 (12.86–19.62)	0.11 (0.09–0.12)	0.94 (0.93–0.95)	0.85 (0.82–0.88)	96 (66–139)	6.43 (5.14–8.05)	0.07 (0.05–0.08)
SR+SA	0.89 (0.86–0.92)	0.91 (0.85–0.95)	86 (43–172)	10.27 (5.86–18.01)	0.12 (0.09–0.16)	0.92 (0.90–0.94)	0.87 (0.84–0.90)	83 (60–115)	7.37 (5.80–9.36)	0.09 (0.07–0.12)
SR+Mal	0.93 (0.90–0.95)	0.82 (0.79–0.85)	61 (40–92)	5.23 (4.39–6.24)	0.09 (0.06–0.12)	0.95 (0.92–0.96)	0.77 (0.71–0.81)	57 (40–80)	4.04 (3.28–4.97)	0.07 (0.05–0.10)
LR2	0.87 (0.81–0.91)	0.91 (0.86–0.94)	65 (42–100)	9.60 (6.20–14.70)	0.15 (0.11–0.21)	0.94 (0.92–0.96)	0.86 (0.82–0.89)	92 (55–154)	6.59 (5.23–8.29)	0.07 (0.05–0.10)

Abbreviations: CI, confidence interval; RMI, risk of malignancy index; SA, subjective assessment; SR+SA, simple rules, if inconclusive classified by subjective assessment; SR+Mal, simple rules, if inconclusive classified as malignant; sens., sensitivity; spec., specificity; DOR, diagnostic odds ratio; LR+, positive likelihood ratio; LR−, negative likelihood ratio; LR2, logistic regression model 2.

GYNECOLOGY

Predicting the risk of malignancy in adnexal masses based on the *Simple Rules* from the International Ovarian Tumor Analysis group

Dirk Timmerman, MD, PhD[1]; Ben Van Calster, MSc, PhD[1]; Antonia Testa, MD, PhD; Luca Savelli, MD, PhD; Daniela Fischerova, MD, PhD; Wouter Froyman, MD; Laure Wynants, MSc; Caroline Van Holsbeke, MD, PhD; Elisabeth Epstein, MD, PhD; Dorella Franchi, MD; Jeroen Kaijser, MD, PhD; Artur Czekierdowski, MD, PhD; Stefano Guerriero, MD, PhD; Robert Fruscio, MD, PhD; Francesco P. G. Leone, MD; Alberto Rossi, MD; Chiara Landolfo, MD; Ignace Vergote, MD, PhD; Tom Bourne, MD, PhD; Lil Valentin, MD, PhD

American Journal of Obstetrics & Gynecology APRIL 2016

TABLE 5
Model coefficients for 11 predictors obtained on development data (n = 2445)

Predictor	Coefficient	SE
Intercept	−1.10	0.26
B1 (unilocular cyst)	−3.10	0.34
B2 (solid components present, but <7 mm)	−1.55	0.59
B3 (acoustic shadows)	−1.58	0.27
B4 (smooth multilocular tumor with largest diameter <100 mm)	−3.59	0.60
B5 (no blood flow; color score 1)	−1.96	0.24
M1 (irregular solid tumor)	2.38	0.39
M2 (ascites)	2.87	0.29
M3 (at least 4 papillary structures)	1.72	0.28
M4 (irregular multilocular-solid tumor with largest diameter ≥100 mm)	1.12	0.23
M5 (very strong flow; color score 4)	1.53	0.24
Oncology center	0.95	0.31

Timmerman et al. Simple ultrasound rules to predict risk of malignancy in adnexal masses. Am J Obstet Gynecol 2016.

TABLE 10
Summary classification of Simple Rules risk calculation based on all data (n = 4848)

Features	Observed malignancy rate	Estimated individual risk of malignancy	Classification
No M-features AND >2 B-features	1/175 (0.6%)	<0.01–0.29%	Very low risk
- No M-features AND 2 B-features - No M-features AND feature B1 present	20/1560 (1.3%)	0.19–2.7% 1.2–3.1%	Low risk
No M-features AND 1 B-feature present (except B1)	60/722 (8.3%)	2.4–15.2%	Intermediate risk
- No features - Equal no. of M- and B-features - >0 M-features, but more B- than M-features	451/1096 (41.1%)	27.5–48.7% 5.6–78.1% 1.3–28.4%	Elevated risk
More M- than B-features present	1133/1295 (87.5%)	42.0–>99.9%	Very high risk

This simplified system only provides risk ranges for no. of B- and M-features present, but facilitates clinical triaging in absence of electronic devices. Personalized risk estimates can be obtained in second step.

B-feature, benign feature; *M-feature*, malignant feature.

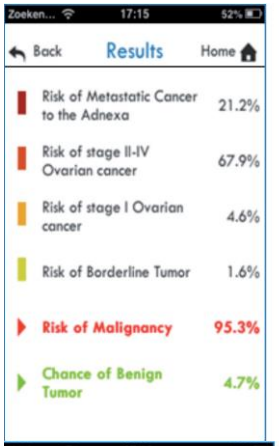
Timmerman et al. Simple ultrasound rules to predict risk of malignancy in adnexal masses. Am J Obstet Gynecol 2016.

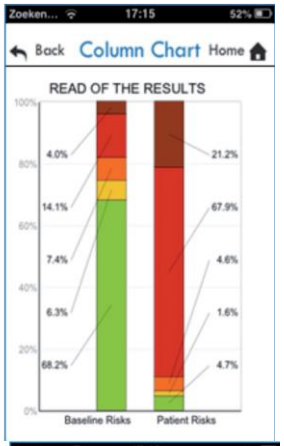
thebmj

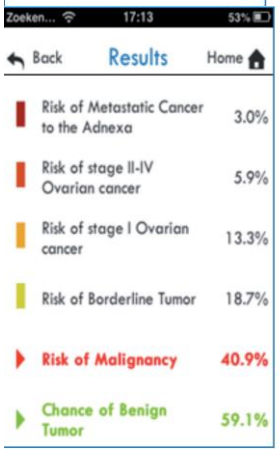
BMJ 2014;349:g5920 doi: 10.1136/bmj.g5920 (Published 16 October 2014) Page 1 of 14

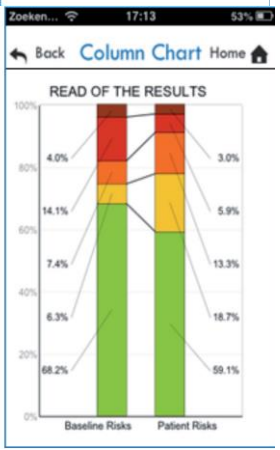
RESEARCH

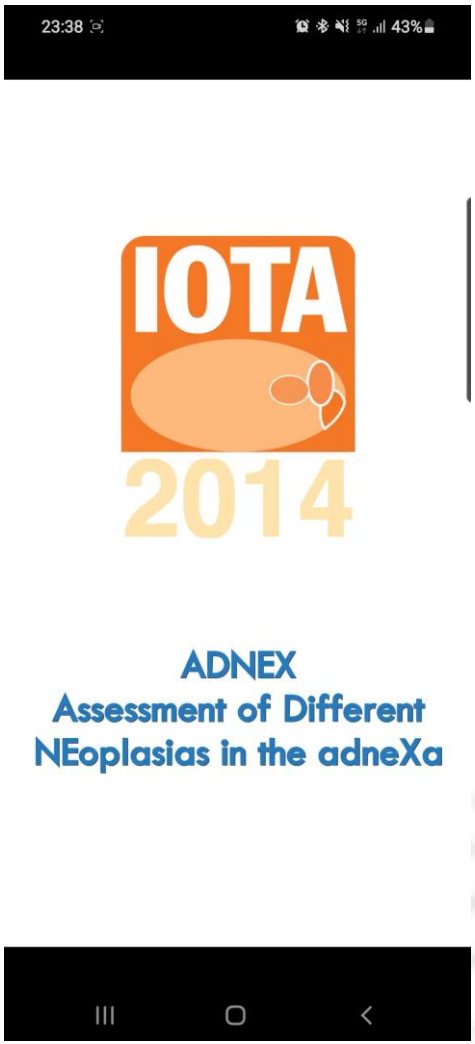
Evaluating the risk of ovarian cancer before surgery using the ADNEX model to differentiate between benign, borderline, early and advanced stage invasive, and secondary metastatic tumours: prospective multicentre diagnostic study











- Patient's age
- Type of center
- Maximum diameter of the lesión
- Number of papillary projections
- Number of locules
- Ascites
- Maximum diameter of solid component
- Acoustic shadowing
- CA-125

Table 3| **Diagnostic performance of ADNEX model when using different thresholds for total probability of malignancy (sum of probabilities of four subtypes of ovarian malignancy)**

Threshold for probability of malignancy*	Development data (n=3506)				Validation data (n=2403)				After updating on pooled data (n=5909)			
	AUC	Sensitivity	Specificity	Diagnostic odds ratio	AUC	Sensitivity	Specificity	Diagnostic odds ratio	AUC	Sensitivity	Specificity	Diagnostic odds ratio
Not applicable	0.954 (0.947 to 0.961)	—	—	—	0.943 (0.934 to 0.952)	—	—	—	0.950 (0.944 to 0.955)	—	—	—
3%	—	98.8 (97.9 to 99.4)	52.3 (50.4 to 54.3)	93.6	—	98.9 (98.0 to 99.4)	46.6 (44.0 to 49.2)	76.8	—	99.1 (98.6 to 99.5)	43.4 (41.8 to 45.0)	86.2
5%	—	97.9 (96.8 to 98.7)	65.4 (63.6 to 67.3)	87.9	—	98.4 (97.4 to 99.1)	59.4 (56.8 to 62.0)	88.1	—	98.0 (97.3 to 98.6)	61.1 (59.5 to 62.6)	78.0
10%	—	95.9 (94.4 to 97.1)	75.5 (73.8 to 77.2)	72.0	—	96.5 (95.2 to 97.6)	71.3 (68.9 to 73.7)	69.2	—	96.4 (95.4 to 97.2)	73.2 (71.8 to 74.6)	72.7
15%	—	94.4 (92.8 to 95.8)	81.0 (79.4 to 82.5)	71.9	—	94.2 (92.5 to 95.6)	77.2 (74.9 to 79.3)	54.7	—	94.5 (93.4 to 95.5)	78.7 (77.4 to 79.9)	63.4

AUC=area under receiver operating characteristic curve.

Exact binomial 95% confidence intervals are reported in parentheses.

*Probability equal to or more than threshold indicates malignancy.

RESEARCH

the bmj | BMJ 2020;370:m2614 | doi: 10.1136/bmj.m2614

Validation of models to diagnose ovarian cancer in patients managed surgically or conservatively: multicentre cohort study

Ben Van Calster[1,2,3] Lil Valentin,[4,5] Wouter Froyman,[1,6] Chiara Landolfo,[1,7] Jolien Ceusters,[8] Antonia C Testa,[9,10] Laure Wynants,[1,11] Povilas Sladkevicius,[4,5] Caroline Van Holsbeke,[12] Ekaterini Domali,[13] Robert Fruscio,[14] Elisabeth Epstein,[15,16] Dorella Franchi,[17] Marek J Kudla,[18] Valentina Chiappa,[19] Juan L Alcazar,[20] Francesco P G Leone,[21] Francesca Buonomo,[22] Maria Elisabetta Coccia,[23] Stefano Guerriero,[24] Nandita Deo,[25] Ligita Jokubkiene,[4,5] Luca Savelli,[26] Daniela Fischerová,[27] Artur Czekierdowski,[28] Jeroen Kaijser,[29] An Coosemans,[6,8,30] Giovanni Scambia,[9,10] Ignace Vergote,[6,8,30] Tom Bourne,[1,6,7] Dirk Timmerman[1,6]

IOTA 5
17 centers
4905 patients

Table 5 | Sensitivity (at 90% specificity) and specificity (at 90% sensitivity) for all prediction models

Model	Sensitivity at 90% specificity (95% CI)	Specificity at 90% sensitivity (95% CI)
RMI	70.1% (63.5 to 76.0)	69.3% (60.1 to 77.3)
LR2	82.4% (76.3 to 87.1)	81.7% (73.2 to 87.9)
SRRisk	88.5% (83.4 to 92.2)	83.8% (74.2 to 90.3)
ADNEX without CA125	85.2% (78.9 to 89.9)	85.7% (78.5 to 90.8)
ADNEX with CA125	86.5% (80.9 to 90.7)	86.6% (80.8 to 90.9)

ADNEX=assessment of different neoplasias in the adnexa; LR2=logistic regression model 2; RMI=risk of malignancy index; SRRisk=simple rules risk model.

Diagnostic Accuracy of the ADNEX Model for Ovarian Cancer at the 15% Cut-Off Value: A Systematic Review and Meta-Analysis

*Xiaotong Huang[1,2], Ziwei Wang[1,2], Meiqin Zhang[1,2] and Hong Luo[1,2]**

frontiers in Oncology

SYSTEMATIC REVIEW
published: 17 June 2021
doi: 10.3389/fonc.2021.684257

10 studies
5170 masses
Cut-off 15%
Pooled sensitivity: 92%
Pooled specificity: 82%

Review > Ultrasound Med Biol. 2022 May;48(5):730-742. doi: 10.1016/j.ultrasmedbio.2022.02.001. Epub 2022 Mar 7.

Value of Assessment of Different Neoplasias in the Adnexa in the Differential Diagnosis of Malignant Ovarian Tumor and Benign Ovarian Tumor: A Meta-analysis

Xiang Yue [1], Lili Zhong [2], Yashan Wang [3], Chenyang Zhang [3], Xiaofei Chen [3], Sor Jiayi Hu [3], Junjun Hu [3], Chunpeng Wang [4], Xin Liu [5]

22 studies
17293 masses

Table 2. The Sensitivity and Specificity of the ADNEX Model With and Without the Inclusion of Serum CA125.

Cut-off value	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)	DOR	AUC
With CA125						
3%	0.99(0.99-0.99)	0.38(0.29-0.49)	1.61(1.37-1.89)	0.02(0.01-0.03)	73.90	0.98
5%	0.97(0.95-0.98)	0.64(0.56-0.72)	2.69(2.16-3.35)	0.05(0.03-0.07)	53.82	0.95
10%	0.94(0.92-0.96)	0.78(0.72-0.83)	4.28(3.39-5.41)	0.07(0.06-0.10)	58.77	0.95
15%	0.92(0.90-0.94)	0.83(0.77-0.88)	5.53(3.97-7.70)	0.09(0.07-0.12)	60.80	0.95
20%	0.90(0.87-0.92)	0.83(0.72-0.90)	5.34(3.13-9.11)	0.12(0.09-0.16)	43.80	0.92
30%	0.86(0.81-0.90)	0.85(0.80-0.88)	5.60(4.47-7.01)	0.17(0.13-0.22)	33.36	0.92
Without CA125						
5%	0.98(0.97-0.99)	0.54(0.43-0.65)	2.15(1.69-2.74)	0.03(0.01-0.07)	76.52	0.98
10%	0.95(0.91-0.97)	0.73(0.64-0.80)	3.51(2.66-4.63)	0.07(0.04-0.11)	51.15	0.93
15%	0.95(0.92-0.97)	0.76(0.64-0.85)	3.90(2.52-6.05)	0.06(0.04-0.11)	61.89	0.96
20%	0.92(0.88-0.95)	0.81(0.73-0.87)	4.77(3.25-7.01)	0.10(0.06-0.16)	48.13	0.94

Ultrasound Obstet Gynecol 2012; **40**: 582–591
Published online in Wiley Online Library (wileyonlinelibrary.com). **DOI:** 10.1002/uog.11177

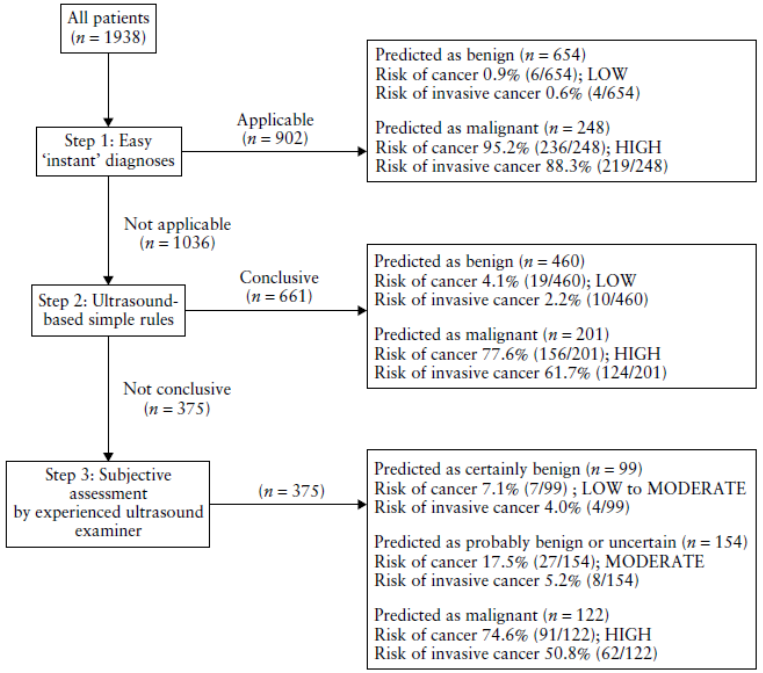
Clinically oriented three-step strategy for assessment of adnexal pathology

L. AMEYE*†, D. TIMMERMAN†, L. VALENTIN‡, D. PALADINI§, J. ZHANG¶, C. VAN HOLSBEKE†**, A. A. LISSONI††, L. SAVELLI‡‡, J. VELDMAN†, A. C. TESTA§§, F. AMANT†, S. VAN HUFFEL* and T. BOURNE†¶¶

Easy descriptors

Table 1 Easy 'instant' diagnoses of adnexal masses identified by six descriptors (Phases 1 and 2, total $n = 3511$)

Descriptor	*Predicted histology*	*Correct outcome (benign or malignant)*	*Correct histology*	*Cases in which descriptor was applicable*
Predicted outcome benign				1066*
Unilocular tumor with ground glass echogenicity in a premenopausal woman	Endometrioma	396/398 (99.5 (98.2–99.9))	360/398 (90.5 (87.2–93.0))	
Unilocular tumor with mixed echogenicity and acoustic shadows in a premenopausal woman	Teratoma	136/136 (100 (97.3–100))	126/136 (92.6 (87.0–96.0))	
Unilocular anechoic tumor with regular walls and maximum diameter of lesion < 10 cm	Simple cyst/ cystadenoma	240/243 (98.8 (96.4–99.6))	210/243 (86.4 (81.5–90.2))	
Remaining unilocular tumor with regular walls		285/289 (98.6 (96.5–99.5))		
Predicted outcome malignant				458†
Tumor with ascites and at least moderate color Doppler blood flow in a postmenopausal woman		194/203 (95.6 (91.8–97.7))		
Age > 50 years and CA 125 > 100 U/mL		386/414 (93.2 (90.4–95.3))		
Total				1518‡



Ultrasound Obstet Gynecol 2016; **48**: 397–402
Published online in Wiley Online Library (wileyonlinelibrary.com). **DOI:** 10.1002/uog.15854

External validation of IOTA simple descriptors and simple rules for classifying adnexal masses

J. L. ALCÁZAR*, M. A. PASCUAL†, B. GRAUPERA†, M. AUBÁ*, T. ERRASTI*, B. OLARTECOECHEA*, A. RUIZ-ZAMBRANA*, L. HERETER†, S. AJOSSA‡ and S. GUERRIERO‡

Sensitivity 92-95%
Specificity 87-98%

Ultrasound Obstet Gynecol 2019; 53: 693–700
Published online in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/uog.20163

Prospective external validation of IOTA three-step strategy for characterizing and classifying adnexal masses and retrospective assessment of alternative two-step strategy using simple-rules risk

J. J. HIDALGO[1,2], F. ROS[2], M. AUBÁ[3], T. ERRASTI[3], B. OLARTECOECHEA[3], Á. RUIZ-ZAMBRANA[3] and J. L. ALCÁZAR[3]

FULL PAPER

BJC

British Journal of Cancer (2014) 111, 680–688 | doi: 10.1038/bjc.2014.333

Keywords: ovarian cancer; diagnosis; meta-analysis

Strategies to diagnose ovarian cancer: new evidence from phase 3 of the multicentre international IOTA study

A Testa[1,17], J Kaijser[2,3,17], L Wynants[4,5], D Fischerova[6], C Van Holsbeke[7], D Franchi[8], L Savelli[9], E Epstein[10], A Czekierdowski[11], S Guerriero[12], R Fruscio[13], F P G Leone[14], I Vergote[3], T Bourne[2,3,15], L Valentin[16], B Van Calster[2] and D Timmerman*[2,3]

Gynecologic Oncology 130 (2013) 140–146

Contents lists available at SciVerse ScienceDirect

Gynecologic Oncology

journal homepage: www.elsevier.com/locate/ygyno

A multicenter prospective external validation of the diagnostic performance of IOTA simple descriptors and rules to characterize ovarian masses

Ahmad Sayasneh [a,b], Jeroen Kaijser [c], Jessica Preisler [b], Susanne Johnson [d], Catriona Stalder [b], Richard Husicka [b], Sharmistha Guha [b], Osama Naji [b], Yazan Abdallah [b], Fateh Raslan [e], Alexandra Drought [e], Alison A. Smith [f], Christina Fotopoulou [b], Sadaf Ghaem-Maghami [a,b,*], Ben Van Calster [g], Dirk Timmerman [c,g], Tom Bourne [a,b,g]

Radiology

ORIGINAL RESEARCH · STATEMENTS AND GU

O-RADS US Risk Stratification and Management System: A Consensus Guideline from the ACR Ovarian-Adnexal Reporting and Data System Committee

Rochelle F. Andreotti, MD • Dirk Timmerman, MD, PhD • Lori M. Strachowski, MD • Wouter Froyman, MD • Beryl R. Benacerraf, MD • Geneviève L. Bennett, MD • Tom Bourne, PhD • Douglas L. Brown, MD • Beverly G. Coleman, MD • Mary C. Frates, MD • Steven R. Goldstein, MD • Ulrike M. Hamper, MD, MBA • Mindy M. Horrow, MD • Marta Hernanz-Schulman, MD • Caroline Reinhold, MD, MSc • Stephen L. Rose, MD • Brad P. Whitcomb, MD • Wendy L. Wolfman, MD • Phyllis Glanc, MD

Radiology 2020; 294:168–185

O-RADS 0 = Incomplete evaluation

O-RADS 1 = normal ovaries

O-RADS 2

O-RADS 3

O-RADS 4

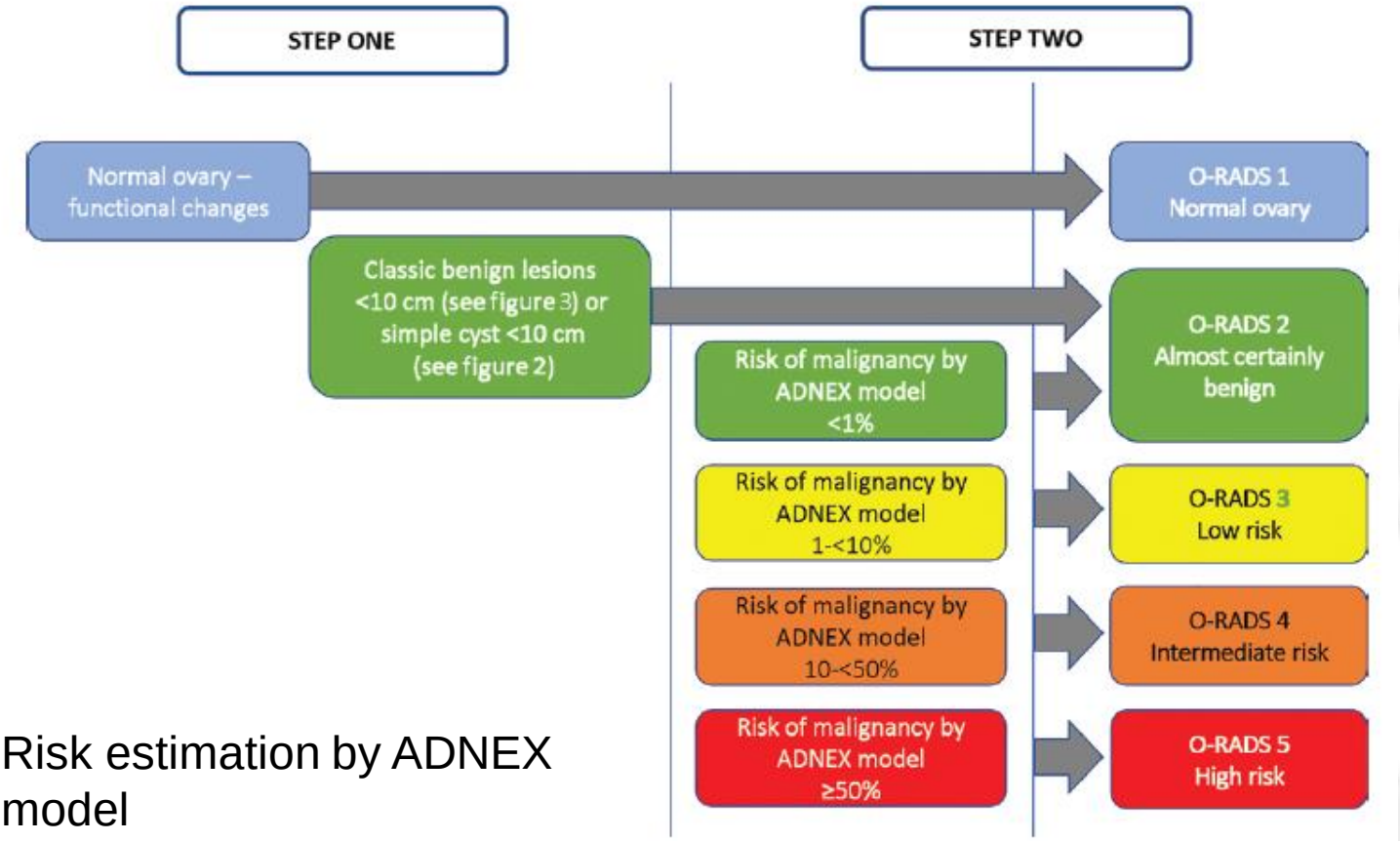
O-RADS 5

Based on IOTA terms and descriptions

Table: IOTA Phase 1–3 Data Used to Define O-RADS Categories

Description	Fulfill Criterion	Malignant
<1%	1452 (24.6)	7 (0.5)
Classic hemorrhagic cyst ≥5 cm to <10 cm	11 (0.2)	0 (0)
Classic dermoid cyst <10 cm	321 (5.4)	0 (0)
Classic endometrioma <10 cm	583 (9.9)	4 (0.7)
Unilocular smooth cyst ≤3 cm	54 (0.9)	0 (0)
Other unilocular smooth cyst >3 cm to <10 cm	483 (8.2)	3 (0.6)
1% to <10%	945 (16.0)	34 (3.6)
Unilocular smooth ≥10 cm	185 (3.1)	5 (2.7)
Unilocular irregular wall	101 (1.7)	4 (4.0)
Multilocular smooth CS 1–3 <10 cm	577 (9.8)	19 (3.3)
Solid smooth CS 1	82 (1.4)	6 (7.3)
10% to <50%	1734 (29.3)	516 (29.8)
Multilocular smooth ≥10 cm CS 1–3	227 (3.8)	41 (18.1)
Multilocular smooth CS 4	22 (0.4)	3 (13.6)
Multilocular irregular	182 (3.1)	35 (19.2)
Unilocular-solid no papillary projection	198 (3.4)	58 (29.3)
Unilocular-solid 1–3 papillary projections	338 (5.7)	98 (29.0)
Multilocular-solid CS 1–2	405 (6.9)	126 (31.1)
Solid smooth CS 2–3	362 (6.1)	155 (42.8)
50%–100%	1774 (30.0)	1374 (77.5)
Unilocular-solid with ≥4 papillary projections	94 (1.6)	64 (68.1)
Multilocular-solid CS 3–4	619 (10.5)	372 (60.1)
Solid smooth CS 4	135 (2.3)	104 (77.0)
Solid irregular	206 (3.5)	178 (86.4)
Ascites or metastases	720 (12.2)	656 (91.1)

Note.—Data are the number of lesions that fulfill each criteria, with percentages in parentheses. Source.—References 14, 15. CS = color score, IOTA = International Ovarian Tumor Analysis, O-RADS = Ovarian-Adnexal Reporting and Data System.



Risk estimation by ADNEX model

O-RADS Score	Risk Category [IOTA Model]	Lexicon Descriptors		Management	
				Pre-menopausal	Post-menopausal
0	Incomplete Evaluation [N/A]	N/A		Repeat study or alternate study	
1	Normal Ovary [N/A]	Follicle defined as a simple cyst ≤ 3 cm		None	N/A
		Corpus Luteum ≤ 3cm			
2	Almost Certainly Benign [< 1%]	Simple cyst	≤ 3 cm	N/A	None
			> 3 cm to 5 cm	None	Follow up in 1 year. *
			> 5 cm but < 10 cm	Follow up in 8 - 12 weeks	
		Classic Benign Lesions	See **Figure 3** for separate descriptors	See **Figure 3** for management strategies	
		Non-simple unilocular cyst, smooth inner margin	≤ 3 cm	None	Follow up in 1 year * If concerning, US specialist or MRI
			> 3 cm but < 10 cm	Follow-up in 8 - 12 weeks If concerning, US specialist	US specialist or MRI
3	Low Risk Malignancy [1-<10%]	Unilocular cyst ≥ 10 cm (simple or non-simple)		US specialist or MRI Management by gynecologist	
		Typical dermoid cysts, endometriomas, hemorrhagic cysts ≥ 10 cm			
		Unilocular cyst, any size with irregular inner wall <3 mm height			
		Multilocular cyst < 10 cm, smooth inner wall, CS = 1-3			
		Solid smooth, any size, CS = 1			
4	Intermediate Risk [10- < 50%]	Multilocular cyst, no solid component	≥ 10 cm, smooth inner wall, CS = 1-3	US specialist or MRI Management by gynecologist with GYN-oncologist consultation or solely by GYN-oncologist	
			Any size, smooth inner wall, CS = 4		
			Any size, irregular inner wall and/or irregular septation, any color score		
		Unilocular cyst with solid component	Any size, 0-3 papillary projections, CS = any		
		Multilocular cyst with solid component	Any size, CS = 1-2		
		Solid	Smooth, any size, CS = 2-3		
5	High Risk [≥ 50%]	Unilocular cyst, any size, ≥ 4 papillary projections, CS = any		GYN-oncologist	
		Multilocular cyst with solid component, any size, CS = 3-4			
		Solid smooth, any size, CS = 4			
		Solid irregular, any size, CS = any			
		Ascites and/or peritoneal nodules**			

Reservations Regarding O-RADS Recommendations

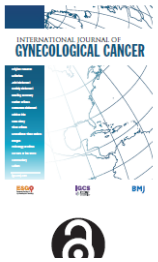
From
Elizabeth Suh-Burgmann, MD,* Tracy Flanagan, MD,† and Natasha Brasic, MD‡
Departments of Gynecologic Oncology,* Women's Health,† and Radiology,‡ Kaiser Permanente Northern California, 1425 S Main St, Walnut Creek, CA 94596
e-mail: *Betty.Suh-Burgmann@kp.org*

1. MRI is recommended for further evaluation of low and intermediate risk masses. However, to our knowledge, no prospective community-based studies have been performed that demonstrate that MRI improves clinical outcomes when used in populations with low cancer prevalence. Therefore, its cost effectiveness for triage in these settings is questionable, and raises concerns about potential added risks of incidental findings

2. Repeat US, which has been shown to help avoid surgical intervention for transient or benign masses (4), is not included as a management option for O-RADS 3 and 4 lesions, despite that this is the most common strategy used to further evaluate adnexal masses considered low or intermediate risk, because observation of stability versus growth is clinically meaningful.

3. The definition of low risk as cancer risk of 1% to less than 10% is inconsistent with most views of ovarian cancer risk. A woman with a 9% risk of cancer would not generally be considered low risk given that prophylactic surgery is recommended for women with *BRCA2* mutations who are considered high risk on the basis of a lifetime risk of ovarian cancer of 10%–28% (5).

4. The authors' statement that "the O-RADS system has been developed for the patient at average risk" seems incongruent with the acknowledgment that the International Ovarian Tumor Analysis data on which the system is based is drawn from preoperative US in high-risk European referral populations, which is expected to inflate the observed risk.



ESGO/ISUOG/IOTA/ESGE Consensus Statement on pre-operative diagnosis of ovarian tumors

Dirk Timmerman [1,2] François Planchamp,[3] Tom Bourne [1,2,4] Chiara Landolfo [5] Andreas du Bois,[6] Luis Chiva [7] David Cibula [8] Nicole Concin [6,9] Daniela Fischerova [8] Wouter Froyman [1] Guillermo Gallardo Madueño [10] Birthe Lemley,[11,12] Annika Loft,[13] Liliana Mereu,[14] Philippe Morice,[15] Denis Querleu [16,17] Antonia Carla Testa [5,18] Ignace Vergote,[19] Vincent Vandecaveye [20,21] Giovanni Scambia,[5,18] Christina Fotopoulou [22]

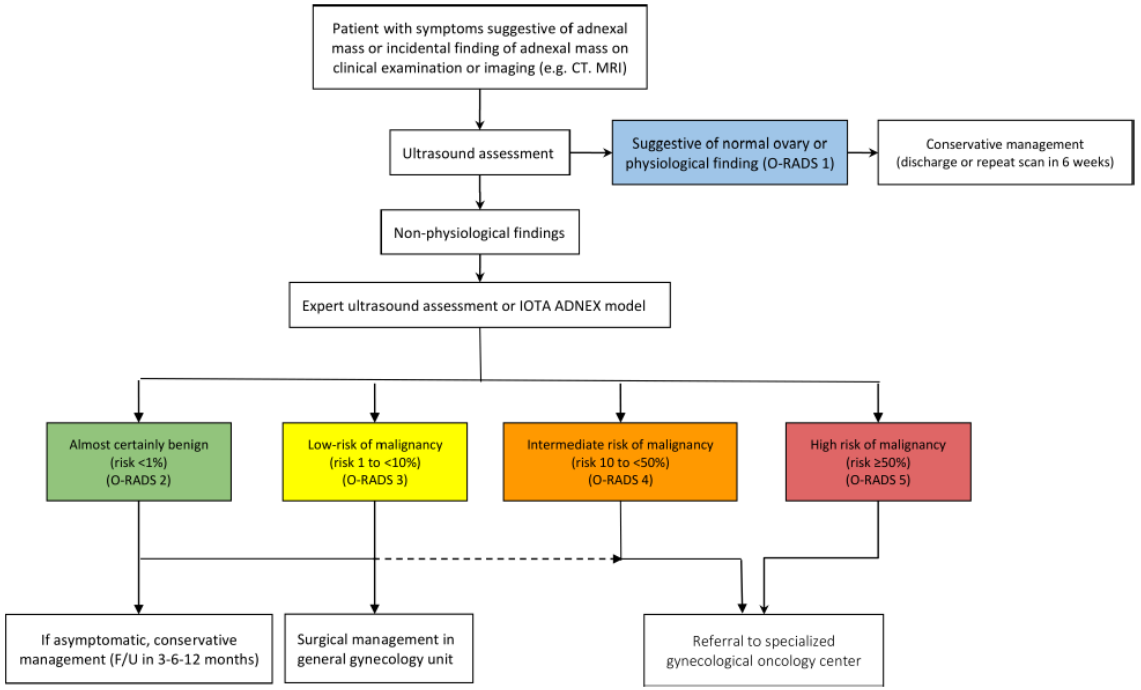


Figure 2 Flowchart of steps recommended to distinguish between benign and malignant tumors and to direct patients towards appropriate treatment pathway. CT, computed tomography; F/U, follow-up; IOTA ADNEX, International Ovarian Tumour Analysis Group Assessment of Different NEoplasias in the adneXa; MRI, magnetic resonance imaging; O-RADS, Ovarian-Adnexal Reporting and data system.

Ovarian-Adnexal Reporting and Data System (O-RADS)

The Ovarian-Adnexal Reporting and Data System (O-RADS) lexicon for ultrasound was published in 2018, providing a standardized glossary that includes all appropriate descriptors and definitions of the characteristic ultrasound appearance of normal ovaries and various adnexal lesions.[72 73] The O-RADS ultrasound working group developed an adnexal mass triage system based either on the O-RADS descriptors or on the risk of malignancy assigned to the mass using the IOTA ADNEX model to classify ovarian tumors into different risk categories.[74] However, at present, neither the triage system nor the O-RADS descriptors have been externally validated. Basha et al[75] determined the malignancy rates, validity, and reliability of the O-RADS approach when applied to a database of 647 adnexal masses collected before the development of the O-RADS system. In this retrospective study, the O-RADS system had significantly higher sensitivity than did the GI-RADS system and the IOTA Simple Rules, with a non-significant slightly lower specificity compared with both GI-RADS and IOTA Simple Rules, and with similar reliability.

Letter

Correspondence on "ESGO/ISUOG/IOTA/ESGE consensus statement on pre-operative diagnosis of ovarian tumors" by Timmerman et al

We read with much interest the consensus paper by Timmerman et al[1] on the pre-operative diagnosis of ovarian tumors, which included representatives from ESGO/ISUOG/IOTA/ESGE. Although the paper details the important value of ultrasonography in this clinical setting, it does not adequately address the 25% of women in whom the assessment with ultrasound remains insufficient, as has been demonstrated by multiple publications.[2 3]

In the setting of sonographically indeterminate adnexal lesions, magnetic resonance imaging (MRI) has been shown to be very useful for improving characterization, with a large evidence-based body of literature that includes investigational studies, meta-analyses, and systematic reviews.[3–11] Figure 2 in the paper deals with the management of women according to the O-RADS classification system. However, it is important to point out that a separate O-RADS score exists for ultrasonography (O-RADS US) and for MRI (O-RADS MRI). In this Figure, the authors use the term O-RADS; it is important for the reader to note that this article refers to the O-RADS US scoring system. Furthermore, the American College of Radiology consensus, in which the IOTA group was involved, recommends MRI for further evaluation of O-RADS US 3 and 4 category lesions. In the O-RADS US 3 lesion score category, the positive predictive value for malignancy is only approximately 10%, and proceeding to surgical resection without any additional imaging evaluation may potentially result in 90% of surgeries being performed for physiologic lesions and benign tumors. To date, MRI has largely demonstrated its utility by reclassifying ovarian lesions that would fall into the O-RADS US 3 or 4. This is particularly relevant for pre-menopausal women where a crucial issue is the preservation of fertility, as well as the poor outcomes for potentially tipping a young patient into early hormonal menopause. Indeed, gynecologists Anthoulakis and Nikoloudis wrote: "the preponderant contribution of MRI in adnexal mass evaluation is its specificity because it provides confident diagnosis of many benign adnexal lesions".[12]

Most ovarian lesions are benign, and many are found incidentally as part of investigations for other clinical reasons. Figure 2 in the paper by Timmerman et al suggests that, when an ovarian mass is discovered at MRI, ultrasonography would be useful to further characterize the adnexal lesion. However, for the characterization of ovarian lesions, if the MRI was protocoled as a pelvic MRI for adnexal lesion characterization, ultrasonography may not be required. As pointed out by the authors: "O-RADS MRI score is another modern diagnostic development with promising potential for the differentiation of benign from malignant adnexal masses". This statement is supported by a large prospective European cohort (n=1194 patients) that validated the very high level of sensitivity and specificity (higher than 90%) of MRI in women with indeterminate or complex ovarian masses at ultrasonography.[13] Our patients are served best when both gynecologists and radiologists play a complementary role in the pre-operative evaluation of women with ovarian lesions. This critical point is not sufficiently addressed in this consensus statement which, in addition to the lack of representation by a radiological society, may limit wide acceptance of this statement in clinical practice.

Isabelle Thomassin-Naggara ,[1,2] **Elizabeth Sadowski**,[3] **Andrea Rockall**,[4,5] **Caroline Reinhold**[6]

[1]IRIS Department of Radiology, Tenon Hospital, APHP, Paris, France
[2]Sorbonne Université, Paris, France
[3]Department of Radiology and Obstetrics and Gynecology, Madison, Wisconsin, USA
[4]Faculty of Medicine, Department of Surgery and Cancer, Imperial College London, London, UK
[5]Department of Radiology, Imperial College Healthcare NHS Trust, London, UK
[6]Department of Radiology, Augmented Intelligence & Precision Health Laboratory of the Research Institute of McGill University Health Centre, Montreal, Québec, Canada

Correspondence to Professor Isabelle Thomassin-Naggara, APHP, 75184 Paris, France; isabelle.thomassin@aphp.fr

Twitter Isabelle Thomassin-Naggara @Naggaraisabelle

Contributors All authors contributed to writing the letter.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests None declared.

Patient consent for publication Not required.

Provenance and peer review Not commissioned; internally peer reviewed.

To cite Thomassin-Naggara I, Sadowski E, Rockall A, *et al*. *Int J Gynecol Cancer* 2021;31:1394–1395.

Accepted 12 August 2021
Published Online First 20 August 2021

Linked

► http://dx.doi.org/10.1136/ijgc-2021-002565

► http://dx.doi.org/10.1136/ijgc-2021-003013

Int J Gynecol Cancer 2021;31:1394–1395. doi:10.1136/ijgc-2021-002910

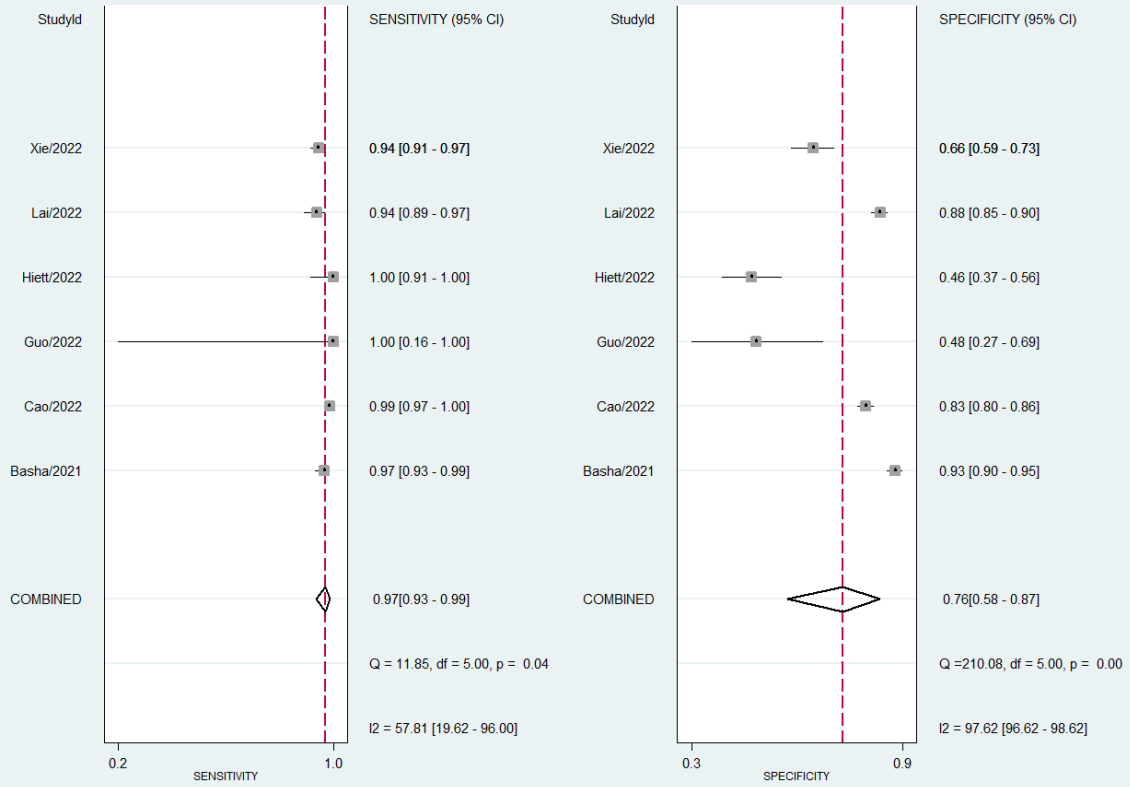
ORCID iD
Isabelle Thomassin-Naggara http://orcid.org/0000-0002-0919-0133

REFERENCES

1 Timmerman D, Planchamp F, Bourne T, *et al*. ESGO/ISUOG/IOTA/ESGE consensus statement on pre-operative diagnosis of ovarian tumors. *Int J Gynecol Cancer* 2021;31:961–82.
2 Chilla B, Hauser N, Singer G, *et al*. Indeterminate adnexal masses at

TITLE: Diagnostic accuracy of ultrasound O-RADS for classifying adnexal mass: systematic review and meta-analysis

AUTHORS (FIRST NAME, LAST NAME): Julio Vara[1], Ana Lopez-Picazo[1], M. Angela Pascual[2], Stefano Guerriero[3], Juan Luis Alcazar[1]



6 studies
3006 masses
Pooled sensitivity: 97%
Pooled specificity: 76%

All studies IOTA Lexicon

TITLE: O-RADS classification for ultrasound assessment of adnexal masses: agreement between IOTA lexicon and ADNEX model

AUTHORS (FIRST NAME, LAST NAME): Julio Vara[1], Serena Springer[3], Mariachiara Pagliuca[2], Juan Gonzalez de Canales[1], Javiera Yakcich[4], M. Angela Pascual[5], Stefano Guerriero[2], Juan Luis Alcazar[1]

- Ambispective study
- 454 masses in 412 patients
  - Weighted Kappa: 0.42
  - Percentage of agreement: 39%

Distribution of cases using O-RADS classification

		O-RADS using ADNEX model			
		O-RADS 2	O-RADS 3	O-RADS 4	O-RADS 5
O-RADS using IOTA lexicon	O-RADS 2	30	218	2	0
	O-RADS 3	6	75	10	0
	O-RADS 4	0	23	40	7
	O-RADS 5	0	1	11	31

TITLE: Diagnostic performance of ultrasound O-RADS classification for adnexal masses: A comparative study between using IOTA lexicon and ADNEX model

AUTHORS (FIRST NAME, LAST NAME): Julio Vara[1], Serena Springer[5], Mariachiara Pagliuca[3], Juan Gonzalez de Canales[1], Javiera Yakcich[6], Silvia Ajossa[3], M. Angela Pascual[4], Stefano Guerriero[2], Juan Luis Alcazar[1]

Diagnostic performance of different approaches

	Sensitivity	Specificity	LR +	LR -
O-RADS using IOTA lexicon (A)	92% (95%CI: 83%-97%)	86% (95%CI: 82%-89%)	6.6 (95%CI: 5.1-8.6)	0.09(95%CI: 0.04-0.21)
O-RADS using Adnex model (B)	86% (95%CI: 75%-92%)	88% (95%CI: 84%-91%)	7.1 (95%CI: 5.3-9.5)	0.16 (95%CI: 0.09-0.30)
Expert subjective impression (C)	97% (95%CI: 89%-99%)	94% (95%CI: 91%-96%)	16.4 (95%CI: 11.0-24-5)	0.03 (95%CI: 0.008-0.13)

A vs B, p=0.065. C vs B, p = 0.01. C vs A, p=0.048

- Conclusions
  - There are serveral IOTA models
  - Some of them are complex, other are user-friendly
  - Need for app or calculator for some of them
- Questions
  - Which is the best in terms of performance?
  - Are they reproducible?
  - Which one should be used?
  - Can any replace expert evaluation?

- LR1/LR2
  - Adequate external validation
  - Good performance
  - Reproducible
  - Need for app/calculator
- Simple rules
  - Adequate external validation
  - Good performance
  - Reproducible
  - No Need for app/calculator

- SR risk assessment
  - Not yet adequate external validation
  - Good performance
  - Reproducibility??
  - Need for app/calculator
- ADNEX model
  - Adequate external validation
  - Good performance (sensitivity!!)
  - Reproducibility??
  - Need for app/calculator

- Three step strategy
  - Adequate external validation
  - Good performance
  - Reproducibility???
  - No Need for app/calculator

- O-RADS classification
  - Not yet adequate external validation
  - Good performance (sensitivity!!)
  - Reproducibility??
  - Discrepancies depending on criteria used
  - May need for app/calculator

- Questions
  - Which is the best in terms of performance?
    - All of them offer excelent sensitivity
    - SR and Three-steps offer better specificity
  - Are they reproducible?
    - Only SR and LR1 / 2 have been demonstrated as reproducible
  - Can any replace expert evaluation?
    - No
  - Which one should be used?

YOU CHOOSE

CUN Pamplona

CUN Madrid

Grazie

jlalcazar@unav.es
@Juan_L_Alcazar