10.5005/jp-journals-10009-1514

ORIGINAL RESEARCH

External Validation of Gynecological Imaging and Reporting Data System for Sonographic Evaluation of Adnexal Masses

[1]Linder Diaz, [2]Belkys Zambrano, [3]Fabio J Adami, [4]Juan L Alcázar

ABSTRACT

Objective: To perform an external validation of Gynecological Imaging and Reporting Data System (GI-RADS) and to assess how referring clinicians value this reporting system in their daily practice.

Materials and methods: Prospective observational study comprising 257 women (mean age 40.3 years) and 281 adnexal masses, referred by 20 clinicians to an ultrasound referral center. All women underwent transvaginal or transrectal ultrasound. Presumed diagnosis of the adnexal mass was based on examiner's subjective impression according to patter recognition analysis. Reporting was performed according to GI-RADS classification. Sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), positive likelihood ratio (LR+), and negative likelihood ratio (LR−) of the GI-RADS were calculated. The gold standard was histologic diagnosis (benign or malignant) or spontaneous resolution of the cyst during follow-up (benign). Referring clinicians were asked for completing a survey in order to assess how useful they considered this reporting system.

Results: In this study, 56 masses were classified as GI-RADS 2, 174 masses were classified as GI-RADS 3, 19 masses were classified as GI-RADS 4, and 32 masses were classified as GI-RADS 5. Among them, 230 masses were removed surgically and 51 masses resolved spontaneously. There were 35 malignant lesions. Sensitivity, specificity, LR+, and LR− were 97.1% (95% confidence interval (CI): 85.5–99.5%), 93.1% (95%CI: 89.2–95.6%), 14.1 (95%CI: 8.8–22.3), and 0.03 (95%CI: 0.004–0.21) respectively. All 20 referring clinicians reported that GI-RADS was very useful for their clinical decision-making process.

Conclusion: GI-RADS shows a high diagnostic performance and is helpful for referring clinicians for taking clinical decisions.

Keywords: Adnexal masses, Diagnosis, Management, Reporting, Ultrasound.

How to cite this article: Diaz L, Zambrano B, Adami FJ, Alcázar JL. External Validation of Gynecological Imaging and Reporting Data System for Sonographic Evaluation of Adnexal Masses. Donald School J Ultrasound Obstet Gynecol 2017;11(2):135-140.

[1-4]Consultant

[1,2]Department of Obstetrics and Gynecology, Centro Clinico Carali 2, Barquisimeto, Venezuela

[3]Department of Radiology, Imaging Unit, Instituto Diagnóstico Barquisimeto, Barquisimeto,Venezuela

[4]Department of Obstetrics and Gynecology, Clinica Universidad de Navarra, University of Navarra, Pamplona, Spain

Corresponding Author: Juan L Alcázar, Consultant, Department of Obstetrics and Gynecology, Clinica Universidad de Navarra University of Navarra, Pamplona, Spain, Phone: +34948245400 email: jlacazar@unav.es

Source of support: Nil

Conflict of interest: None

INTRODUCTION

Adnexal masses constitute one of the commonest clinical problems in gynecology. Accurate diagnosis is essential for adequate management.[1] Transvaginal ultrasound is considered as the first line imaging technique for discriminating benign from malignant adnexal masses.[2] Subjective impression of the examiner performing ultrasound evaluation remains as the best approach for this task.[3] However, diagnostic performance of this approach depends heavily on the examiner's experience and confidence.[4,5]

An additional problem is the quality of reporting.[6] Inappropriate or low-quality reporting may lead to unwarranted concern by the patient and referring clinician, which could lead to unnecessary additional tests and surgery.[7] In 2009, Amor et al[8] proposed the so-called Gynecological Imaging and Reporting Data System (GI-RADS) for reporting ultrasound evaluation of adnexal masses. This reporting system is based on the idea of breast imaging (BI-RADS) Lexicon for breast ultrasound.[9] As a matter of fact, similar lexicons have been proposed for imaging evaluation of other organs, such as thyroid (TI-RADS),[10] liver (LI-RADS),[11] or prostate (PI-RADS).[12]

To date, studies validating GI-RADS are very scanty. The objective of the present study is to perform an external validation of GI-RADS and to assess how referring clinicians value this reporting system in their daily practice.

MATERIALS AND METHODS

This is a prospective observational study performed at one single institution (Centro Clinico Carali 2, Barquisimeto, Venezuela) from December 2011 to December 2015. This center is a private practice specialized on gynecological ultrasound where patients are usually referred for ultrasound evaluation.

Eligible patients were all consecutive patients referred during study period to this center for gynecological ultrasound for the clinical suspicion of the presence of an adnexal mass (clinical symptoms, such as pain, suspicion

on pelvic bimanual examination) or for a second opinion for a mass diagnosed at ultrasound by referring clinician.

All women gave verbal informed consent after the nature of the study was explained by one of the authors, Linder Diaz. Institutional Review Board was waived because all procedures performed were done on routine basis.

Exclusion criteria were as follows: Patients under 18 years old, pregnancy at the moment of ultrasound evaluation, surgery not performed in case of surgical indication, and lost to follow-up.

All women underwent transvaginal or transrectal ultrasound (e.g., women with virgo intacta) using a ACUSON X150 ultrasound system (Siemens Health care, Mountain View, CA, USA) equipped with a 4 to 9 MHz transvaginal transducer. Large masses were also evaluated by transabdominal ultrasound. All examinations were performed by one of the authors (LD) with more than 5 years' experience on gynecologic ultrasound and who were specifically trained for sonographic evaluation of adnexal masses before starting the study.[13]

Presumed diagnosis of the adnexal mass was based on examiner's subjective impression according to patter recognition analysis.[14] Reporting was performed according to GI-RADS classification[8] (Table 1). Findings were classified as GI-RADS 1, GI-RADS 2 (Fig. 1), GI-RADS 3 (Fig. 2), GI-RADS 4 (Fig. 3), and GI-RADS 5 (Fig. 4). During the examination, tumor size was estimated measuring the diameters in the three orthogonal planes.

Reporting was done using the model proposed by Amor et al.[8] The referring clinicians received this report and they were asked for managing the patients according to the protocol suggested by Amor et al.[15] Those patients classified as GI-RADS 1 needed no follow-up (they were excluded from the study and from further analysis). GI-RADS 2 patients were considered for expectant management by follow-up sonography. Based on that these lesions were assumed to be functional. GI-RADS 3 patients underwent surgery by general gynecologists or general surgeons. Based on that these lesions were considered probably benign and expected to persist over time. Patients classified as GI-RADS 4 and 5 were referred to gynecological oncologists for appropriate surgical management; based on that these lesions were considered probably or very probably malignant.

Table 1: Gynecologic Imaging Report and Data System (GI-RADS) classification system for adnexal masses

GI-RADS grade	*Diagnosis*	*EPM*	*Description*
1	Definitive benign	0%	Normal ovaries identified and no adnexal mass seen
2	Very probable benign	<1%	Adnexal lesions thought to be of functional origin, e.g., follicles, corpora lutea, hemorrhagic cysts
3	Probably benign	1–5%	Neoplastic adnexal lesions thought to be benign, such as endometrioma, teratoma, simple cyst, hydrosalpinx, paraovarian cyst, peritoneal pseudocyst, pedunculated myoma, or findings suggestive of pelvic inflammatory disease
4	Probably malignant	5–20%	Any adnexal lesion not included in GI-RADS 1 to 3 and with one or two findings suggestive of malignancy*
5	Very probable malignant	>20%	Adnexal masses with three or more findings suggestive of malignancy*

*Thick papillary projections, thick septations, solid areas and/or ascites, and vascularization within solid areas, papillary projections or central area of a solid tumor on color or power Doppler assessment; EPM: Estimated probability of malignancy

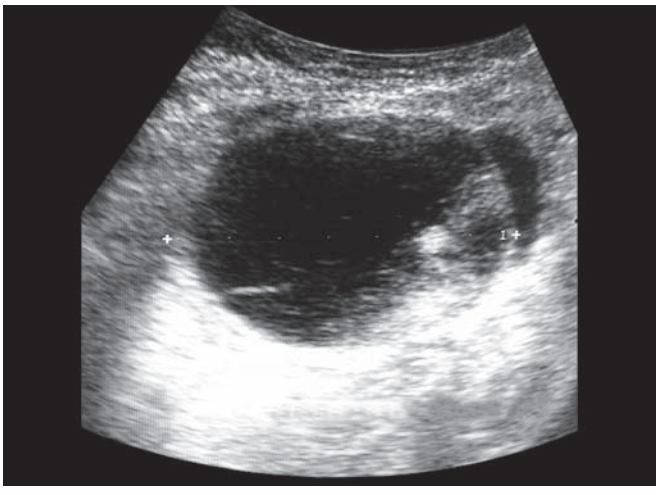

Fig. 1: An adnexal mass classified as GI-RADS 2. It was considered as a hemorrhagic cyst. Spontaneous resolution was confirmed at follow-up scan 3 months after diagnosis

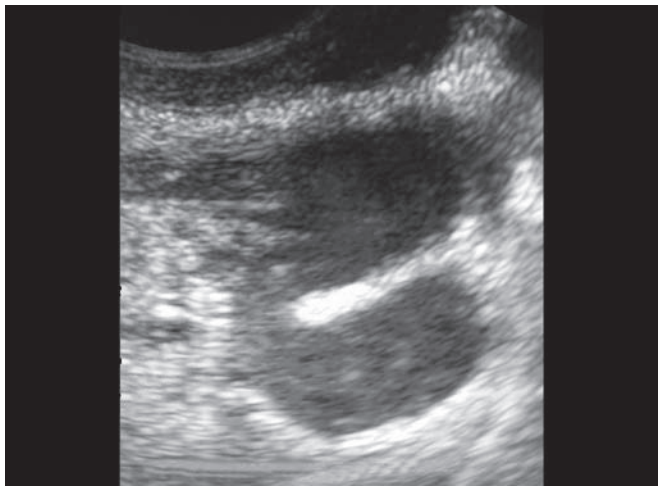

Fig. 2: Adnexal classified as GI-RADS 3. Examiner's diagnosis was hydrosalpinx. It was confirmed after surgical removal

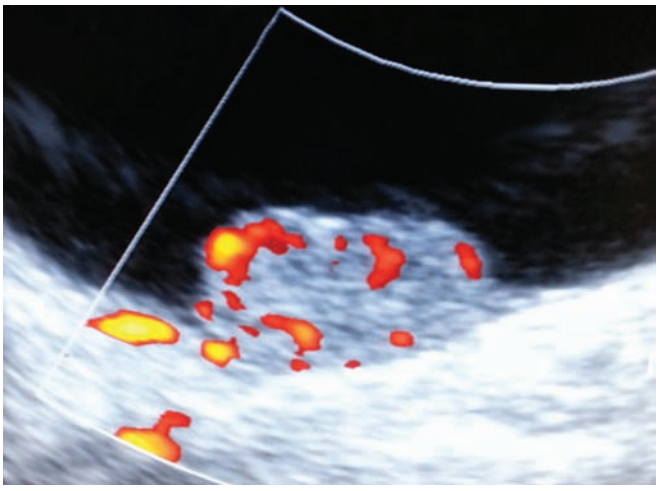

Fig. 3: Adnexal mass classified as GI-RADS 4. There were two suspicious features (solid area and vascularization within solid area). Examiner's diagnosis was borderline tumor. Histology revealed a cystadenofibroma

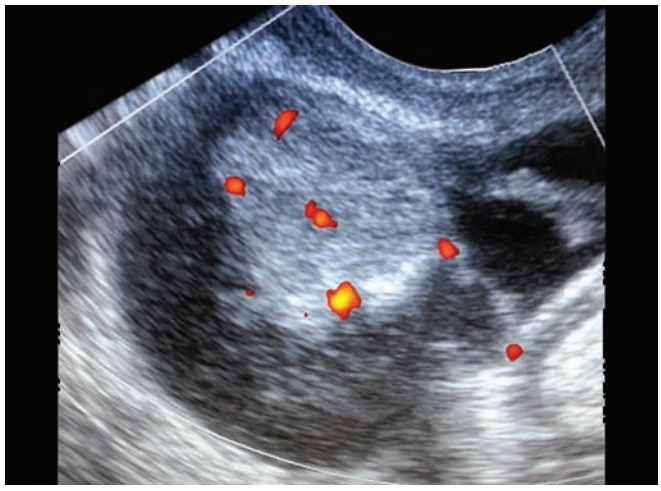

Fig. 4: Adnexal mass classified as GI-RADS 5. There were three suspicious features (solid area, vascularization within solid area, and ascites). Examiner's diagnosis was ovarian malignancy. It was confirmed histologically as serous carcinoma after surgical removal

When surgical removal of the tumor was performed, a definitive histologic diagnosis was obtained. Tumors were classified according to World Health Organization.[16] Borderline tumors were considered as malignant for analytic purposes.

We calculated the sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), positive likelihood ratio (LR+), and negative likelihood ratio (LR−) of the GI-RADS system for identifying adnexal masses at high risk of malignancy, considering GI-RADS 2 and 3 as low risk and GI-RADS 4 and 5 as high risk. The gold standard was histologic diagnosis (benign or malignant) or spontaneous resolution of the cyst during follow-up (benign).

We assessed how useful referring clinicians found the GI-RADS for making decisions regarding patient management asking them to complete a simple survey. We used the same single question stated by Amor et al[15]: "How useful do you think GI-RADS reporting system is for understanding ultrasound findings and giving confidence in clinical decisions regarding your patient?" and there were five possible answers: (A) Totally useful; (B) quite useful; (C) neither useful nor useless; (D) useless; (E) completely useless.

To assess interobserver reproducibility of GI-RADS classification, two examiners (LD and BZ) performed a separate analysis in the first 43 women included in the study. Both examiners performed a transvaginal scan, blinded to each other's results, and each one provided a GI-RADS report. To determine the concordance between examiners, we used Cohen's weighted Kappa index (κ) with 95% confidence interval. A κ of <0.20 indicates poor agreement; 0.21 to 0.40 indicates fair agreement; 0.41 to 0.60 indicates moderate agreement; 0.61 to 0.80 indicates good agreement; and 0.81 to 1.00 indicates very good agreement.[17]

RESULTS

During study period, 441 women were referred for gynecological ultrasound for the clinical suspicion of the presence of an adnexal mass or for a second opinion after a diagnosis of an adnexal mass. Twenty referring clinicians participated in this study (10 general gynecologists, 3 oncologic gynecologists, 3 general surgeons, and 4 general practitioners).

One hundred and eighty-four women were excluded for the following reasons: Four women were under 18 years old, two were pregnant at the time of diagnosis, 80 were classified as GI-RADS 1, 46 women classified as GI-RADS 3 did not undergo surgery, and 52 women were lost to follow-up.

A total of 257 women were ultimately included in the analysis. Twenty-four women had bilateral masses, giving a total number of 281 masses analyzed. Patients' mean age was 40.3 years old (SD: 15.1), ranging from 18 to 87 years old; 200 (77.8%) were premenopausal and 57 (22.2%) were postmenopausal.

Median tumor size was 5.0 cm, ranging from 1 to 40 cm.

Fifty-six masses (19.9%) were classified as GI-RADS 2, 174 masses (61.9%) were classified as GI-RADS 3, 19 masses (6.8%) were classified as GI-RADS 4, and 32 masses (11.4%) were classified as GI-RADS 5.

Two hundred and thirty (81.8%) were removed surgically and 51 masses resolved spontaneously. Histologic diagnoses are shown in Table 2. The correlation between GI-RADS classification and final histological diagnosis is shown in Table 3.

Table 2: Histologic diagnosis of masses surgically removed (n = 230)

Histology	*n*	*Percentage*
Endometrioma	53	23.0
Serous cyst	32	13.9
Dermoid cyst	22	9.6
Hydrosalpinx	17	7.4
Mucinous cyst	16	6.9
Peritoneal cyst	15	6.5
Leiomyoma	10	4.3
Fibroma/Fibrothecoma	8	3.5
Cystadenofibroma	6	2.6
Para-ovarian cyst	6	2.6
Tubo-ovarian abscess	6	2.6
Hemorrhagic cyst	3	1.3
Functional cyst	1	0.4
Borderline ovarian tumor	6	2.6
Primary invasive ovarian cancer	26	11.3
Metastatic cancer	3	1.3

Flow Chart 1 shows how patients (and masses) were managed according to GI-RADS classification.

Considering GI-RADS 4 and 5 as malignant, the sensitivity, specificity, LR+, and LR- were 97.1% (95%CI: 85.5–99.5%), 93.1% (95%CI: 89.2–95.6%), 14.1 (95%CI: 8.8–22.3), and 0.03 (95%CI: 0.004–0.21) respectively.

Interobserver agreement for classifying adnexal masses according to GI-RADS system was good (weighted Kappa: 0.804, 95% CI: 0.656–0.952).

All 20 referring clinicians reported that GI-RADS was very useful for their clinical decision-making process.

DISCUSSION

In the present study, we have performed an external validation of the GI-RADS. Our results show that this system has a very good performance for classifying adnexal masses. But, more importantly in our opinion, we found out that this reporting system was highly appreciated by referring clinicians as they could understand better the report of the ultrasound examination and make clearer to them the findings observed as well as they felt more confident to take clinical decisions with their patients.

Since the pioneering report from Amor et al,[8] only four studies have been reported evaluating the diagnostic performance of GI-RADS.[15,18-20] The results of these studies are summarized in Table 4. All these studies reported

Table 3: Final diagnosis according to GI-RADS classification (n = 281)

	GI-RADS 2	*GI-RADS 3*	*GI-RADS 4*	*GI-RADS 5*	*Total*
Functional cyst	39	0	0	0	39
Endometrioma	1	52	0	0	53
Serous cyst	3	27	2	0	32
Dermoid cyst	0	20	2	0	22
Hydrosalpinx	0	17	0	0	17
Mucinous cyst	0	12	3	1	16
Peritoneal cyst	0	15	0	0	15
Leiomyoma	0	9	0	1	10
Fibroma/Fibrothecoma	0	6	2	0	8
Cystadenofibroma	0	3	1	2	6
Para-ovarian cyst	0	6	0	0	6
Tubo-ovarian abscess	0	3	1	2	6
Hemorrhagic cyst	13	3	0	0	16
Borderline ovarian tumor	0	1	3	2	6
Primary invasive ovarian cancer	0	0	4	22	26
Metastatic cancer	0	0	1	2	3
Total	56	174	19	32	281

Flow Chart 1: Masses management according to GI-RADS classification and final results

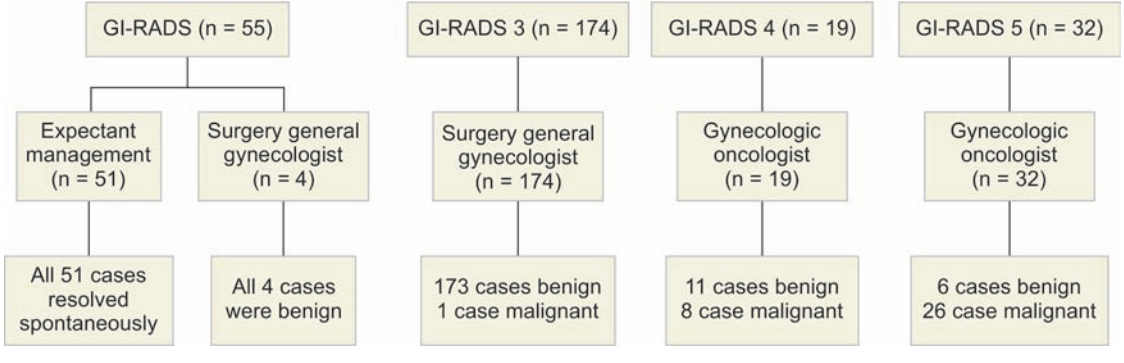

Table 4: Summary of studies reported evaluating GI-RADS reporting system

Author	*Year*	*n*	*Cancer prevalence (%)*	*Sensitivity (%)*	*Specificity (%)*	*LR+*	*LR−*
Amor et al[15]	2011	432	26.0	99.1	85.9	7.05	0.01
Qiuyue and Guorong[18]	2013	999	10.0	99.0	90.0	10.73	0.01
Rams et al[19]	2015	98	14.2	100	89.2	9.33	NA
Orozco Fernández et al[20]	2015	387	16.8	84.9	96.3	22.70	0.20

similar diagnostic performance, with sensitivity ranging from 85 to 100% and specificity ranging from 86 to 96%. Our results are in agreement with these data.

Like in all these four studies, we observed that many cases classified as GI-RADS 4 were actually benign. Therefore, this group represents a significant source of false-positive cases. We think that the definition of this group should be reconsidered. According to Amor et al,[8] GI-RADS 4 is defined by the presence of one or two criteria suspicious of malignancy. Probably, splitting in this group into two (e.g., GI-RADS 4a – one single suspicious feature – and GI-RADS 4b – two suspicious features – assigning a higher risk for malignancy to GI-RADS 4b) could help in reducing the false-positive rate of this group. We have not made any attempt to make this as subgroup analysis due to the small number of cases we had in this group.

We consider more encouraging results regarding how referring clinicians evaluated this reporting system and its utility for clinical decision-making. All 20 referring clinicians considered this reporting system as very useful. We think this may have important implications in clinical practice, since reporting in ultrasound assessment of adnexal masses has been acknowledged as a significant problem due to poor quality and lack of standardization.[6,7]

The main limitation of our study is that all examinations were performed by a single examiner, specifically trained for assessing adnexal masses by ultrasound. Therefore, we do not know whether this reporting system would render similar results in hands of less skilled in ultrasound performance and image interpretation. However, we observed that reproducibility of this reporting system was high.

In summary, we found that GI-RADS reporting system is useful for classifying adnexal masses and is helpful for referring clinicians for taking clinical decisions and managing women with adnexal masses.

REFERENCES

1. Liu JH, Zanotti KM. Management of the adnexal mass. Obstet Gynecol 2011 Jun;117(6):1413-1428.
2. American College of Obstetricians and Gynecologists. ACOG Practice Bulletin. Management of adnexal masses. Obstet Gynecol 2007 Jul;110(1):201-214.
3. Valentin L, Hagen B, Tingulstad S, Eik-Nes S. Comparison of 'pattern recognition' and logistic regression models for discrimination between benign and malignant pelvic masses: a prospective cross validation. Ultrasound Obstet Gynecol 2001 Oct;18(4):357-365.
4. Guerriero S, Alcazar JL, Pascual MA, Ajossa S, Gerada M, Bargellini R, Virgilio B, Melis GB. Intraobserver and interobserver agreement of grayscale typical ultrasonographic patterns for the diagnosis of ovarian cancer. Ultrasound Med Biol 2008 Nov;34(11):1711-1716.
5. Van Holsbeke C, Daemen A, Yazbek J, Holland TK, Bourne T, Mesens T, Lannoo L, Boes AS, Joos A, Van De Vijver A, et al. Ultrasound experience substantially impacts on diagnostic performance and confidence when adnexal masses are classified using pattern recognition. Gynecol Obstet Invest 2010 Jan;69(3):160-168.
6. Le T, Fayadh RA, Menard C, Hicks-Boucher W, Faught W, Hopkins L, Fung-Kee-Fung M. Variations in ultrasound reporting on patients referred for investigation of ovarian masses. J Obstet Gynaecol Can 2008 Oct;30(10):902-906.
7. Brown DL, Dudiak KM, Laing FC. Adnexal masses: US characterization and reporting. Radiology 2010 Feb;254(2): 342-354.
8. Amor F, Vaccaro H, Alcázar JL, León M, Craig JM, Martinez J. Gynecologic imaging reporting and data system: a new proposal for classifying adnexal masses on the basis of sonographic findings. J Ultrasound Med 2009 Mar;28(3):285-291.
9. American College of Radiology. BI-RADS: ultrasound. In: Breast imaging reporting and data system: BI-RADS atlas. 4th ed. Reston (VA): American College of Radiology; 2003.
10. Xu SY, Zhan WW, Wang WH. Evaluation of thyroid nodules by a scoring and categorizing method based on sonographic features. J Ultrasound Med 2015 Dec;34(12):2179-2185.
11. Darnell A, Forner A, Rimola J, Reig M, García-Criado Á, Ayuso C, Bruix J. Liver imaging reporting and data system with MR imaging: evaluation in nodules 20 mm or smaller detected in cirrhosis at screening US. Radiology 2015 Jun;275(3):698-707.
12. Cash H, Maxeiner A, Stephan C, Fischer T, Durmus T, Holzmann J, Asbach P, Haas M, Hinz S, Neymeyer J, et al. The detection of significant prostate cancer is correlated with the Prostate Imaging Reporting and Data System (PI-RADS) in MRI/transrectal ultrasound fusion biopsy. World J Urol 2016 Apr;34(4):525-532.
13. Alcázar JL, Díaz L, Flórez P, Guerriero S, Jurado M. Intensive training program for ultrasound diagnosis of adnexal masses: protocol and preliminary results. Ultrasound Obstet Gynecol 2013 Aug;42(2):218-223.
14. Guerriero S, Ajossa S, Gerada M, Virgilio B, Pilloni M, Galvan R, Laparte MC, Alcázar JL, Melis GB. Transvaginal ultrasonography in the diagnosis of extrauterine pelvic diseases. Expert Rev Obstet Gynecol 2008;3(6):731-752.
15. Amor F, Alcázar JL, Vaccaro H, León M, Iturra A. GI-RADS reporting system for ultrasound evaluation of adnexal masses in clinical practice: a prospective multicenter study. Ultrasound Obstet Gynecol 2011 Oct;38(4):450-455.

16. Scully, RE.; Sobin, LH. WHO histological classification of ovarian tumors. Geneva: World Health Organization; 1999.
17. Brennan P, Silman A. Statistical methods for assessing observer variability in clinical measures. BMJ 1992 Jun;304(6840): 1491-1494.
18. Qiuyue C, Guorong L. The applicational value of GI-RADS ultrasonographic stratification in diagnosing adnexal masses. Chinese J Ultrasound Med 2013;29(6):527-530.
19. Rams N, Muñoz R, Soler C, Parra J. Results of the Gynecologic Imaging Reporting and Data System classification in adnexal masses. Prog Obstet Gynecol 2015 Mar;58(3):125-129.
20. Orozco Fernández R, Peces Rama A, Llanos Llanos MC, Martinez Mendoza A, Machado Linde F, Nieto Diaz A. Clinical application of the gynecologic imaging reporting and data system (GI-RADS) for the evaluation of adnexal masses. SM J Gynecol Obstet 2015 Jan;1(2):1009-1012.