Ontario
Cancer Care Ontario
Action Cancer Ontario

Evidence-Based Series 4-15 Version 2 ARCHIVED

A Quality Initiative of the
Program in Evidence-Based Care (PEBC), Cancer Care Ontario (CCO)

Management of a Suspicious Adnexal Mass

Members of the Gynecology Cancer Disease Site Group

An assessment conducted in November 2020 ARCHIVED Evidence-based Series (EBS) 4-15 Version 2. This means that the document will no longer be maintained but may still be useful for academic or other information purposes. The PEBC has a formal and standardized process to ensure the currency of each document (PEBC Assessment & Review Protocol).

EBS 4-15 Version 2 is comprised of 4 sections. You can access the summary and full report here:

https://www.cancercareontario.ca/en/guidelines-advice/types-of-cancer/466

Section 1: Guideline Recommendations (ENDORSED)
Section 2: Evidentiary Base
Section 3: Guideline Development Methods and External Review Process
Section 4: Document Assessment and Review

September 9, 2016

For further information about this series and all PEBC reports, including previous versions, please contact the PEBC office at:
Phone: 905-527-4322, ext. 42842 Fax: 905-526-6775 E-mail: ccopgi@mcmaster.ca

PEBC Report Citation (Vancouver Style): Dodge J, Covens A, Lacchetti C, Elit L, Le T, Devries-Aboud M, et al. Management of a suspicious adnexal mass. Le T, Coakley N, reviewers. Toronto (ON): Cancer Care Ontario; 2011 Jul 07 [Archived Nov 2020] Program in Evidence-based Care Evidence-Based Series No.: 4-15 Version 2 ARCHIVED.

Journal Citations (Vancouver Style): Dodge JE, Covens AL, Lacchetti C, Elit LM, Le T, Devries-Aboud M, et al. Management of a suspicious adnexal mass: a clinical practice guideline. Curr Oncol. 2012 Aug;19(4):e244-57.

Covens AL, Dodge JE, Lacchetti C, Elit LM, Le T, Devries-Aboud M, et al. Surgical management of a suspicious adnexal mass: a systematic review. Gynecol Oncol. 2012 Jul;126(1):149-56.
Dodge JE, Covens AL, Lacchetti C, Elit LM, Le T, Devries-Aboud M, et al. Preoperative identification of a suspicious adnexal mass: a systematic review and meta-analysis. Gynecol Oncol. 2012 Jul;126(1):157-66.

Guideline Report History

GUIDELINE VERSION	SYSTEMATIC REVIEW		PUBLICATIONS	NOTES AND KEY CHANGES
	SEARCH DATES	DATA		
Original Version July 2011	2004-2009	Full report	Web and print Publications	NA
Current Version 2 September 2016	2009-2016	New data found in section 4: Document Assessment and Review	Updated web publications	2011 Recommendations are **ENDORSED**

Ontario
Cancer Care Ontario
Action Cancer Ontario

Evidence-Based Series 4-15 Version 2: Section 1

Management of a Suspicious Adnexal Mass: Guideline Recommendations

J Dodge, A Covens, C Lacchetti, L Elit, T Le, M Devries-Aboud, M Fung Kee Fung and the Gynecology Cancer Disease Site Group

A Quality Initiative of the
Program in Evidence-Based Care (PEBC), Cancer Care Ontario (CCO)

September 9, 2016

These guideline recommendations have been ENDORSED, which means that the recommendations are still current and relevant for decision making.

Please see Section 4: Document Review Summary and Tool for a summary of updated evidence published between 2009 and 2016, and for details on how this Clinical Practice Guideline was ENDORSED

QUESTIONS

1. What is the optimal strategy for preoperative identification of the adnexal mass suspicious for ovarian cancer?
2. What is the most appropriate surgical procedure for a woman who presents with an adnexal mass suspicious for ovarian cancer?

TARGET POPULATION

The target population of this guideline is adult women presenting with a suspicious adnexal mass, either symptomatic or asymptomatic.

INTENDED USERS

This guideline is targeted for clinicians managing the care of women with a suspicious adnexal mass, specifically general gynecologists and gynecologic oncologists.

RECOMMENDATIONS

Identification of an Adnexal Mass Suspicious for Ovarian Cancer

- **Sonography, particularly three-dimensional (3D) sonography, magnetic resonance imaging (MRI), and computerized tomography (CT) imaging are each recommended for differentiating malignant from benign ovarian masses. However, the working group offers the following further recommendations, based on their expert consensus opinion and the consideration of availability, access, and harm:**
  - **Transvaginal sonography should be the first modality of choice, where technically feasible, in patients with a suspicious, isolated ovarian mass.**
  - **MRI is the most appropriate test to help clarify the malignant potential in patients where ultrasound may be unreliable.**
  - **CT is most useful in cases where metastatic disease is suspected or needs to be ruled out.**

  Key Evidence
  - The diagnostic performance of each diagnostic technology was compared and contrasted based on the summary data on sensitivity and specificity obtained from the meta-analysis.
  - A meta-analysis of six cohort studies that investigated 3D sonography (1-6) indicated an enhanced sensitivity of 93.5% and specificity of 91.5% with 3D technology (Section 2, Figure 2A).
  - A meta-analysis of 22 cohort studies with 24 data sets that investigated the effectiveness of MRI in the diagnosis of adnexal masses (7-28) found an overall sensitivity of 91.9% and specificity of 88.4% (Section 2, Figure 2A).
  - A meta-analysis of seven studies with eight data sets considering CT technology (2,10,12,14,22,29-30) yielded an overall sensitivity of 87.2% and specificity of 84.0% (Section 2, Figure 2A).

- **Evaluation of an adnexal mass by Doppler technology alone is not recommended. Doppler technology should be combined with a morphological assessment.**

  Key Evidence
  - This recommendation is based on the results of several meta-analyses on Doppler indices, but not direct comparisons between them. Rather, the summary data from these meta-analyses were inspected and reasonable sensitivities and specificities were noted.
  - A meta-analysis of the resistance index (RI) included 35 cohort studies (2,5,17,30-61) with 42 data sets and yielded an overall sensitivity of 77.2% and specificity of 89.8% (Section 2, Figure 2C).
  - A meta-analysis of 21 cohort studies with 22 data sets that evaluated the Pulsatility Index (PI) found an overall sensitivity of 80.6% and specificity of 79.9% (Section 2, Figure 2C).
  - A meta-analysis of the peak systolic velocity (PSV) included seven cohort studies (32-33,37,42,50-51,62) and found an overall sensitivity of 80.0% and specificity of 84.2% (Section 2, Figure 2C).

  Qualifying Statement
  - Assessment of an adnexal mass by colour Doppler technology, using the RI, PI, and PSV indices, was neither as sensitive nor specific as simple ultrasonography. Furthermore, because of the overlap of vascular parameters between malignant and benign masses, a firm diagnosis based on Doppler evaluation alone can be problematic.

- **Ultrasound-based morphological scoring systems can be used to differentiate benign from malignant adnexal masses. These systems are based on specific ultrasound parameters, each with several scores according to determined features. All evaluated scoring systems were found to have an acceptable level of sensitivity and specificity; therefore, the choice of scoring system may be made based on clinician preference. More information on the characteristics of these scoring systems can be found in Appendix 1.**

  Key Evidence

  - Direct comparisons between ultrasound-based morphological scoring systems were not performed in this review. Instead, the assessment was based on summary data on sensitivity and specificity obtained from the meta-analyses conducted. The meta-analyses found summary sensitivities ranging from 83.5% (Finkler) (63) to 91% (DePriest) (64) and specificities ranging from 63% (Lerner) (65) to 85.9% (Ferrazzi) (66) (Section 2, Figure 2B).
  - The Risk of Malignancy Index (RMI) (67) is a clinical prediction rule that includes CA-125 and menopausal status, in addition to ultrasound-based morphology. In a meta-analysis of data from the 13 RMI studies (67-79), with 15 data sets, employing a cutoff of 200 to be indicative of malignancy, the summary sensitivity and specificity were 79.2% and 91.7%, respectively (Section 2, Figure 2B). RMI2 (74) and RMI3 (80) are newer versions of this tool, with comparable levels of sensitivity and specificity. The choice of version of RMI should be based on clinician preference.

  Qualifying Statement

  - Ultrasound diagnostic criteria using a set of simple rules to distinguish between benign and malignant masses, and the IOTA (International Ovarian Tumour Analysis) predictive adnexal model had been extensively studied with acceptable sensitivity and specificity. This can serve as potential alternative diagnostic strategy to the RMI score. (81-83)

- **As a stand-alone modality, serum CA-125 is not recommended for distinguishing between benign and malignant adnexal masses.**

  Key Evidence

  - This recommendation is based on a meta-analysis of 49 cohort studies (17,31,35,39,52,62-63,70,72,77-78,84-121) and two case-control studies (122-123) with a total of 52 data sets that found, at a threshold of 35 U/mL, an overall sensitivity of 78.7% and specificity of 77.9% (Section 2, Figure 2D).

  Qualifying Statement

  - Elevated serum CA-125 levels have been reported in a variety of benign conditions. Because the incidence of ovarian cancer relative to benign gynecologic conditions is lower in premenopausal women, CA-125 values are of limited use in this population (124). CA-125 levels are elevated in only 50% of early stage ovarian cancers (125). Caution should be used in interpreting values in such patients.

- **Frozen section for the intraoperative diagnosis of a suspicious adnexal mass is recommended in settings where availability and patient preferences allow.**

Key Evidence

- This recommendation is based on a meta-analysis of frozen section diagnoses that included 15 cohort studies (7,126-139) and yielded an overall sensitivity of 89.2% and specificity of 97.9% (Section 2, Figure 2D).

Surgical Procedures for an Adnexal Mass Suspicious for Malignancy

- **Comprehensive surgical staging with lymphadenectomy is recommended for the surgical management of patients with early-stage ovarian cancer to improve survival.**

Key Evidence

- This recommendation is based on the results of five retrospective cohort studies (140-144).
- Two large population-based studies (140,141) found that surgical staging with lymphadenectomy was associated with improved three-year ($p<0.001$) (141) and five-year disease specific survival ($p<0.001$) (140) compared to staging procedures without lymphadenectomy.
- Oksefjell et al (142) reported a statistically significant improvement in five-year overall survival rates in patients undergoing a lymphadenectomy versus those that did not (87% versus [vs.] 64%, respectively; $p=0.02$).
- Survival analyses performed by both Skirnisdottir et al (143) and Hornung et al (144) also demonstrated a statistically significant benefit in disease-free survival ($p=0.004$ and $p=0.0007$, respectively) for patients undergoing a lymphadenectomy versus those that did not.
- Hornung and colleagues (144) also considered overall survival and reported a statistically significant difference ($p=0.0008$) in the two patient groups in favour of the patients undergoing a lymphadenectomy.
- One randomized controlled trial (RCT) (145) was identified and reported no statistically significant effect of lymphadenectomy on progression-free (hazard ratio [HR]= 0.72; 95% confidence interval [CI], 0.46 to 1.14) or overall (HR=0.85; 95% CI, 0.49 to 1.47) survival. However, this study was underpowered to detect a difference in survival, the study's secondary outcome. Rather, the sample-size calculation for this RCT was undertaken to detect a difference in prevalence of lymph node positivity. It was deemed inadequate to inform the recommendation.

- **Laparoscopy is a reasonable alternative to laparotomy, provided appropriate surgery and/or staging can be done. The choice between laparoscopy and laparotomy should be based on patient and clinician preferences. Discussion with a gynecologic oncologist is recommended.**

Key Evidence

- This recommendation is based on the results of six retrospective cohort studies (146-151).
- In the three studies (146,147,150) that considered patients with early epithelial ovarian cancer, no statistical difference in survival rates was detected between patients undergoing a laparoscopy versus laparotomy.
- In the management of patients with early borderline ovarian tumours, Romangnolo et al (149), Park et al (151) and Desfeux et al (148) found that a laparoscopic versus laparotomic surgical approach did not appear to influence survival rates.

- **Fertility-preserving surgery is an acceptable alternative to more extensive surgery in patients with low-malignant potential (LMP) tumours and those with well-differentiated surgically staged 1 ovarian cancer. Discussion with a gynecologic oncologist is recommended.**

Key Evidence

- This recommendation is based on two cohort studies that compared the impact of conservative fertility-sparing surgeries versus more radical surgical approaches. Yinon et al (152) specifically compared rates of recurrence in 40 patients who underwent unilateral salpingo-oophorectomy versus 22 patients who underwent cystectomy only. No statistical difference in recurrence rates was detected (27.5% vs. 22.7%, respectively; p=0.8). Similarly, in a larger study of 360 women with LMP tumours, Park et al (151) found no difference in disease-free survival between patients who underwent radical or fertility-sparing surgery (p=0.651).

Qualifying Statement

- The Gynecology Cancer Disease Site Group (DSG) acknowledges that, despite definitions and criteria, it is unrealistic to expect that 100% of ovarian cancers will be identified as suspicious preoperatively. Pathology remains the gold standard.

RELATED GUIDELINES

- PEBC EBS 4-4 2004 Update. *Management Options for Women with a Hereditary Predisposition to Ovarian Cancer*.
- PEBC EBS 4-6a 2004 Update. *Screening Postmenopausal Women for Ovarian Cancer*
- PEBC EBS 4-6b 2004. *Screening High Risk Women for Ovarian Cancer*

Funding

The PEBC is a provincial initiative of Cancer Care Ontario supported by the Ontario Ministry of Health and Long-Term Care through Cancer Care Ontario. All work produced by the PEBC is editorially independent from its funding source.

PEBC Contact Information

For further information about this series and all PEBC reports, including previous versions, please contact the PEBC office at:

Phone: 905-527-4322, ext. 42842 Fax: 905-526-6775 E-mail: ccopgi@mcmaster.ca

REFERENCES

1. Geomini P, Kluivers K, Moret E, Bremer GL, Kruitwagen R, Mol B. Evaluation of adnexal masses with three-dimensional ultrasonography. Obstet Gynecol. 2006;108(5):1167-75.
2. Laban M, Metawee H, Elyan A, Kamal M, Kamel M, Mansour G. Three-dimensional ultrasound and three-dimensional power Doppler in the assessment of ovarian tumors. Int J Gynaecol Obstet. 2007;99(3):201-5.
3. Alcazar JL, Galan MJ, Garcia-Manero M, Guerriero S. Three-dimensional sonographic morphologic assessment in complex adnexal masses: preliminary experience. J Ultrasound Med. 2003;22(3):249-54.
4. Kurjak A, Kupesic S. Three dimensional ultrasound and power doppler in assessment of uterine and ovarian angiogenesis: a prospective study. Croat Med J. 1999;40(3):413-20.
5. Kurjak A, Kupesic S, Sparac V, Kosuta D. Three-dimensional ultrasonographic and power Doppler characterization of ovarian lesions. Ultrasound Obstet Gynecol. 2000;16(4):365-71.
6. Alcázar JL, Castillo G. Comparison of 2-dimensional and 3-dimensional power-Doppler imaging in complex adnexal masses for the prediction of ovarian cancer. Am J Obstet Gynecol. 2005;192(3):807-12.
7. Bazot M, Nassar-Slaba J, Thomassin-Naggara I, Cortez A, Uzan S, Darai E. MR imaging compared with intraoperative frozen-section examination for the diagnosis of adnexal tumors; correlation with final histology. Eur Radiol. 2006;16(12):2687-99.
8. Chen M, Wang WC, Zhou C, Zhou NN, Cai K, Yang ZH, et al. Differentiation between malignant and benign ovarian tumors by magnetic resonance imaging. Chin Med Sci J. 2006;21(4):270-5.
9. Guerra A, Cunha TM, Felix A. Magnetic resonance evaluation of adnexal masses. Acta Radiol. 2008;49(6):700-9.
10. Tsili AC, Tsampoulas C, Argyropoulou M, Navrozoglou I, Alamanos Y, Paraskevaidis E, et al. Comparative evaluation of multidetector CT and MR imaging in the differentiation of adnexal masses. Eur Radiol. 2008;18(5):1049-57.
11. Booth SJ, Turnbull LW, Poole DR, Richmond I. The accurate staging of ovarian cancer using 3T magnetic resonance imaging--a realistic option. BJOG. 2008;115(7):894-901.
12. Umemoto M, Shiota M, Shimono T, Hoshiai H. Preoperative diagnosis of ovarian tumors, focusing on the solid area based on diagnostic imaging. J Obstet Gynaecol Res. 2006;32(2):195-201.
13. Scoutt LM, McCarthy SM, Lange R, Bourque A, Schwartz PE. MR evaluation of clinically suspected adnexal masses. J Comput Assist Tomogr. 1994;18(4):609-18.
14. Buist MR, Golding RP, Burger CW, Vermorken JB, Kenemans P, Schutter EM, et al. Comparative evaluation of diagnostic methods in ovarian carcinoma with emphasis on CT and MRI. Gynecol Oncol. 1994;52(2):191-8.
15. Fenchel S, Grab D, Nuessle K, Kotzerke J, Rieber A, Kreienberg R, et al. Asymptomatic adnexal masses: correlation of FDG PET and histopathologic findings. Radiology. 2002;223(3):780-8.
16. Grab D, Flock F, Stohr I, Nussle K, Rieber A, Fenchel S, et al. Classification of asymptomatic adnexal masses by ultrasound, magnetic resonance imaging, and positron emission tomography. Gynecol Oncol. 2000;77(3):454-9.
17. Hata K, Hata T, Manabe A, Sugimura K, Kitao M. A critical evaluation of transvaginal Doppler studies, transvaginal sonography, magnetic resonance imaging, and CA 125 in detecting ovarian cancer. Obstet Gynecol. 1992;80(6):922-6.

18. Hricak H, Chen M, Coakley FV, Kinkel K, Yu KK, Sica G, et al. Complex adnexal masses: detection and characterization with MR imaging--multivariate analysis. Radiology. 2000;214(1):39-46.
19. Huber S, Medl M, Baumann L, Czembirek H. Value of ultrasound and magnetic resonance imaging in the preoperative evaluation of suspected ovarian masses. Anticancer Res. 2002;22(4):2501-7.
20. Jain KA, Friedman DL, Pettinger TW, Alagappan R, Jeffrey RB Jr, Sommer FG. Adnexal masses: comparison of specificity of endovaginal US and pelvic MR imaging. Radiology. 1993;186(3):697-704.
21. Komatsu T, Konishi I, Mandai M, Togashi K, Kawakami S, Konishi J, et al. Adnexal masses: transvaginal US and gadolinium-enhanced MR imaging assessment of intratumoral structure. Radiology. 1996;198(1):109-15.
22. Kurtz AB, Tsimikas JV, Tempany CM, Hamper UM, Arger PH, Bree RL, et al. Diagnosis and staging of ovarian cancer: comparative values of Doppler and conventional US, CT, and MR imaging correlated with surgery and histopathologic analysis--report of the Radiology Diagnostic Oncology Group. Radiology. 1999;212(1):19-27.
23. Medl M, Kulenkampff KJ, Stiskal M, Peters-Engl C, Leodolter S, Czembirek H. Magnetic resonance imaging in the preoperative evaluation of suspected ovarian masses. Anticancer Res. 1995;15(3):1123-5.
24. Reuter M, Steffens J, Schuppler U, Luttges J, Muhle C, Brinkmann G, et al. Critical evaluation of the specificity of MRI and TVUS for differentiation of malignant from benign adnexal lesions. Eur Radiol. 1998;8(1):39-44.
25. Yamashita Y, Torashima M, Hatanaka Y, Harada M, Higashida Y, Takahashi M, et al. Adnexal masses: accuracy of characterization with transvaginal US and precontrast and postcontrast MR imaging. Radiology. 1995;194(2):557-65.
26. Kawahara K, Yoshida Y, Kurokawa T, Suzuki Y, Nagahara K, Tsuchida T, et al. Evaluation of positron emission tomography with tracer 18-fluorodeoxyglucose in addition to magnetic resonance imaging in the diagnosis of ovarian cancer in selected women after ultrasonography. J Comput Assist Tomogr. 2004;28(4):505-16.
27. Sohaib SA, Mills TD, Sahdev A, Webb JA, Vantrappen PO, Jacobs IJ, et al. The role of magnetic resonance imaging and ultrasound in patients with adnexal masses. Clin Radiol. 2005;60(3):340-8.
28. Rieber A, Nussle K, Stohr I, Grab D, Fenchel S, Kreienberg R, et al. Preoperative Diagnosis of Ovarian Tumors with MR Imaging: Comparison with Transvaginal Sonography, Positron Emission Tomography, and Histologic Findings. Am J Roentgenol. 2001;177(1):123-9.
29. Kitajima K, Murakami K, Yamasaki E, Kaji Y, Fukasawa I, Inaba N, et al. Diagnostic accuracy of integrated FDG-PET/contrast-enhanced CT in staging ovarian cancer: comparison with enhanced CT. Eur J Nucl Med Mol Imaging. 2008;35(10):1912-20.
30. Lin JY, Angel C, DuBeshter B, Walsh CJ. Diagnoses after laparotomy for a mass in the pelvic area in women. Surg Gynecol Obstet. 1993;176(4):333-8.
31. Alcazar JL, Errasti T, Zornoza A, Minguez JA, Galan MJ. Transvaginal color Doppler ultrasonography and CA-125 in suspicious adnexal masses. Int J Gynaecol Obstet. 1999;66(3):255-61.
32. Alcazar JL, Ruiz-Perez ML, T E. Transvaginal color Doppler sonography in adnexal masses: which parameter performs best? Ultrasound Obstet Gynecol. 1996;8(2):114-9.
33. Alcazar JL, Lopez-Garcia G. Transvaginal color Doppler assessment of venous flow in adnexal masses. Ultrasound Obstet Gynecol. 2001;17(5):434-8.

34. Anandakumar C, Chew S, Wong YC, Chia D, Ratnam SS. Role of transvaginal ultrasound color flow imaging and Doppler waveform analysis in differentiating between benign and malignant ovarian tumors. Ultrasound Obstet Gynecol. 1996;7(4):280-4.
35. Berlanda N, Ferrari MM, Mezzopane R, Boero V, Grijuela B, Ferrazzi E, et al. Impact of a multiparameter, ultrasound-based triage on surgical management of adnexal masses. Ultrasound Obstet Gynecol. 2002;20(2):181-5.
36. Bromley B, Goodman H, Benacerraf BR. Comparison between sonographic morphology and Doppler waveform for the diagnosis of ovarian malignancy. Obstet Gynecol. 1994;83(3):434-7.
37. Buy JN, Ghossain MA, Hugol D, Hassen K, Sciot C, Truc JB, et al. Characterization of adnexal masses: combination of color Doppler and conventional sonography compared with spectral Doppler analysis alone and conventional sonography alone. AJR Am J Roentgenol. 1996;166(2):385-93.
38. Carter JR, Lau M, Fowler JM, Carlson JW, Carson LF, Twiggs LB. Blood flow characteristics of ovarian tumors: implications for ovarian cancer screening. Am J Obstet Gynecol. 1995;172(3):901-7.
39. Chou CY, Chang CH, Yao BL, Kuo HC. Color Doppler ultrasonography and serum CA 125 in the differentiation of benign and malignant ovarian tumors. J Clin Ultrasound. 1994;22(8):491-6.
40. Franchi M, Beretta P, Ghezzi F, Zanaboni F, Goddi A, Salvatore S. Diagnosis of pelvic masses with transabdominal color Doppler, CA 125 and ultrasonography. Acta Obstet Gynecol Scand. 1995;74(9):734-9.
41. Guerriero S, Ajossa S, Piras S, Gerada M, Floris S, Garau N, et al. Three-dimensional quantification of tumor vascularity as a tertiary test after B-mode and power Doppler evaluation for detection of ovarian cancer. J Ultrasound Med. 2007;26(10):1271-8.
42. Hata K, Hata T, Kitao M. Intratumoral peak systolic velocity as a new possible predictor for detection of adnexal malignancy. Am J Obstet Gynecol. 1995;172(5):1496-500.
43. Jain KA. Prospective evaluation of adnexal masses with endovaginal gray-scale and duplex and color Doppler US: correlation with pathologic findings. Radiology. 1994;191(1):63-7.
44. Kurjak A, Zalud I, Alfirevic Z. Evaluation of adnexal masses with transvaginal color ultrasound. J Ultrasound Med. 1991;10(6):295-7.
45. Kurjak A, Predanic M. New scoring system for prediction of ovarian malignancy based on transvaginal color Doppler sonography. J Ultrasound Med. 1992;11(12):631-8.
46. Leeners B, Schild RL, Funk A, Hauptmann S, Kemp B, Schroder W, et al. Colour Doppler sonography improves the pre-operative diagnosis of ovarian tumours made using conventional transvaginal sonography. Eur J Obstet Gynecol Reprod Biol. 1996;64(1):79-85.
47. Marchesini ACdS, Magrio FAA, Berezowski AT, Neto OBP, Nogueira AA, Candido dos Reis FJ. A critical analysis of Doppler velocimetry in the differential diagnosis of malignant and benign ovarian masses. J Womens Health (Larchmt). 2008;17(1):97-102.
48. Marret H, Sauget S, Giraudeau B, Brewer M, Ranger-Moore J, Body G, et al. Contrast-enhanced sonography helps in discrimination of benign from malignant adnexal masses. J Ultrasound Med. 2004;23(12):1629-39; quiz 41-42.
49. Merce LT, Caballero RA, Barco MJ, Bau S, Lopez G. B-mode, utero-ovarian and intratumoural transvaginal colour Doppler ultrasonography for differential diagnosis of ovarian tumours. Eur J Obstet Gynecol Reprod Biol. 1998;76(1):97-107.

50. Mousavi AS, Borna S, Moeinoddini S. Estimation of probability of malignancy using a logistic model combining color Doppler ultrasonography, serum CA125 level in women with a pelvic mass. Int J Gynecol Cancer. 2006;16 Suppl 1:92-8.
51. Prompeler HJ, Madjar H, Sauerbrei W. Classification of adnexal tumors by transvaginal color Doppler. Gynecol Oncol. 1996;61(3):354-63.
52. Schneider VL, Schneider A, Reed KL, Hatch KD. Comparison of Doppler with two-dimensional sonography and CA 125 for prediction of malignancy of pelvic masses. Obstet Gynecol. 1993;81(6):983-8.
53. Stein SM, Laifer-Narin S, Johnson MB. Differentiation of benign and malignant adnexal masses: relative value of gray-scale, color Doppler, and spectral Doppler sonography. AJR Am J Roentgenol. 1995;164 (2):381-6.
54. Takac I. Analysis of blood flow in adnexal tumors by using color Doppler imaging and pulsed spectral analysis. Ultrasound Med Biol. 1998;24(8):1137-41.
55. Tekay A, Jouppila P. Validity of pulsatility and resistance indices in classification of adnexal tumors with transvaginal color Doppler ultrasound. Ultrasound Obstet Gynecol. 1992;2(5):338-44.
56. Tepper R, Lerner-Geva L, Altaras MM, Goldberger S, Ben-Baruch G, Markov S, et al. Transvaginal color flow imaging in the diagnosis of ovarian tumors. J Ultrasound Med. 1995;14(10):731-4.
57. Timor-Tritsch LE, Lerner JP, Monteagudo A, Santos R. Transvaginal ultrasonographic characterization of ovarian masses by means of color flow-directed Doppler measurements and a morphologic scoring system. Am J Obstet Gynecol. 1993;168(3 Pt 1):909-13.
58. Valentin L. Comparison of Lerner score, Doppler ultrasound examination, and their combination for discrimination between benign and malignant adnexal masses. Ultrasound Obstet Gynecol. 2000;15(2):143-7.
59. Valentin L. Pattern recognition of pelvic masses by gray-scale ultrasound imaging: the contribution of Doppler ultrasound. Ultrasound Obstet Gynecol. 1999;14(5):338-47.
60. Wu CC, Lee CN, Chen TM, Lai JI, Hsieh CY, Hsieh FJ. Factors contributing to the accuracy in diagnosing ovarian malignancy by color Doppler ultrasound. Obstet Gynecol. 1994;84(4):605-8.
61. Zanetta G, Vergani P, Lissoni A. Color Doppler ultrasound in the preoperative assessment of adnexal masses. Acta Obstet Gynecol Scand. 1994;73(8):637-41.
62. Hillaby K, Aslam N, Salim R, Lawrence A, Raju KS, Jurkovic D. The value of detection of normal ovarian tissue (the 'ovarian crescent sign') in the differential diagnosis of adnexal masses. Ultrasound Obstet Gynecol. 2004;23(1):63-7.
63. Finkler NJ, Benacerraf B, Lavin PT, Wojciechowski C, Knapp RC. Comparison of serum CA 125, clinical impression, and ultrasound in the preoperative evaluation of ovarian masses. Obstet Gynecol. 1988;72(4):659-64.
64. DePriest PD, Shenson D, Fried A, Hunter JE, Andrews SJ, Gallion HH, et al. A morphology index based on sonographic findings in ovarian cancer. Gynecol Oncol. 1993;51(1):7-11.
65. 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(1 Pt 1):81-5.
66. 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(3):192-7.

67. Jacobs I, Oram D, Fairbanks J, Turner J, Frost C, Grudzinskas JG. A risk of malignancy index incorporating CA 125, ultrasound and menopausal status for the accurate preoperative diagnosis of ovarian cancer. Br J Obstet Gynaecol. 1990;97(10):922-9.
68. Bensaid C, Le Frere Belda MA, Metzger U, Larousserie F, Clement D, Chatellier G, et al. Performance of laparoscopy in identifying malignant ovarian cysts. Surg Endosc. 2006;20(9):1410-4.
69. Van Holsbeke C, Van Calster B, Valentin L, Testa AC, Ferrazzi E, Dimou I, et al. External validation of mathematical models to distinguish between benign and malignant adnexal tumors: a multicenter study by the International Ovarian Tumor Analysis Group. Clin Cancer Res. 2007;13(15 Pt 1):4440-7.
70. Asif N, Sattar A, Dawood MM, Rafi T, Aamir M, Anwar M. Pre-operative evaluation of ovarian mass: risk of malignancy index. J Coll Physicians Surg Pak. 2004;14(3):128-31.
71. Davies AP, Jacobs I, Woolas R, Fish A, Oram D. The adnexal mass: benign or malignant? Evaluation of a risk of malignancy index. Br J Obstet Gynaecol. 1993;100(10):927-31.
72. Manjunath AP, Pratapkumar, Sujatha K, Vani R. Comparison of three risk of malignancy indices in evaluation of pelvic masses. Gynecol Oncol. 2001;81(2):225-9.
73. Morgante G, la Marca A, Ditto A, De Leo V. Comparison of two malignancy risk indices based on serum CA125, ultrasound score and menopausal status in the diagnosis of ovarian masses. Br J Obstet Gynaecol. 1999;106(6):524-7.
74. Tingulstad S, Hagen B, Skjeldestad FE, Onsrud M, Kiserud T, Halvorsen T, et al. Evaluation of a risk of malignancy index based on serum CA125, ultrasound findings and menopausal status in the pre-operative diagnosis of pelvic masses. Br J Obstet Gynaecol. 1996;103(8):826-31.
75. Aslam N, Tailor A, Lawton F, Carr J, Savvas M, Jurkovic D. Prospective evaluation of three different models for the pre-operative diagnosis of ovarian cancer. Br J Obstet Gynaecol. 2000;107(11):1347-53.
76. Mol BW, Boll D, De Kanter M, Heintz AP, Sijmons EA, Oei SG, et al. Distinguishing the benign and malignant adnexal mass: an external validation of prognostic models. Gynecol Oncol. 2001;80(2):162-7.
77. Timmerman D, Bourne TH, Tailor A, Collins WP, Verrelst H, Vandenberghe K, et al. A comparison of methods for preoperative discrimination between malignant and benign adnexal masses: the development of a new logistic regression model. Am J Obstet Gynecol. 1999;181(1):57-65.
78. Engelen MJA, Bongaerts AHH, Sluiter WJ, de Haan HH, Bogchelman DH, Tenvergert EM, et al. Distinguishing benign and malignant pelvic masses: the value of different diagnostic methods in everyday clinical practice. Eur J Obstet Gynecol Reprod Biol. 2008;136(1):94-101.
79. Obeidat BR, Amarin ZO, Latimer JA, Crawford RA. Risk of malignancy index in the preoperative evaluation of pelvic masses. Int J Gynecol Obstet. 2004;85(3):255-8.
80. Tingulstad S, Hagen B, Skjeldestad FE, Halvorsen T, Nustad K, Onsrud M. The risk-of-malignancy index to evaluate potential ovarian cancers in local hospitals. Obstet Gynecol. 1999;93(3):448-52.
81. Van Calster B, Van Hoorde K, Froyman W, Kauser J, Landolfo C,Anthoulakis, et al. Practical guidance for applying the ADNEX model from the IOTA group to discriminate between different subtypes of adnexal tumors. Facts Views Vis OBGYN. 2015;7(1):32-41.
82. Timmerman D, Van Calster B, Testa A, Savelli L, Fischerova d, Froyman W, et al. Predicting the risk of malignancy in adnexal masses based on the Simle Rules from the International Ovarian Tumor Analysis group. Am J Obstet Gynecol. 2016;214(4):424-437.

83. Timmerman D, Testa A, Bourne T, Ameye L, Jurkovic D, Van Holsbeke C, et al. Simple ultrasound-based rules for the diagnosis of ovarian cancer. Ultrasound Obstet Gynecol. 2008;31:681-690.
84. Patsner B, Mann WJ. The value of preoperative serum CA 125 levels in patients with a pelvic mass. Am J Obstet Gynecol. 1988;159(4):873-6.
85. Adonakis G, Paraskevaidis E, Tsiga S. A combined approach for the early detection of ovarian cancer in a symptomatic women. Eur J Obstet Gynecol Reprod Biol. 1996;65(2):221-5.
86. Benjapibal M, Neungton C. Pre-operative prediction of serum CA125 level in women with ovarian masses. J Med Assoc Thai. 2007;90(10):1986-91.
87. Romagnolo C, Trivella G, Bonacina M, Fornale M, Maggino T, Ferrazzi E. Preoperative diagnosis of 221 consecutive ovarian masses: scoring system and expert evaluation. Eur J Gynaecol Oncol. 2006;27(5):487-9.
88. Zhang Z, Yu Y, Xu F, Berchuck A, van Haaften-Day C, Havrilesky LJ, et al. Combining multiple serum tumor markers improves detection of stage I epithelial ovarian cancer. Gynecol Oncol. 2007;107(3):526-31.
89. Milojkovic M, Hrgovic Z, Hrgovic I, Jonat W, Maass N, Bukovic D. Significance of CA 125 serum level in discrimination between benign and malignant masses in the pelvis. Arch Gynecol Obstet. 2004;269(3):176-80.
90. Erdogan N, Ozcelik B, Serin IS, Akgun M, Ozturk F. Doppler ultrasound assessment and serum cancer antigen 125 in the diagnosis of ovarian tumors. Int J Gynaecol Obstet. 2005;91(2):146-50.
91. Balbi GC, Musone R, Menditto A, Balbi F, Corcioni C, Calabria G, et al. Women with a pelvic mass: indicators of malignancy. Eur J Gynaecol Oncol. 2001;22(6):459-62.
92. Chalas E, Welshinger M, Engellener W, Chumas J, Barbieri R, Mann WJ. The clinical significance of thrombocytosis in women presenting with a pelvic mass. Am J Obstet Gynecol. 1992;166(3):974-7.
93. Chen DX, Schwartz PE, Li XG, Yang Z. Evaluation of CA 125 levels in differentiating malignant from benign tumors in patients with pelvic masses. Obstet Gynecol. 1988;72(1):23-7.
94. Dowd JR QM, Rome R, et al. . Women with a pelvic mass--when to perform an ultrasound. . Aust N Z J Obstet Gynaecol. 1993;33(4):404-7.
95. Einhorn N, Bast RC Jr, Knapp RC, Tjernberg B, Zurawski VR Jr. Preoperative evaluation of serum CA 125 levels in patients with primary epithelial ovarian cancer. Obstet Gynecol. 1986;67(3):414-6.
96. Gadducci A, Capriello P, Bartolini T, Barale E, Cappelli N, Facchini V, et al. The association of ultrasonography and CA-125 test in the preoperative evaluation of ovarian carcinoma. Eur J Gynaecol Oncol. 1988;9(5):373-6.
97. Gadducci A, Ferdeghini M, Prontera C, Moretti L, Mariani G, Bianchi R, et al. The concomitant determination of different tumor markers in patients with epithelial ovarian cancer and benign ovarian masses: relevance for differential diagnosis. Gynecol Oncol. 1992;44(2):147-54.
98. Gadducci A, Ferdeghini M, Rispoli G, Prontera C, Bianchi R, Fioretti P. Comparison of tumor-associated trypsin inhibitor (TATI) with CA125 as a marker for diagnosis and monitoring of epithelial ovarian cancer. Scand J Clin Lab Invest Suppl. 1991;207:19-24.
99. Hogdall EV, Hogdall CK, Tingulstad S, Hagen B, Nustad K, Xu FJ, et al. Predictive values of serum tumour markers tetranectin, OVX1, CASA and CA125 in patients with a pelvic mass. Int J Cancer. 2000;89(6):519-23.

100. Woolas RP, Conaway MR, Xu F, Jacobs IJ, Yu Y, Daly L, et al. Combinations of multiple serum markers are superior to individual assays for discriminating malignant from benign pelvic masses. Gynecol Oncol. 1995;59(1):111-6.
101. Zhang Z, Barnhill SD, Zhang H, Xu F, Yu Y, Jacobs I, et al. Combination of multiple serum markers using an artificial neural network to improve specificity in discriminating malignant from benign pelvic masses. Gynecol Oncol. 1999;73(1):56-61.
102. Hurteau JA, Woolas RP, Jacobs IJ, Oram DC, Kurman CC, Rubin LA, et al. Soluble interleukin-2 receptor alpha is elevated in sera of patients with benign ovarian neoplasms and epithelial ovarian cancer. Cancer. 1995;76(9):1615-20.
103. Kawai M, Kikkawa F, Ishikawa H, Tamakoshi K, Maeda O, Hasegawa N, et al. Differential diagnosis of ovarian tumors by transvaginal color-pulse Doppler sonography. Gynecol Oncol. 1994;54(2):209-14.
104. Maggino T, Gadducci A, D'Addario V, Pecorelli S, Lissoni A, Stella M, et al. Prospective multicenter study on CA 125 in postmenopausal pelvic masses. Gynecol Oncol. 1994;54(2):117-23.
105. Malkasian GD Jr, Knapp RC, Lavin PT, Zurawski VR Jr, Podratz KC, Stanhope CR, et al. Preoperative evaluation of serum CA 125 levels in premenopausal and postmenopausal patients with pelvic masses: discrimination of benign from malignant disease. Am J Obstet Gynecol. 1988;159(2):341-6.
106. Mancuso A, De Vivo A, Triolo O, Irato S. The role of transvaginal ultrasonography and serum CA 125 assay combined with age and hormonal state in the differential diagnosis of pelvic masses. Eur J Gynaecol Oncol. 2004;25(2):207-10.
107. O'Connell GJ RE, Murphy KJ, et al. ;. Predictive value of CA 125 for ovarian carcinoma in patients presenting with pelvic masses. . Obstet Gynecol. 1987;70 (6):930-2.
108. Pyrgiotis E, Salamalekis E, Loghis C, Dima C, Zourlas PA. CA 125 in preoperative evaluation of pelvic masses. Eur J Gynaecol Oncol. 1993;14(4):279-82.
109. Roman LD, Muderspach LI, Stein SM, Laifer-Narin S, Groshen S, Morrow CP. Pelvic examination, tumor marker level, and gray-scale and Doppler sonography in the prediction of pelvic cancer. Obstet Gynecol. 1997;89(4):493-500.
110. Schutter EM, Davelaar EM, van Kamp GJ, Verstraeten RA, Kenemans P, Verheijen RH. The differential diagnostic potential of a panel of tumor markers (CA 125, CA 15-3, and CA 72-4 antigens) in patients with a pelvic mass. Am J Obstet Gynecol. 2002;187(2):385-92.
111. Schutter EM, Kenemans P, Sohn C, Kristen P, Crombach G, Westermann R, et al. Diagnostic value of pelvic examination, ultrasound, and serum CA 125 in postmenopausal women with a pelvic mass. An international multicenter study. Cancer. 1994;74(4):1398-406.
112. Schutter EM, Sohn C, Kristen P, Mobus V, Crombach G, Kaufmann M, et al. Estimation of probability of malignancy using a logistic model combining physical examination, ultrasound, serum CA 125, and serum CA 72-4 in postmenopausal women with a pelvic mass: an international multicenter study. Gynecol Oncol. 1998;69(1):56-63.
113. Sengoku K, Satoh T, Saitoh S, Abe M, Ishikawa M. Evaluation of transvaginal color Doppler sonography, transvaginal sonography and CA 125 for prediction of ovarian malignancy. Int J Gynaecol Obstet. 1994;46(1):39-43.
114. Smikle CB, Lunt CC, Hankins GD. Clinical predictors in the evaluation of a pelvic mass. Mil Med. 1995;160(5):233-5.
115. Soper JT, Hunter VJ, Daly L, Tanner M, Creasman WT, Bast RC Jr. Preoperative serum tumor-associated antigen levels in women with pelvic masses. Obstet Gynecol. 1990;75(2):249-54.

116. Tay SK, Chua EK. Correlation of serum, urinary and salivary CA 125 levels in patients with adnexal masses. Ann Acad Med Singapore. 1994;23(3):311-4.
117. Tian J, Zhang J, Jiao L, Li Y, Cao L. A prospective study of Tc-99m MIBI in the differential diagnosis of pelvic masses in female patients. Clin Nucl Med. 2000;25(8):614-8.
118. Torres JC, Derchain SF, Faundes A, Gontijo RC, Martinez EZ, Andrade LA. Risk-of-malignancy index in preoperative evaluation of clinically restricted ovarian cancer. Sao Paulo Med J. 2002;120(3):72-6.
119. Vasilev SA, Schlaerth JB, Campeau J, Morrow CP. Serum CA 125 levels in preoperative evaluation of pelvic masses. Obstet Gynecol. 1988;71(5):751-6.
120. Wakahara F, Kikkawa F, Nawa A, Tamakoshi K, Ino K, Maeda O, et al. Diagnostic efficacy of tumor markers, sonography, and intraoperative frozen section for ovarian tumors. Gynecol Obstet Invest. 2001;52(3):147-52.
121. Weiner Z, Thaler I, Beck D, Rottem S, Deutsch M, Brandes JM. Differentiating malignant from benign ovarian tumors with transvaginal color flow imaging. Obstet Gynecol. 1992;79(2):159-62.
122. El-Shalakany A, Abou-Talib Y, Shalaby HS, Sallam M. Preoperative serum inhibin levels in patients with ovarian tumors. J Obstet Gynaecol Res. 2004;30(2):155-61.
123. Nakae M, Iwamoto I, Fujino T, Maehata Y, Togami S-i, Yoshinaga M, et al. Preoperative plasma osteopontin level as a biomarker complementary to carbohydrate antigen 125 in predicting ovarian cancer. J Obstet Gynaecol Res. 2006;32(3):309-14.
124. Myers E, Bastian L, Havrilesky L, Kulasingam S, Terplan M, Cline K, et al. Management of Adnexal Mass. Rockville (MD): Agency for Healthcare Research and Quality; 2006 Feb. Report No.: Evidence Report/Technology Assessment No.: 130. AHRQ Publication No.: 06-E004. Contract No.: 290-02-0025.
125. National Institutes of Health (NIH) Consensus Development Panel on Ovarian Cancer. NIH consensus conference, ovarian cancer: screening, treatment and follow-up. JAMA. 1995;273:491-7.
126. Naik R, Cross P, Lopes A, Godfrey K, Hatem MH. "True" versus "apparent" stage I epithelial ovarian cancer: value of frozen section analysis. Int J Gynecol Cancer. 2006;16 Suppl 1:41-6.
127. Boriboonhirunsarn D, Sermboon A. Accuracy of frozen section in the diagnosis of malignant ovarian tumor. J Obstet Gynaecol Res. 2004;30(5):394-9.
128. Brun J-L, Cortez A, Rouzier R, Callard P, Bazot M, Uzan S, et al. Factors influencing the use and accuracy of frozen section diagnosis of epithelial ovarian tumors. Am J Obstet Gynecol. 2008;199(3):244.e1-7.
129. Wootipoom V, Dechsukhum C, Hanprasertpong J, Lim A. Accuracy of intraoperative frozen section in diagnosis of ovarian tumors. J Med Assoc Thai. 2006;89(5):577-82.
130. Yarandi F, Eftekhar Z, Izadi-Mood N, Shojaei H. Accuracy of intraoperative frozen section in the diagnosis of ovarian tumors. Aust NZ J Obstet Gynaecol. 2008;48(4):438-41.
131. Ghaemmaghami F, Fakour F, Karimi Zarchi M, Behtash N, Modares Gilani M, Mousavi A, et al. Clinical assessment, gross examination, frozen section of ovarian masses: do patients benefit? Arch Gynecol Obstet. 2008;278(3):209-13.
132. Wasinghon P, Suthippintawong C, Tuipae S. The accuracy of intraoperative frozen sections in the diagnosis of ovarian tumors. J Med Assoc Thai. 2008;91(12):1791-5.
133. Ilvan S, Ramazanoglu R, Ulker Akyildiz E, Calay Z, Bese T, Oruc N. The accuracy of frozen section (intraoperative consultation) in the diagnosis of ovarian masses. Gynecol Oncol. 2005;97(2):395-9.

134. Stewart CJR, Brennan BA, Hammond IG, Leung YC, McCartney AJ. Intraoperative assessment of ovarian tumors: a 5-year review with assessment of discrepant diagnostic cases. Int J Gynecol Pathol. 2006;25(3):216-22.
135. Geomini P, Zuurendonk LD, Bremer GL, de Graaff J, Kruitwagen R, Mol BW. The impact of size of the adnexal mass on the accuracy of frozen section diagnosis. Gynecol Oncol. 2005;99(2):362-6.
136. Fanfani F, Zannoni GF, Fagotti A, Gagliardi ML, Masciullo V, Testa AC, et al. Importance of a specialized pathologist for the examination of frozen sections of adnexal masses. Int J Gynecol Cancer. 2007;17(5):1034-9.
137. Tangjitgamol S, Jesadapatrakul S, Manusirivithaya S, Sheanakul C. Accuracy of frozen section in diagnosis of ovarian mass. Int J Gynecol Cancer. 2004;14(2):212-9.
138. Taskiran C, Erdem O, Onan A, Bozkurt N, Yaman-Tunc S, Ataoglu O, et al. The role of frozen section evaluation in the diagnosis of adnexal mass. Int J Gynecol Cancer. 2008;18(2):235-40.
139. Canis M, Mashiach R, Wattiez A, Botchorishvili R, Rabischong B, Jardon K, et al. Frozen section in laparoscopic management of macroscopically suspicious ovarian masses. J Am Assoc Gynecol Laparosc. 2004;11(3):365-9.
140. Chan JK, Munro EG, Cheung MK, Husain A, Teng NN, Berek JS, et al. Association of lymphadenectomy and survival in stage I ovarian cancer patients. Obstet Gynecol. 2007;109(1):12-9.
141. Chan J, Fuh K, Shin J, Cheung M, Powell C, Chen L, et al. The treatment and outcomes of early-stage epithelial ovarian cancer: have we made any progress? Br J Cancer. 2008;98(7):1191-6.
142. Oksefjell H, Sandstad B, Trope C. Is the watch and wait approach adequate after comprehensive surgical staging in invasive stage I epithelial ovarian cancer? The Norwegian Radium Hospital experience. Eur J Gynaecol Oncol. 2008;29(6):583-9.
143. Skirnisdottir I, Sorbe B. Lymph node sampling is of prognostic value in early stage epithelial ovarian carcinoma. Eur J Gynaecol Oncol. 2005;26(2):181-5.
144. Hornung R, Urs E, Serenella E, Edward W, Ursula S, Urs H, et al. Analysis of potential prognostic factors in 111 patients with ovarian cancer. Cancer Lett. 2004;206(1):97-106.
145. Maggioni A, Benedetti Panici P, Dell'Anna T, Landoni F, Lissoni A, Pellegrino A, et al. Randomised study of systematic lymphadenectomy in patients with epithelial ovarian cancer macroscopically confined to the pelvis. Br J Cancer. 2006;95(6):699-704.
146. Ghezzi F, Cromi A, Uccella S, Bergamini V, Tomera S, Franchi M, et al. Laparoscopy versus laparotomy for the surgical management of apparent early stage ovarian cancer. Gynecol Oncol. 2007;105(2):409-13.
147. Lecuru F, Desfeux P, Camatte S, Bissery A, Blanc B, Querleu D. Impact of initial surgical access on staging and survival of patients with stage I ovarian cancer. Int J Gynecol Cancer. 2006;16(1):87-94.
148. Desfeux P, Camatte S, Chatellier G, Blanc B, Querleu D, Lecuru F. Impact of surgical approach on the management of macroscopic early ovarian borderline tumors. Gynecol Oncol. 2005;98(3):390-5.
149. Romagnolo C, Gadducci A, Sartori E, Zola P, Maggino T. Management of borderline ovarian tumors: results of an Italian multicenter study. Gynecol Oncol. 2006;101(2):255-60.
150. Park JY, Bae J, Lim MC, Lim SY, Seo SS, Kang S, et al. Laparoscopic and laparotomic staging in stage I epithelial ovarian cancer: a comparison of feasibility and safety. Int J Gynecol Cancer. 2008;18(6):1202-9.

151. Park JY, Kim DY, Kim JH, Kim YM, Kim YT, Nam JH. Surgical management of borderline ovarian tumors: The role of fertility-sparing surgery. Gynecol Oncol. 2009;113(1):75-82.
152. Yinon Y, Beiner ME, Gotlieb WH, Korach Y, Perri T, Ben-Baruch G. Clinical outcome of cystectomy compared with unilateral salpingo-oophorectomy as fertility-sparing treatment of borderline ovarian tumors. Fertil Steril. 2007;88(2):479-84.
153. 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(1):70-6.

Appendix 1. Scoring systems for distinguishing benign from malignant adnexal masses.

Ultrasound-based morphological scoring systems

Table 1. Detailed description of ultrasound scoring systems (121).

Scoring system	**Score**				
Sassone et al., 1991 (153)					
Morphology	1	2	3	4	5
Inner wall structure	Smooth	Irregularities ≤ 3mm	Papillarities > 3 mm	Not applicable, mostly solid	-
Wall thickness(mm)	Thin (≤ 3)	Thick (> 3)	Not applicable, mostly solid	-	-
Septa (mm)	None	Thin (≤ 3)	Thick (> 3)		-
Echogenicity	Sonolucent	Low echogenicity	Low echogenicity with ochogenic core; mixed echogenicity	-	High echogenicity
DePriest et al., 1993 (64)					
Morphology	0	1	2	3	4
Cystic wall structure	Smooth (<3 mm thick)	Smooth (> 3 mm thick)	Papillary projection (< 3 mm)	Papillary projection (≥3 mm)	Predominately solid
Volume (cm3)	< 10	10-50	> 50-200	> 200-500	> 500
Septum structure	No septa	Thin septa (< 3 mm)	Thick septa (3 mm to 1 cm)	Solid area (≥ 1 cm)	Predominately solid
Ferrazzi et al., 1997 (66)					
Morphology	1	2	3	4	5
Wall	≤ 3 mm	> 3 mm	-	Irregular, mostly solid	Irregular, not applicable
Septa	None	≤ 3 mm	> 3 mm		
Vegetations	None	-	-	≤ 3 mm	> 3 mm
Echogenicity	Sonolucent	Low echogenicity	-	With echogenic areas	With heterogeneous echogenic areas, solid
Lerner et al., 1994 (65)					
Morphology	0	1	2	3	
Wall structure	Smooth or small irregularities <3 mm	-	- Solid or not applicable	Papillarities ≥ 3 mm	
Shadowing	Yes	No	-	-	
Septa	None or thin (<3 mm)	Thick (≥ 3 mm)	-	-	
Echogenicity	Sonolucent or low-level echo or echogenic core	-	-	Mixed or high	

Cutoffs suggestive of malignancy: Sassone: >9, DePriest: ≥5, Ferrazzi: >9. Lerner: ≥3

Table 2. Finkler ultrasound-based morphological scoring system (63).*

Clear cyst and smooth borders or fibroid (ovaries normal), or tubular cyst such as hydrosalpinx	1
Clear cyst with slightly irregular border; cyst with smooth walls but low-level echoes (i.e., endometrioma)	2
Cyst with low-level echoes with slightly irregular border but no nodularity (i.e. endometrioma); clear cyst in postmenopausal patient	3
Equivocal, nonspecific ultrasound appearance: solid ovarian enlargement or small cyst with irregular borders and internal echoes (hemorrhagic cyst or benign ovarian tumour)	4-6
Multiseptated or irregular cystic mass consistent in appearance with ovarian tumour (7 = less nodularity, 8-9 = more nodularity)	7-9
Pelvic mass as above, with ascites	10
*1 = benign, 10 = malignant, ≥7 indicative of probable malignancy	

Risk of Malignancy Index (RMI)

The RMI (67, 124), is a clinical prediction rule that calculates a numeric score based on the tumour marker CA-125, which may be elevated in the blood of some cancer patients, multiplied by a menopausal score and an ultrasound morphology score. The most common threshold for probability of malignancy is 200. Scores are calculated as follows:

RMI = U × M × CA-125

where ultrasound (transabdominal) is scored 1 point for each of the following characteristics: multilocular cyst, evidence of solid areas, evidence of metastases, presence of ascites, and bilateral lesions.

U = 0 for ultrasound score of 0
= 1 for ultrasound score of 1
= 3 for ultrasound score of ≥ 2

CA-125 = serum CA-125 in U/ml

And menopausal status is defined as:
M = 1 if premenopausal
= 3 if postmenopausal

RMI2 (74) is calculated in the same way as the original RMI, except that new weights were used for the ultrasound and menopause components:

U = 1 for ultrasound score of 0-1
= 4 for ultrasound score of ≥ 2

M = 1 if premenopausal
= 4 if postmenopausal

RMI3 (80) is a further refinement to the RMI and RMI2, using the same definitions, but with an adjustment to the ultrasound and menopause components:

U = 1 for ultrasound score of 0-1
= 3 for ultrasound score of ≥ 2

M = 1 if premenopausal
= 3 if postmenopausal

Ontario
Cancer Care Ontario
Action Cancer Ontario

Evidence-Based Series 4-15 Version 2: Section 2

Management of a Suspicious Adnexal Mass: Evidentiary Base

J Dodge, A Covens, C Lacchetti, L Elit, T Le, M Devries-Aboud, M Fung Kee Fung and the Gynecology Cancer Disease Site Group

A Quality Initiative of the
Program in Evidence-Based Care (PEBC), Cancer Care Ontario (CCO)

Report Date: July 7, 2011

These guideline recommendations have been ENDORSED, which means that the recommendations are still current and relevant for decision making.

Please see Section 4: Document Review Summary and Tool for a summary of updated evidence published between 2009 and 2016, and for details on how this Clinical Practice Guideline was ENDORSED

QUESTIONS

1. What is the optimal strategy for preoperative identification of the adnexal mass suspicious for ovarian cancer?
2. What is the most appropriate surgical procedure for a woman who presents with an adnexal mass suspicious for ovarian cancer?

INTRODUCTION

In Canada in 2010, an estimated 2600 new cases of ovarian cancer will be diagnosed and, of those cases, 1750 women will die, making ovarian cancer the seventh most prevalent form of cancer in Canadian women and their fifth leading cause of cancer death (1). Women with ovarian cancer typically have subtle, non-specific symptoms such as abdominal pain, bloating, changes in bowel frequency, and urinary and/or pelvic symptoms (2), making early detection difficult. Thus, the majority of ovarian cancer cases are diagnosed at an advanced stage when the cancer has spread outside the pelvis (3). Due to the late diagnosis of this disease, the five-year relative survival ratio for ovarian cancer in Canada is only 40% (1). Unfortunately, due to the low-positive predictive values of potential screening tests (CA-125 and ultrasound), there is currently no screening strategy for ovarian cancer (4).

Palpation using a bimanual pelvic examination or by radiological imaging (3) can identify an adnexal mass, which is defined as an enlarged lump near the uterus, usually in the ovary or fallopian tube. Adnexal masses include both benign (ovarian cysts, fibroids, and endometriomas) and malignant tumours.

There are numerous methods that have been tested in the preoperative identification of adnexal masses suspicious for malignancy. These methods include CA-125, transvaginal and transabdominal ultrasound, MRI, CT scans and the risk of malignancy index (RMI). However, the most appropriate identification method has yet to be determined, mainly due to poor positive and negative predictive values associated with each test in differentiating a benign from a malignant mass.

Once the diagnosis of ovarian cancer is confirmed, the patient must undergo surgical staging or debulking. However, which surgical staging and debulking procedure should be used to improve overall survival, progression-free survival, and quality of life in women with ovarian cancer is also unknown. The purpose of this document is to identify evidence that would inform optimal recommended protocols for the identification and surgical management of adnexal masses suspicious for malignancy.

METHODS

The evidence-based series (EBS) guidelines developed by the CCO PEBC use the methods of the Practice Guidelines Development Cycle (5). For this project, the core methodology used to develop the evidentiary base was an update of two previously published systematic reviews: the Agency for Healthcare Research and Quality (AHRQ) report, 2006 (3) and Australian Cancer Network (ACN) Clinical Practice Guideline, 2004 (6). Evidence was selected and reviewed by five members of the PEBC Gynecology Cancer DSG and one methodologist.

The systematic review is a convenient and up-to-date source of the best available evidence on the management of an adnexal mass suspicious for malignancy. The body of evidence in this review is primarily comprised of prospective and retrospective cohort studies. That evidence forms the basis of the recommendations developed by the Gynecology Cancer DSG and published in Section 1. The systematic review and companion recommendations are intended to promote evidence-based practice in Ontario, Canada. The PEBC is supported by the Ontario Ministry of Health and Long-Term Care through Cancer Care Ontario. All work produced by the PEBC is editorially independent from its funding source.

Literature Search Strategy

Environmental Scan

As a first step, an internet search of Canadian and international health organizations and the National Guidelines Clearinghouse (see Appendix 1 for full list) was conducted for existing guidelines and systematic reviews relevant to our research question. Guidelines were included if they were published since 1999 in English. This initial environmental scan yielded 11 practice guidelines; however, one guideline was excluded because the full guideline was available only in French, and another guideline was excluded because only the National Guidelines Clearinghouse summary was available. One evidence report/technology assessment and one clinical practice guideline identified through this environmental scan were deemed to be the most appropriate to answer the guideline questions. The 2006 AHRQ report (3) addressed the identification of an adnexal mass suspicious for malignancy question. The 2004 ACN Clinical Practice Guideline (6) addressed the surgical management of an adnexal mass suspicious for malignancy question.

Update Literature Search Strategy

The literature search from the AHRQ report was updated (Appendix 2) using MEDLINE (OVID: January 2004 through week 3, March 2009). In addition, as an exact search strategy for the Australian Cancer Network report was not available, an update of that literature search (Appendix 2) was approximated using the keywords provided in the report using MEDLINE (OVID: January 2004 through week 3, April 2009). This literature search combined disease-specific terms ('pelvic mass,' 'adnexal mass,' 'pelvic neoplasms,' 'ovarian cancer,' 'ovarian neoplasm,' 'ovarian carcinoma,' 'epithelial ovarian cancer,' 'borderline ovarian tumours' and 'tumours of low malignant potential') with surgical specific terms ('intraoperative pathological examination,' 'frozen section,' 'debulking surgery,' 'fertility sparing,' 'surgical staging,' 'bilateral salpingo-oophorectomy,' 'total hysterectomy,' 'node or nodal dissection,' 'surgical management,' 'treatment,' 'cytoreduction,' 'secondary cytoreduction,' 'interval cytoreduction,' 'laparotomy,' and laparoscopy') for all study designs.

Relevant articles and abstracts were selected and reviewed by two reviewers. The reference lists of included studies along with the personal reference lists of the guideline working group were searched for additional studies.

Study Selection Criteria

Articles were eligible for inclusion in this systematic review if they were systematic reviews, meta-analyses, clinical practice guidelines, randomized trials, or comparative cohort studies. Studies indentified in the update of the AHRQ report literature search were included based on the same inclusion criteria put forth in the AHRQ report (3).

For studies investigating single modality identification of an adnexal mass, the inclusion criteria were:

1) comparison of the test (e.g., bimanual pelvic exam or ultrasound, to histology or negative surgery
2) greater than 20 patients included in study
3) able to construct a 2-by-2 table, which compares the results of the diagnostic test with the definitive histological diagnosis.

For studies investigating the use of multi-modality scoring systems (i.e., RMI), the inclusion criteria were:

1) patients with suspicion of cancer
2) studies with scoring, risk score, combined modality approach
3) assesses predictive value of two or more variables using multivariable model
4) greater than 50 patients included in study.

Studies identified in the update of the Australian Cancer Network (6) guideline were based on the following selection criteria:

1) greater than 20 patients included in study
2) patients with an adnexal mass suspicious for early stage (I-II) malignancy,
3) two-armed (or greater) study design with a comparison of surgical procedures/techniques/approaches
4) report on at least one of the following outcomes: optimal surgery, overall survival, progression-free or disease-free survival, reduction in the number of surgeries, morbidity, adverse events, quality of life.

Synthesizing the Evidence

A bivariate, random-effects meta-regression model was used to produce summary estimates of sensitivity and specificity and to plot summary receiver operating characteristic (ROC) curves with 95% confidence regions. This model, described in detail elsewhere (7-9),

has several advantages over the standard summary ROC approach. Chief among these is the preservation of the two-dimensional nature of the data and the incorporation of any correlation that might exist between sensitivity and specificity (8). The model assumes that the logit sensitivities and specificities are normally distributed and makes use of the variance estimates to compute study weights (7). Heterogeneity in the current review was assessed visually. Given that between-study heterogeneity is widespread for measures of diagnostic accuracy (10), a random-effects model was used for all pooling. This bivariate, random-effects model takes into account the difference in precision by which sensitivity and specificity have been measured within and across studies, and it incorporates and estimates the amount of between-study variability (8). Statistical analyses were executed with the statistical software package STATA version 11 (11) using the metandi command. The outcomes of the meta-analyses were plotted as summary ROC curves and can be seen in Figures 2A-D.

The Gynecology Cancer DSG decided not to pool the surgical studies, but rather to present the results of each study individually in a descriptive fashion.

Quality Appraisal and Data Extraction

The Appraisal of Guidelines Research and Evaluation (AGREE) tool (12) was used to evaluate the quality of identified evidence-based guidelines. While all scoring domains of the AGREE tool were considered in the evaluation of guidelines, the Rigour of Development domain, describing the rigour of systematic methods in identifying and evaluating evidence, was considered to be most relevant in application for this systematic review. Systematic reviews and meta-analyses were assessed for quality using the AMSTAR tool (13). The quality of primary studies included assessments for study design, type of data collection, sampling method, and blinding.

RESULTS

Updated Literature Search Results

1. Identification of an adnexal mass literature

Four meta-analyses (14-17) and 67 primary studies (Table 1), pertaining to the identification of an adnexal mass suspicious for malignancy, met the inclusion criteria and are included in this review.

Table 1. New primary studies on identification of a suspicious adnexal mass.

	Study Design	Data Collection	Diagnostic Test
Alcazar, 2005 (18)	Cohort	prospective	3D US+Doppler
Bazot, 2006 (19)	Cohort	retrospective	MRI, Frozen section
Benjapibal, 2007 (20)	Cohort	prospective	CA-125
Bensaid, 2006 (21)	Cohort	prospective	RMI
Booth, 2008 (22)	Cohort	retrospective	MRI
Boriboonhirunsarn, 2004 (23)	Cohort	retrospective	Frozen section
Brun, 2008 (24)	Cohort	retrospective	Frozen section
Canis, 2004 (25)	Cohort	retrospective	Frozen section
Chen, 2006 (26)	Cohort	prospective	MRI
Choudhury, 2005 (27)	Cohort	prospective	2-D PD

Dai, 2008 (28)	Cohort	prospective	3-D PD
Daponte, 2007 (29)	Cohort	prospective	Ferrazzi
Dearking, 2007 (30)	Cohort	prospective	ACOG/SGO referral
El-shalakany, 2004 (31)	Case Control	prospective	CA-125
Engelen, 2008 (32)	Cohort	prospective	RMI, CA-125
Erdogan, 2005 (33)	Cohort	prospective	US+Doppler, CA-125
Exacoustos, 2005 (34)	Cohort	retrospective	2-D PD
Fanfani, 2007 (35)	Cohort	retrospective	Frozen section
Ferrazzi, 2005 (36)	Cohort	prospective	Ferrazzi
Geomini, 2005 (37)	Cohort	retrospective	Frozen section
Geomini, 2006 (38)	Cohort	prospective	US
Ghaemmaghami, 2008 (39)	Cohort	retrospective	Frozen section
Guerra, 2008 (40)	Cohort	retrospective	MRI
Guerriero, 2007 (41)	Cohort	prospective	RI, 3-D PD
Hata, 1999 (42)	Cohort	prospective	Sassone, US
Ilvan, 2005 (43)	Cohort	retrospective	Frozen section
Im, 2005 (44)	Cohort	retrospective	ACOG/SGO referral
Jokubkiene, 2007 (45)	Cohort	prospective	LR
Kitajima, 2008 (46)	Cohort	prospective	CT
Kurjak, 2000 (47)	Cohort	prospective	3-D PD
Laban, 2007 (48)	Cohort	prospective	US, RI, US+Doppler, 2-D PD, 3-D PD, 3-D US+Doppler, CT
Lee, 2005 (49)	Cohort	prospective	Lerner, DePriest
Leelahakorn, 2005 (50)	Cohort	prospective	Ferrazzi
Marchesini, 2008 (51)	Cohort	prospective	RI, PI, US+Doppler
Marret, 2005 (52)	Cohort	prospective	RI
Milojkovic, 2004 (53)	Cohort	retrospective	CA-125
Moszynski, 2006 (54)	Cohort	prospective	ANN
Mousavi, 2006 (55)	Cohort	prospective	RI, PSV
Naik, 2006 (56)	Cohort	retrospective	Frozen section
Nakae, 2006 (57)	Case Control	prospective	CA-125
Romagnolo, 2006 (58)	Cohort	prospective	CA-125
Sladkevicius, 2007 (59)	Cohort	prospective	LR
Sohaib, 2005 (60)	Cohort	prospective	US
Stewart, 2006 (61)	Cohort	retrospective	Frozen section
Szpurek, 2005 (62)	Cohort	NR	DePriest
Szpurek, 2005 (63)	Cohort	NR	ANN
Tan, 2007 (64)	Cohort	retrospective	US
Tangjitgamol, 2004 (65)	Cohort	retrospective	Frozen section
Taskiran, 2008 (66)	Cohort	retrospective	Frozen section
Tempe, 2006 (67)	Cohort	prospective	Sassone
Testa, 2007 (68)	Cohort	prospective	Doppler visualization

Timmerman, 2005 (69)	Cohort	prospective	LR
Tongsong, 2007 (70)	Cohort	prospective	US
Topuz, 2005 (71)	Cohort	prospective	Sassone
Tsili, 2008 (72)	Cohort	prospective	MRI, CT
Umemoto, 2006 (73)	Cohort	retrospective	MRI, CT
Valentin , 2006 (74)	Cohort	prospective	LR
Van Calster, 2007 (75)	Cohort	retrospective	2DPD
Van Calster, 2007 (76)	Cohort	prospective	2DPD
Van Holsbeke, 2007(77)	Cohort	prospective	ANN, LR
Wanapirak, 2006(78)	Cohort	prospective	US
Wasinghon, 2008(79)	Cohort	retrospective	Frozen section
Wilson, 2006(80)	Cohort	prospective	2DPD
Wootipoom, 2006(81)	Cohort	retrospective	Frozen section
Yarandi, 2008(82)	Cohort	retrospective	Frozen section
Yazbek, 2007(83)	Cohort	prospective	US
Zhang, 2007(84)	Cohort	retrospective	CA125

Abbreviations: US = ultrasound; MRI = magnetic resonance imaging; CA-125 = cancer antigen -125; RMI = Risk of Malignancy Index; 2-D PD = two-dimensional power Doppler; 3-D PD = three-dimensional power Doppler; ACOG = American College of Obstetricians and Gynecologists; SGO = Society of Gynecologic Oncologists; LR = logistic regression; CT= computed tomography; ANN = artificial neural network; PSV = peak systolic velocity; RI = resistance index; PI = pulsatility index

2. Surgical procedures literature

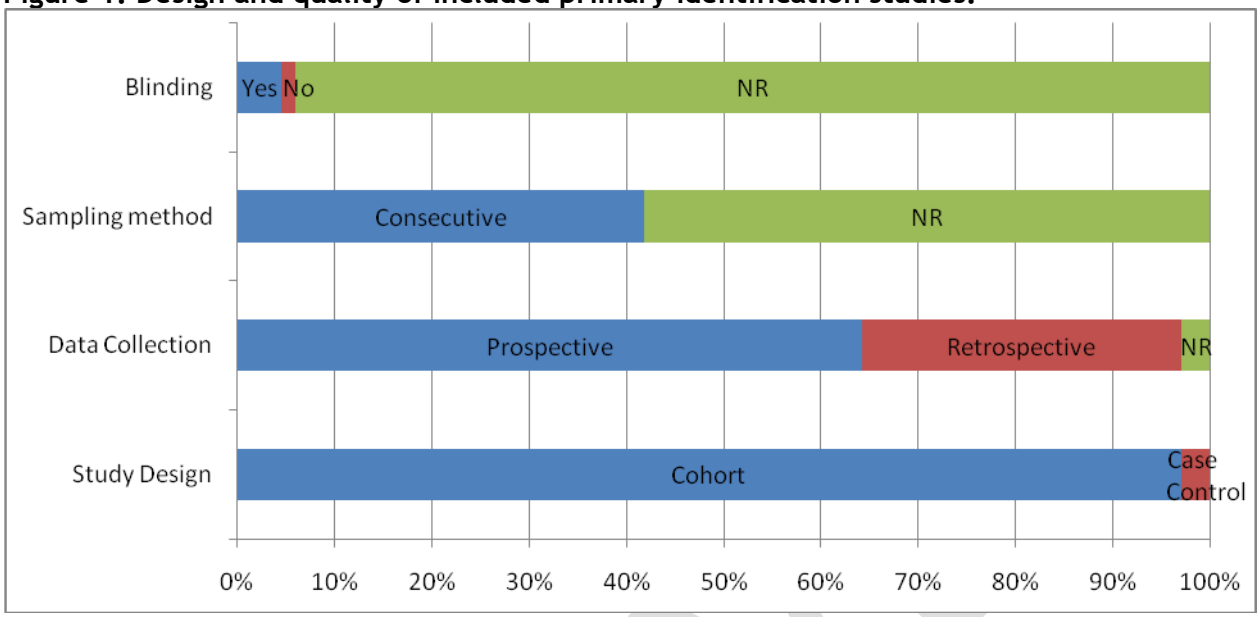
A total of 1809 articles were identified in the updated search for the most appropriate surgical procedure, of which 16 met the inclusion criteria (85-100).

Study Design and Quality

The ACN evidence-based guideline (6) earned 80.2% for the Rigour of Development domain of the AGREE tool. Since the role of AGREE in the assessment of health technology assessments has not yet been formally evaluated, the AHRQ report was not rated with AGREE but with AMSTAR for systematic reviews. It received an overall quality score of 90%. The meta-analyses by Geomini et al (14), Liu et al (15), Medieros et al (17), and Geomini et al (16) received scores of 73%, 55%, 82%, and 82% respectively.

Given that, at the time of inclusion, all patients either have ovarian cancer or they do not, diagnostic accuracy studies are, in principle, cross-sectional in nature (101). The diagnostic tests under study are intended to reduce the clinical uncertainty about their status (101). As such, the terms “cohort-type” or “case-control—type” accuracy studies have been proposed depending on the method of patient recruitment. Using such terminology, two case-control type studies and 65 cohort-type accuracy studies were included in this review (Figure 1). Data collection occurred prospectively in 64% of the studies, retrospectively in 33%, and not described in 3%. The sampling method was consecutive in 42% of the studies and not reported in the remaining 58%. Pathologists were blinded to the preoperative index test findings in only 4% of studies and not blinded in 1% of studies, with the remaining 94% of studies not describing any such blinding.

Figure 1. Design and quality of included primary identification studies.



In general, the quality of evidence for studies identified for the surgical question was poor. No meta-analyses or systematic reviews were identified. Only one of the 16 eligible surgical studies was an RCT (89), and another was designed as a "historically controlled trial" (95), defined as per the Cochrane Collaboration definition of study designs (102). The remaining 14 eligible studies were all retrospective in nature. As the quality of data collection in a retrospective study is often compromised, it is important to take the inherent limitations of such retrospective review designs into consideration. Furthermore, with the exception of 2 large SEER database reviews, the studies tended to be small and were likely underpowered to detect statistically significant differences in survival outcomes.

Outcomes

1. Identification of an adnexal mass suspicious for malignancy

Ultrasonography

Ultrasonography is the most common diagnostic imaging technique for the noninvasive assessment of adnexal masses and is believed to be both a reliable and reproducible method for preoperative discrimination between malignant and benign pelvic masses (103-104).

2D versus 3D

In addition to the four studies considered in the AHRQ report that analyzed both 2D and 3D ultrasonography, two new prospective studies were identified through the updated literature search (38,48). Analyzing the 2D ultrasonography data from the six studies using a bivariate random effects model yielded an overall sensitivity of 85.3% and specificity of 87.4% (Table 2, Figure 2A). When the data were pooled from the six 3D ultrasonography studies, the estimates improved to an overall sensitivity of 93.5% and specificity of 91.5% (Table 2, Figure 2A).

Table 2. Summary results of bivariate analysis of 2D and 3D ultrasound.

	Number of studies	Sensitivity (95% CI)	Specificity (95% CI)
2D Ultrasonography	6	85.3% (95% CI, 69.0-93.8%)	87.4% (95% CI, 75.0-94.1%)
3D Ultrasonography	6	93.5% (95% CI, 74.1-98.6%)	91.5% (95% CI, 80.0-96.6%)

Ultrasound Morphology

Morphological scoring systems, based on parameters observed from gray-scale sonography, were developed to overcome limitations with operator subjectivity and tumour variability (3). These scoring systems assign and then sum up the score for established sonographic variables. A predetermined specific cutoff value classifies the mass as either malignant or benign.

Sassone Scoring System

The model originally described by Sassone et al (105) is based on a weighted sum of the following four morphological features: inner wall structure, wall thickness, septa, and echogenicity. A score of greater than 9 is suggestive of malignancy. The AHRQ report identified and conducted a meta-analysis of 15 studies that explicitly used Sassone's criteria. The meta-analysis by Geomini et al (16) included 18 studies with cutoffs that ranged from 4 to 15, although 9 was the most common cutoff as in the original report. While our updated literature search did not identify any additional studies reporting on the Sassone et al model, we elected to conduct a meta-analysis of studies from both the AHRQ report and the Geomini et al meta-analysis that met our inclusion criteria, because neither report alone captured all the available data. A bivariate random effects model of these 22 studies generated a pooled sensitivity of 90.4% and specificity of 76.4%. Limiting our analysis to the 17 studies (42,62,67, 71,105-117) reporting a cutoff point of 9, yielded an overall sensitivity of 88.6% and specificity of 77.5% (Table 3, Figure 2B).

Lerner Scoring System

Lerner et al (118) devised a scoring system based on a modification to the Sassone model. The Lerner model differs from Sassone's in several ways, including weighted point value assignments, fewer point values per variable studied, the deletion of one variable found not to be significant (wall thickness), and the inclusion of a new variable called shadowing. A cutoff of 3 was determined to best differentiate benign from malignant masses. The AHRQ report did not separately evaluate the Lerner model. Geomini et al (16) reported eight studies that evaluated the Lerner model, in addition to the original, and found a pooled sensitivity of 90% and specificity of 63% (Table 3). Our updated literature search identified one additional study by Lee et al (49) where 137 masses were evaluated in 123 women, and a sensitivity of 82.1% and specificity of 69.7% were reported. We do not believe that the addition of this one study would change the overall estimates reported by Geomini et al and, thus, did not rerun the analysis.

DePriest Scoring System

The model described by DePriest et al (119) uses the weighted sum of cystic wall structure, volume, and septum structure. The score ranges from 0 to 12, with the most common cutoff set at 5. The AHRQ report identified six studies that made use of the DePriest scoring system. Geomini et al (16) evaluated 10 studies and reported a pooled sensitivity of 91% and specificity of 69% (Table 3). Our literature search identified an additional study (49) not included in either the AHRQ report or the meta-analysis by Geomini et al. In 123 women

with 137 masses, Lee and colleagues reported a sensitivity of 92.9% and specificity of 59.6%. We do not believe the addition of this one study to the Geomini et al meta-analysis would have changed the overall estimates and, therefore, did not repeat the analysis.

Ferrazzi Scoring System

Ferrazzi et al (109) developed their model to include four morphological features: wall structure, septa, vegetations and echogenicity. The weighted sum can range from 4 to 18, with 9 used as a cutoff to suggest malignancy. The AHRQ report identified three studies explicitly using the Ferrazzi et al criteria and the meta-analysis by Geomini et al (16) pooled seven studies. An updated search of the literature identified 1 additional study reporting on the Ferrazzi et al scoring system (58). Pooling the nine available, qualifying studies generated an overall sensitivity of 85.2% and specificity of 85.9% (Table 3, Figure 2B).

Finkler Scoring System

Finkler et al (120) developed a 10-point scoring system, where scores of 7 or more indicate malignancy. The AHRQ report identified three studies that evaluated the Finkler system, as did Geomini et al (16). However, Geomini et al found that the summary estimates varied widely in the ROC curve. Despite not having identified any additional studies, we felt it was worthwhile to combine all available data in overall summary estimates. Our bivariate random-effects analysis of the five studies (120-124) allowed in our model generated an overall sensitivity of 83.5% and specificity of 78.2% (Table 3, Figure 2B).

Other Scoring Systems

The meta-analysis by Geomini et al (16) also considered a model developed by Alcazar and Jurado that made use of the Sassone score as a variable in a logistic regression. Although the model was evaluated in four studies, there was too much variability in the sensitivity and specificity to pool the summary estimates. The AHRQ report identified 53 studies that evaluate ultrasonography in the assessment of adnexal mass morphology. These publications included unique, modified, or unclear scoring systems that did not fit into the other scoring system categories. While there was significant heterogeneity in the criteria used for diagnosis, the report did go ahead and pool sensitivity and specificity. The resulting summary estimates were 86% for sensitivity and 83% for specificity.

Table 3. Summary results of bivariate analysis of morphological scoring systems.

	Number of studies	Sensitivity (95% CI)	Specificity (95% CI)
Sassone Scoring System (cutoff 9)	17	88.6% (95% CI, 81.3-93.3%)	77.5% (95% CI, 70.0-83.6%)
Lerner Scoring System	9*	90% (95% CI, 87-98%)	63% (95% CI, 40-81%)
DePriest Scoring System	10*	91% (95% CI, 85-97%)	69% (95% CI, 60-78%)
Ferrazzi Scoring System	9	85.2% (95% CI, 76.4-91.1%)	85.9% (95% CI, 71.9-93.5%)
Finkler Scoring System	5	83.5% (95% CI, 73.1-90.4%)	78.2% (95% CI, 59.0-89.9%)

*From Geomini et al (16)

Explicit Scoring Systems

The AHRQ report considered scoring systems that included data combined from the following categories: 1) imaging, including ultrasound, CT, and MRI; 2) patient risk factors, such as age and menopausal status; and 3) laboratory data, primarily CA-125.

Risk of Malignancy Index (RMI)

The RMI, first published by Jacobs and colleagues (125), is based on scores from ultrasound (U), menopausal status (M), and CA-125 data in the following manner: RMI = U x M x CA-125. A cutoff of 200 was used to differentiate between malignant and benign masses in the original study. The AHRQ report (3) identified 11 studies in addition to the original that assessed the diagnostic performance of RMI. The meta-analysis by Geomini et al (16) identified 16 studies in which the RMI was evaluated but considered only the 13 studies that used the original cutoff of 200.

An updated literature search identified two new studies (21,32) that considered the diagnostic accuracy of RMI. When a bivariate random effects model was used to pool the data from the 13 studies, with 15 data sets, employing a cutoff of 200 to be indicative of malignancy, the summary sensitivity and specificity were 79.2% and 91.7%, respectively (Table 4, Figure 2B). When the analysis was extended to the 23 studies considering a cutoff level of 50, the summary estimates were 82.1% for sensitivity and 87.8% for specificity.

RMI2

Several years after the RMI was published, improvements to the model were attempted. Tingulstad et al (126) proposed the RMI2, which gives new weights to the ultrasound and menopause components of the original model. The same cutoff level of 200 was recommended for the RMI2 model.

The AHRQ report identified four studies in addition to the Tingulstand et al original, while Geomini et al (16) evaluated seven studies and, at a cutoff level of 200, reported pooled estimates for sensitivity and specificity of 79% and 81%, respectively (Table 4). The updated literature search indentified only one additional study evaluating the diagnostic accuracy of RMI2. In 194 women, Bensaid et al (21) reported a sensitivity of 92% (95% CI, 65 to 100%) and specificity of 80% (95% CI, 75 to 84%) at a cutoff of 125. Because data were not available at the recommended cutoff of 200, we did not repeat the analysis to include the data from Bensaid et al.

RMI3

In 1999, Tingulstad et al (127) suggested a further refinement to the two previous RMI models, again with suggested new weights given to the ultrasound and menopause components. Two hundred was once again the recommended cutoff level for this third model.

Both the AHRQ report and the meta-analysis by Geomini et al identified only one additional study (128) that evaluated the RMI3 and reported that this validation study found a sensitivity of 74% and specificity of 91%. These results were very similar to the original RMI3 report, which found a sensitivity of 71% and specificity of 92%. We did not identify any additional literature evaluating RMI3.

Tailor's Model

Tailor et al (129) developed a scoring system based on an artificial neural-network analysis that incorporates age, menopausal status, morphological features, and Doppler indices. The AHRQ report found four additional studies that evaluated this model but did not pool the data sets. They did note, however, that the subsequent studies each reported poorer performance than did the original publication. Geomini et al (16) found five publications that evaluated the original Tailor model. At a suggested cutoff of 50%, the pooled sensitivity and specificity of four data sets were 60% and 93%, respectively (Table 4). Our updated search of the literature did not identify any other studies that evaluated the diagnostic performance of the Tailor et al model.

Table 4. Summary results of bivariate analysis of explicit scoring systems.

	Number of studies	Sensitivity (95% CI)	Specificity (95% CI)
RMI (cutoff 200)	15	79.2% (95% CI, 73.6-83.9%)	91.7% (95% CI, 87.2-94.6%)
RMI 2 (cutoff 200)	7*	79% (95% CI, 71-87%)	81% (95% CI, 72-90%)
Tailor's Model	4*	60% (95% CI, 20-100%)	93% (95% CI, 82-100%)

*Geomini et al meta-analysis (16)

Artificial Neural Network (ANN) 1 and 2

Timmerman et al developed two artificial neural-network (ANN) models (130) in the late 1990s to predict the malignancy of an adnexal mass. Through complex modelling techniques, the ANN models investigate the possible existence of non-linear interactions or correlation between variables. The ANN1 model combines the predictive value of menopausal status, serum CA-125, presence/absence of papillary structures, and colour score, and uses a cutoff of 45%. The ANN2 model includes, in addition to menopausal status, serum CA-125, and papillary structures as in ANN1, the following set of morphological parameters: smoothness of internal walls, unilocularity, presence of ascites, and whether the mass is bilateral. A cutoff of 60% was used. Timmerman et al found the sensitivity and specificity of ANN1 to be 94% and 90%, respectively. ANN2 performed slightly better with a sensitivity of 96% and specificity of 94%. Two separate studies validated the models (77,115) and also reported a better performance with ANN2 over ANN1, although performance in subsequent replications was poorer. Mol et al (115) reported ANN2 to have a sensitivity and specificity of 90% and 46%, respectively, while Van Holsbeke et al (77) reported 98% and 34%, respectively. Our updated literature search did not identify any additional studies that validated the Timmerman ANN models. While other artificial neural networks have been developed over the years, external validation of these additional models is scarce, and they will not be discussed further.

Logistic Regression Models (LR) 1 and 2

The development of algorithms through statistical modeling that assess the probability of malignancy can also be used to distinguish malignant from benign masses preoperatively. One such logistic regression model (LR1) (131), developed in the late 1990s by Timmerman et al, included the following variables in the analysis: menopausal status, CA-125 level, presence of ≥1 papillary growth (>3mm in length), and a colour score indicative of tumour vascularity and blood flow. At a cutoff of 25%, the sensitivity of the LR1 was 95.9% and specificity was 87.1%. Two external validation studies (77,132) found lower estimates upon replication, where sensitivity was found to be 62% by Valentin et al and 78.1% by Van Holsbeke et al, and specificity was 79% in both studies.

A second logistic regression model (LR2) developed by the Timmerman group incorporates the same variables as ANN2 (i.e., menopausal status, serum CA-125, papillary structures, smoothness of internal walls, unilocularity, presence of ascites, and whether the mass is bilateral). Timmerman et al (130) reported, at a cutoff of 60%, a sensitivity and specificity of 95.9% and 85.5%, respectively. Again, at external validation, estimates were considerably lower (sensitivity 90% and 84% and specificity 86% and 75% in Mol et al (115) and Van Holsbeke et al (77), respectively).

A number of other logistic regression prediction models are found in the literature, but many are without external validation and will not be considered further in this document.

Doppler Sonography

Regular gray-scale sonography can be enhanced with Doppler measurements, which assess the direction of blood flow and its relative velocity. Colour Doppler imaging and pulsed Doppler spectral analysis enable evaluation of ovarian tumour blood flow, analysis of the distribution of blood vessels, and quantitative measurement of blood-flow velocity waveforms. These parameters increase the sensitivity and specificity of ultrasound evaluation of ovarian tumours (133).

2D Power Doppler (2D PD)

The updated search of the literature identified five qualifying studies that evaluated 2D PD technology in the preoperative discrimination between malignant and benign adnexal masses. The sensitivity in these studies ranged from 49% to 100%, while specificity ranged from 74% to 100% (Table 5).

Table 5. Performance of 2D PD from updated literature.

Study	Number of patients or masses	Sensitivity	Specificity
Exacoustos et al. 2005 (34)	452	48.5%	95.9%
Van Calster et al. 2007 (75)	809	88.4%	94.9%
Laban et al. 2007 (48)	50	100%	73.7%
Choudhury et al. 2005 (27)	40	90%	100%
Wilson et al. 2006 (80)	38	75.0%	76.7%

3D Power Doppler (3D PD)

The evaluation of 3D PD technology in the differential assessment of adnexal masses was considered in four qualifying studies identified through the updated literature search. The sensitivity in these studies ranged from 68% to 100%, while specificity ranged from 40% to 98% (Table 6).

Table 6. Performance of 3D PD from updated literature.

Study	Number of patients or masses	Sensitivity	Specificity
Dai et al. 2008 (28)	36	76.7%	50%
Guerriero et al. 2007 (41)	35	68%	40%
Kurjak et al. 2000 (47)	90	88.9%	97.5%
Laban et al. 2007 (48)	50	100%	73.7%

Resistance Index (RI)

Resistance Index (RI), the difference between peak systolic and maximum end-diastolic flow velocity divided by peak systolic flow velocity, is one of the most common flow criteria in colour Doppler scanning. The AHRQ report identified 32 articles that evaluated RI,

although one study (134) actually considered a morphology index, not RI, and was excluded from further analyses. Our updated literature search identified five new studies (41,48,51-52,55) that assessed the diagnostic accuracy of Doppler scanning using the RI parameter, but the data from Marret et al (52) was not included because this same data was previously published and already included. The analysis of the summary estimates for these four new studies along with the existing qualifying literature included 42 data sets and yielded an overall sensitivity of 77.2% and specificity of 89.8% (Table 7, Figure 2C).

Pulsatility Index

Pulsatility Index (PI) is defined as the difference between peak systolic and enddiastolic flow velocity, divided by the time-averaged flow velocity. The AHRQ report analyzed the PI from 20 studies, while the updated literature search identified only one additional study (51) looking at PI that met the inclusion criteria. When we analysed the 21 studies with 22 data sets, using a bivariate random effects model, we obtained an overall sensitivity of 80.6% and specificity of 79.9% (Table 7, Figure 2C).

Peak Systolic Velocity

The peak or maximum systolic velocity (PSV) is the maximum flow recorded in any visualized artery and, along with RI and PI, it is one of the most common flow criteria. Our analysis of the six studies included in the AHRQ report plus one additional identified paper (55) yielded an overall sensitivity of 80.0% and specificity of 84.2% (Table 7, Figure 2C).

Visualization

The AHRQ report considered Doppler studies that did not measure or calculate waveforms but, rather, looked at the presence of vascularity within the mass or a direct count of the vessels observed. The report identified 10 such studies with sensitivities that ranged from 77% to 100% and specificities that ranged from 30% to 94%. Pooling these studies resulted in an overall sensitivity of 88% and specificity of 78% (Table 7). Our updated literature search did not identify any additional studies investigating Doppler visualization.

Table 7. Summary results of bivariate analysis of Doppler indices.

	Number of studies	Sensitivity (95% CI)	Specificity (95% CI)
RI	35	77.2% (95% CI, 68.7-83.9%)	89.8% (95% CI, 85.6-92.8%)
PI	22	80.6% (95% CI, 74.9-85.2%)	79.9% (95% CI, 69.8-87.2%)
PSV	7	80.0% (95% CI, 67.7-88.5%)	84.2% (95% CI, 69.3-92.7%)
Visualization	10*	88% (80-92%)	78% (95% CI, 65-87%)

*From AHRQ report (3)

Combined Morphology and Doppler

2D Ultrasonography plus Doppler

A combination of morphological and vascular imaging was developed to try to improve the differentiation of malignant and benign adnexal masses. The AHRQ report identified nine studies that described such a combined modality. An additional three studies were identified through our updated search of the literature. When a bivariate random effect model was used to analyses all 12 studies, an overall sensitivity of 91.0% (95% CI, 84.8% to 94.8%) and specificity of 91.7% (95% CI, 81.1% to 96.6%) was obtained. (Figure 2C).

3D Ultrasonography plus Doppler

The updated search of the literature identified two qualifying studies that evaluated 3D ultrasonography plus Doppler technology in the differential assessment of adnexal masses (Table 8).

Table 8. Performance of 3D Ultrasound plus Doppler from updated literature.

Study	Number of patients or masses	Sensitivity	Specificity
Alcazar et al. 2005 (18)	60 pts, 69 masses	97.8%	79.2%
Laban et al. 2007 (48)	50 pts	100%	84.2%

Other Imaging Modalities

MRI

The AHRQ report identified 15 articles investigating the performance of MRI in the diagnosis of adnexal masses. An updated search of the literature since the AHRQ report was published identified a meta-analysis (15) evaluating, among other things, the MRI modality. Liu et al considered 10 studies reporting 13 data. In addition to this meta-analysis, six primary studies, not included in the Liu et al meta-analysis or the AHRQ report, were also identified during the updated literature search. Again, using a bivariate random effects model, we conducted a meta-analysis of the 15 studies from the AHRQ report, the one study exclusive to the Liu et al meta-analysis that qualified for the current analysis, and the six additional primary studies identified. This analysis of 22 studies with 24 data sets yielded an overall sensitivity of 91.9% and a specificity of 88.4% (Table 9, Figure 2A).

CT

The AHRQ report described three studies that looked at the performance of CT in diagnosing adnexal masses. Liu et al (15) evaluated a total of four studies, two of which also appeared in the AHRQ report and two that did not meet the AHRQ inclusion criteria. Our updated search identified an additional four studies evaluating CT in the diagnosis of adnexal masses. A bivariate random effects analysis of these four studies, along with the three studies (with four data sets) considered in the AHRQ report yielded an overall sensitivity of 87.2% and specificity of 84.0% (Table 9, Figure 2A).

PET

Three studies investigating positron emission tomography (PET) were evaluated in the AHRQ report. Our updated search identified two additional studies, both investigating the accuracy of the combined PET/CT modality. Since results for PET alone were not available, no further PET analyses were undertaken. The AHRQ report found a pooled sensitivity of 67% and specificity of 79% in the diagnosis of adnexal masses with PET technology (Table 9).

Table 9. Summary results of bivariate analysis of imaging modalities other than ultrasound.

	Number of studies	Sensitivity (95% CI)	Specificity (95% CI)
MRI	24	91.9% (95% CI, 88.8-94.1%)	88.4% (95% CI, 83.7-91.9%)
CT	8	87.2% (95% CI, 74.2-94.1%)	84.0% (95% CI, 66.6-93.3%)
PET	3*	67% (95% CI, 52-79%)	79% (95% CI, 70-85%)

*From AHRQ report (3)

Serum Marker

CA-125

The AHRQ report identified 66 studies that investigated the use of CA-125 serum marker in the evaluation of an adnexal mass. An updated literature search identified 11 new studies, with 14 data sets, investigating the use of CA-125 in the diagnosis of an adnexal mass that met our inclusion criteria. In keeping with the AHRQ report and the most commonly used threshold of 35 U/mL, we conducted a meta-analysis of the 51 studies, with 52 data sets, that used a threshold of 35 U/mL. Eight of these studies (20,31-33,53,57,58,135) were published since the AHRQ report. The addition of these newer studies to the analysis did little to change the AHRQ summary estimates. We calculated an overall sensitivity of 78.7% and a specificity of 77.9% (Table 10, Figure 2D). Since many of the studies reported a threshold other than 35 U/mL, we re-ran the analysis using 50 U/mL as the cutoff point. The analysis included 66 studies with 72 data sets, including 10 studies published since the AHRQ report. This yielded an overall sensitivity of 79.0% and specificity of 78.3%.

Frozen Section

While considered an intraoperative assessment rather than a preoperative one, frozen-section diagnosis can help to guide further surgical management of ovarian tumours. Accordingly, the accuracy of this technique is of great consequence and, as such, this method of diagnosing a suspicious adnexal mass was deemed valuable for consideration in this report.

The updated search of the literature identified two systematic reviews and 15 primary studies (19,23-25,35,37,39,43,56,61,65,66,79,81,82) published in or since 2004 that considered the diagnostic accuracy of frozen-section diagnosis and were not included in either systematic review. All 15 studies were retrospective cohort-like in design, with seven reporting the selection of consecutive patients. Only one study reported the blinding of pathologists from the final histopathologic diagnosis when interpreting results of frozen sections. One study specifically reported that pathologists were not blinded, and the remaining 13 studies made no reference to blinding status.

Geomini et al considered the accuracy of frozen-section diagnosis in a 2005 systematic review (14). The literature search ranged from 1966 to mid-2003, with 18 primary studies qualifying for inclusion. When borderline tumours were classified as malignant, the sensitivity of frozen-section diagnosis ranged from 65% to 97% and the specificity between 97% and 100%. Classifying borderline tumours as benign resulted in a sensitivity of 71% to 100% and a specificity of 98.3% to 100%.

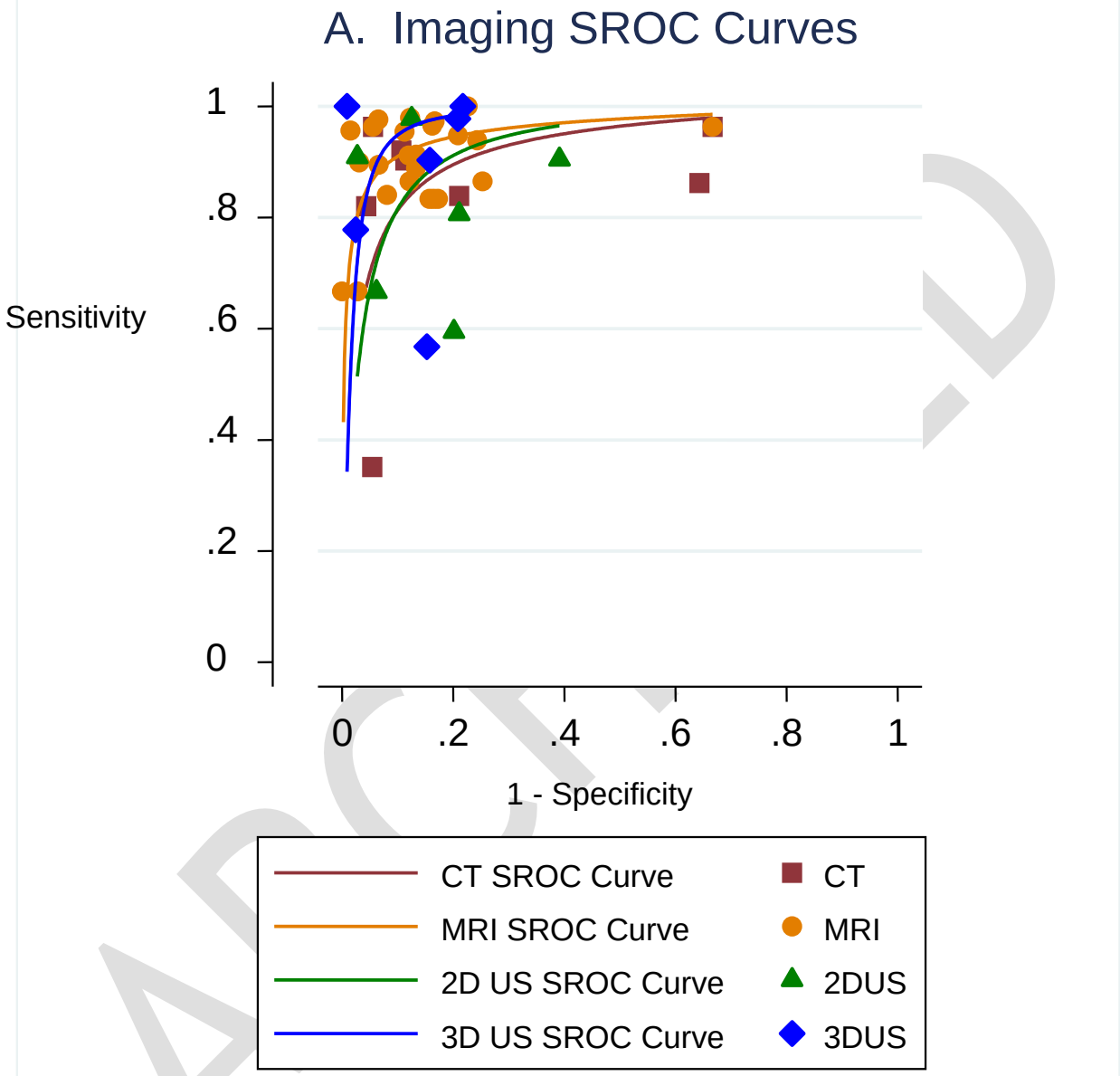
Medeiros and colleagues (17) also conducted a systematic review and meta-analysis on the accuracy of frozen-section analysis that included 14 primary studies. The literature search period in this review ranged from 1984 to the end of 2003. The pooled sensitivity and specificity of frozen-section diagnosis distinguishing between benign and borderline or malignant ovarian tumours was 99% (95% CI, 89 to 99%) and 88% (95% CI, 86 to 90%), respectively.

In our analysis of the 15 primary frozen section studies published since 2004, borderline tumours were considered malignant and counted as such in the 2x2 tables. Furthermore, any deferred cases reported were excluded from the 2x2 tables and, consequently, from the analysis. A bivariate random effects analysis of these 15 studies yielded an overall sensitivity of 89.2% (95% CI, 86.3 to 91.5%) and specificity of 97.9% (95% CI, 96.6 to 98.7%) (Table 10, Figure 2D).

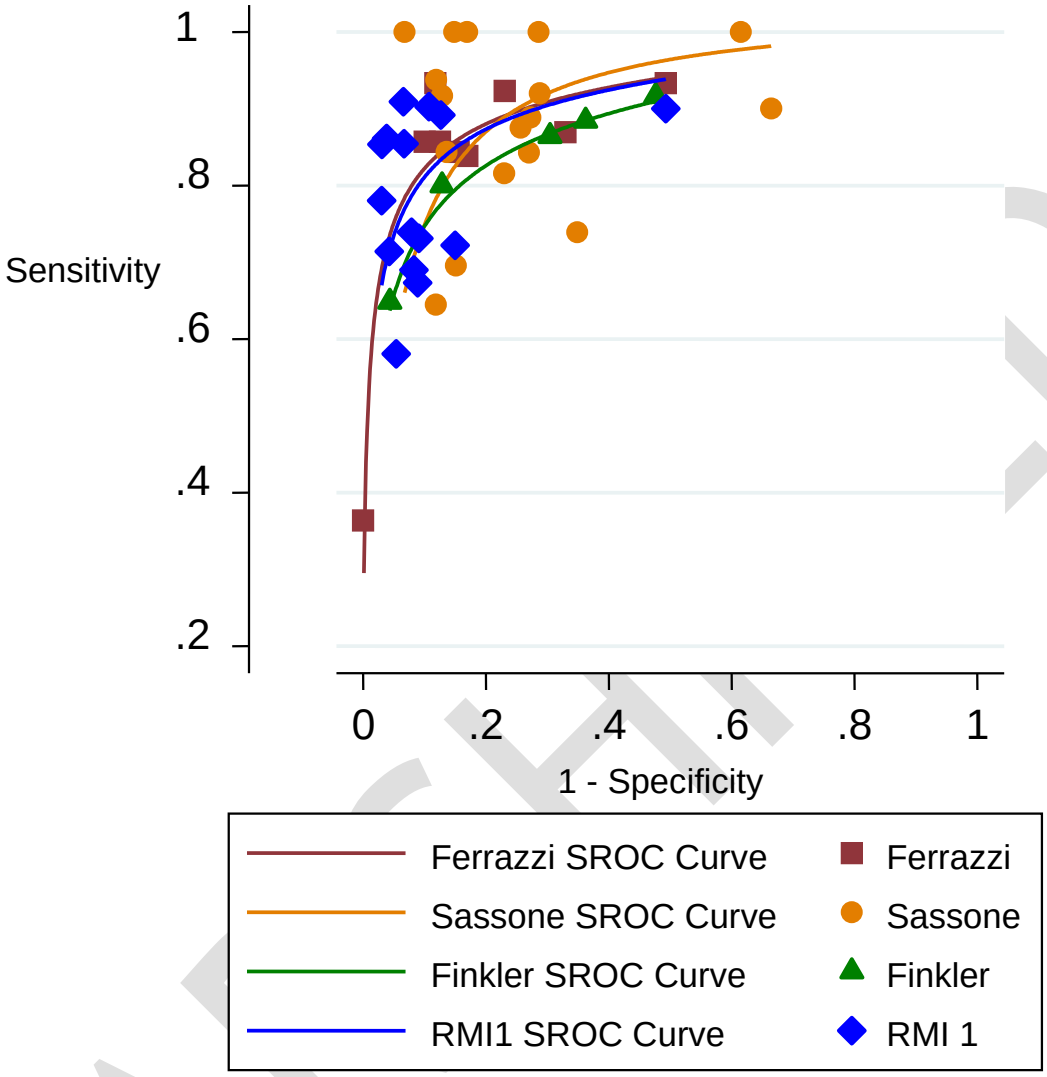
Table 10. Summary results of bivariate analysis of CA-125 and frozen section analysis.

	Number of studies	Sensitivity (95% CI)	Specificity (95% CI)
CA-125	51 studies (52 data sets)	78.7% (95% CI, 75.3-81.7%)	77.9% (95% CI, 73.2-82.0%)
Frozen Section	15	89.2% (86.3-91.5%)	97.9% (86.3-91.5%)

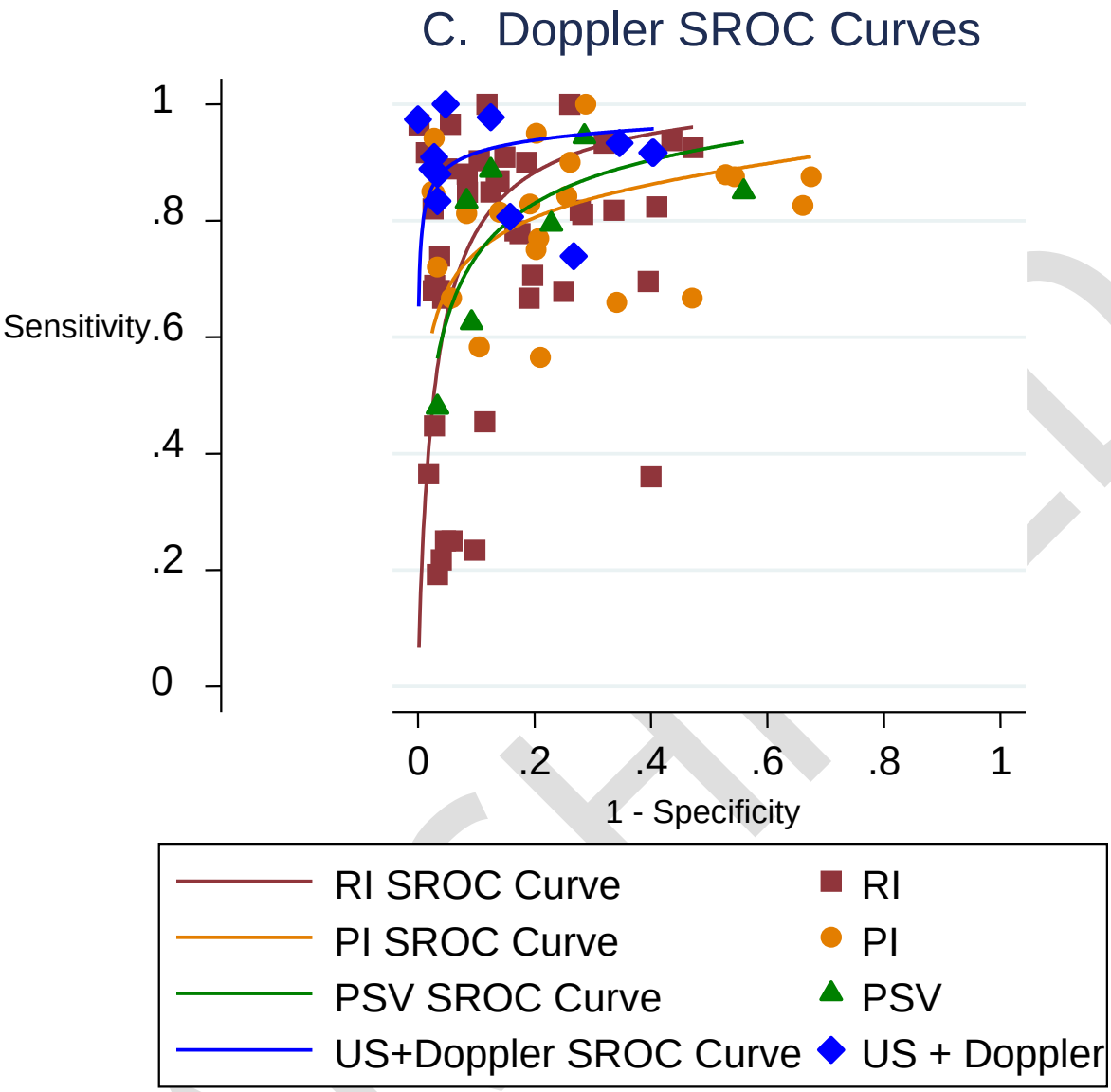
Figure 2. COMBINED Summary ROC curves for the performance of diagnostic tests in differentiating malignant from benign ovarian tumours

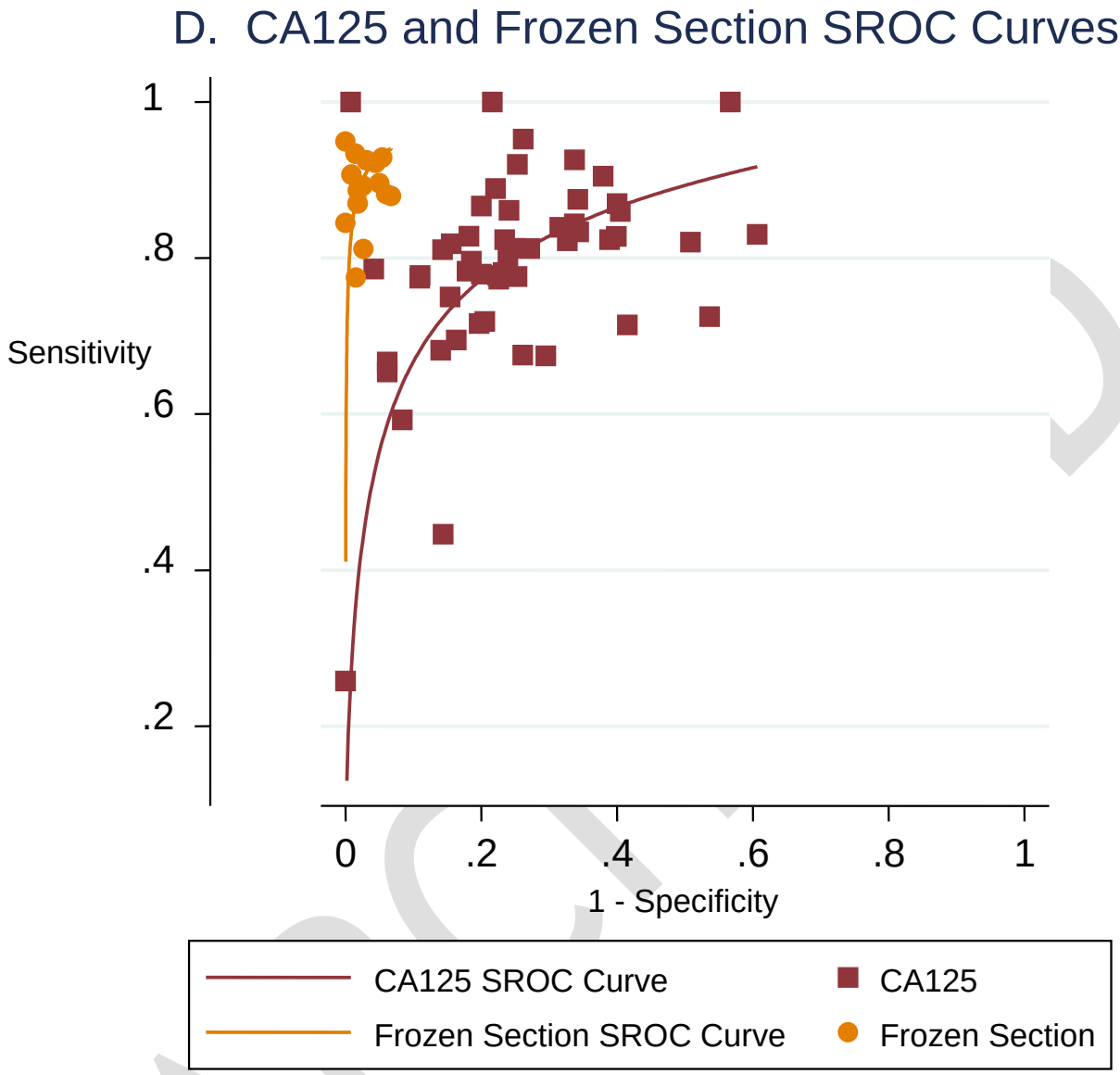


B. Scoring Systems SROC Curves



*Sassone cutoff point of 9; RMI1 cutoff point ≥200


C. Doppler SROC Curves
Sensitivity
1
.8
.6
.4
.2
0
0
.2
.4
.6
.8
1
1 - Specificity
RI SROC Curve
PI SROC Curve
PSV SROC Curve
US+Doppler SROC Curve
RI
PI
PSV
US + Doppler



*CA-125 cutoff point ≥35 U/mL

ACOG/SGO Referral Guidelines

The American College of Obstetricians and Gynecologists (ACOG) and the Society of Gynecologic Oncologists (SGO) jointly published guidelines for the referral of women with pelvic masses that are suspicious for ovarian cancer to gynecologic oncologists (136). The referral guidelines are based on patient age, CA-125 level (>200 U/mL for premenopausal, >35 U/mL for postmenopausal), physical findings, imaging results, and a family history of breast or ovarian cancer in a first-degree relative.

The referral guidelines were validated, and their role in distinguishing benign from malignant masses was tested in a multicentre setting with 1035 patients (44). In premenopausal women, the sensitivity and specificity of the referral guidelines in differentiating benign and malignant masses was 70% and 69%, respectively. Sensitivity was

enhanced in postmenopausal women, where the guideline correctly identified 94% of ovarian cancer patients. Specificity in this group was reported to be 58%.

The predictive value of the referral guidelines was further evaluated in a prospective cohort study of 837 consecutive patients (30). In premenopausal women, the sensitivity and specificity were 79.2% and 69.8%, respectively. In postmenopausal women, sensitivity was found to be 93.2% and specificity was 59.9%. In considering early- versus late-staged disease, the referral guidelines performed better in terms of sensitivity in late-staged disease, especially in postmenopausal women where sensitivity reached 98.3%. In premenopausal women, the referral guidelines were 92.3% sensitive in distinguishing malignant from benign cases.

2. Surgical procedures for an adnexal mass suspicious for malignancy

The Australian Cancer Network (ACN) 2004 guideline (6) on the management of women with epithelial ovarian cancer made recommendations on, among other things, the most appropriate surgical approach to take in such patients. Eight studies were included in the surgery for pelvic mass section of the guideline and are not discussed further here. An updated search of the literature identified an additional 16 studies (Table 11) published since the 2004 ACN guideline that met our inclusion criteria.

Table 11. New primary studies on surgical management of a suspicious adnexal mass categorized by surgical technique.

Study	Study Design	Patients	Procedure	Survival	Adverse Events
Surgical staging					
Chan 2007(85)	Retrospective cohort study SEER database Jan 1, 1988 to Dec 31, 2001 Follow-up not reported	Stage-I ovarian cancer excluding borderline ovarian/low malignant potential (LMP) tumours N=6686 Age <50=2764 Age >50=3922 <u>Age:</u> at diagnosis: 53.7±0.2 (range, 1-99)	Lymphadenectomy vs. No lymphadenectomy	<u>5-year disease specific survival rate:</u> Lymphadenectomy=92.6% No lymphadenectomy=87.0% Log rank p<0.001 Extent of lymphadenectomy on survival: 0 nodes: 87.0% <10 nodes: 91.9% >10 nodes: 93.8% Log rank p<0.001	NR
Chan 2008(86)	Retrospective cohort study SEER database 1988-2001 Follow-up not reported	Stage I-II EOC N=8372 Stage I: n=6152 Stage II: n=2220 <u>Age:</u> at diagnosis: 57 (range, 12-99)	Lymphadenectomy vs. No lymphadenectomy	<u>3-year disease specific survival rate:</u> Lymphadenectomy=93.3±0.5% No lymphadenectomy=82.0±0.6% Log rank p<0.001	NR
Cho 2006(87)	Retrospective cohort study Jan 1990-Oct 2005 Median follow-up (months): BOT staged: 39.0 BOT unstaged: 43.0 OC staged: 36.5 OC unstaged: 73.0	Patients with stage-I mucinous EOC apparently confined to ovaries N=264 <u>Age:</u> BOT staged: 41.9±16.5 BOT unstaged: 37.3±16.3 OC staged: 42.8±18.5 OC unstaged: 36.9±14.6	Surgical staging vs. Unstaged	Report no significant between-group differences in progression-free (p=0.889) and overall survival (p=0.958) rates	<u>Surgical complication rates:</u> Staged: 12.9% Unstaged: 1.0% p<0.001

Study	Study Design	Patients	Procedure	Survival	Adverse Events
Hornung 2004(88)	Retrospective cohort study Follow-up not reported	Primary malignant ovarian tumours N=111 Radical sx: n=54 Non-radical sx: n=38 Age: mean 58 (range, 21-94)	Lymphadenectomy vs. No lymphadenectomy	Report a statistically significant difference in mean overall (p=0.0008) and disease-free survival (p=0.0007) between the two groups, where lymphadenectomy showed a substantial benefit in survival	NR
Maggioni 2006(89)	Randomized controlled trial Median follow-up: 87.8 months	EOC macroscopically confined to pelvis and optimally debulked (RD ≤1cm) N=268 Systemic lymph: n=138 Control (lymph node sampling): n=130	Lymphadenectomy vs. No lymphadenectomy (sampling)	5-yr Progression-free survival: No lymphadenectomy: 73.4% Lymphadenectomy: 78.3% (difference = 4.9%; 95% CI -5.9 to 12.5%) HR (95% CI): 0.72 (0.46-1.14) p=0.16 5-yr Overall survival: No lymphadenectomy: 81.6% Lymphadenectomy: 84.0% (difference = 2.4%; 95% CI -8.3 to 8.9%) HR (95% CI): 0.85 (0.49-1.47) p=0.56 Note: study was underpowered	NR
Oksefjell 2008(90)	Retrospective cohort study 1984-2001 Follow-up not reported	Stage I EOC N=252	Lymph node staging vs. no lymph node staging	5- yr survival rate: Lymph node staging=87% No lymph node staging=64% p=0.02	NR
Skirnisdottir 2005(91)	Retrospective cohort study	Stage IA-IIC OC N=113	Lymph node sampling v. no sampling	Disease-free survival: Lymphadenectomy: 95% No sampling: 62.4%	NR

Study	Study Design	Patients	Procedure	Survival	Adverse Events
	Jan 1, 1994 to Dec 31, 1998 Median follow-up (months): 67 (range, 36-97)	Lymph n=20 No sampling n=93 Age: mean 61 (range, 23-88)		p=0.004 <u>Odds Ratio (95% CI):</u> 0.092 (0.013-0.670) p=0.019	
Suzuki 2008(92)	Retrospective cohort study 1986-2006 Median follow-up 49 months	pTI-IIb pure type CCC N= 205 Grp A: n = 104 Grp B: n = 101 Median Age: Grp A: 52 (range, 30-75) Grp B: 51 (range, 32-75)	Group A= systemic retroperitoneal lymphadenectomy vs. Group B= no systemic lymphadenectomy	<u>5-yr disease specific survival:</u> Grp A: 84.7% Grp B: 85.3% p=0.645 <u>5 –yr disease-free survival:</u> Grp A: 79.7% Grp B: 73.5% p=0.353	NR
Wong 2007(93)	Retrospective cohort study Jan 1991-Dec 2004 Mean follow-up: 21 months (range 2-140)	LMP tumours N=247 Staged: 164 Unstaged: 83 Mean age: 38 (range, 16-89)	Surgical staging vs. unstaged	<u>Overall survival rate:</u> Staged: 97.6% Unstaged: 98.8% p=0.653 <u>Disease-free survival rate:</u> Staged: 97.6% Unstaged: 97.6% p=1.00	NR
Laparoscopy (LPS) vs. Laparotomy (LPT)					
Desfeux 2005(94)	Retrospective Cohort study Jan 1, 1985 to Dec 31,2001 Median follow-up 29 months (interquartile range, 16-55)	LMP tumours with final stage IA-IC N=118 Age (mean): 45±16	Laparoscopy vs. laparotomy	Reported no difference in DFS and OS between the 2 surgical approaches (p=0.6), but number of events was too small and only afforded a 20% power to detect a statistically significant difference.	<u>Intraoperative tumour rupture:</u> p=0.1

Study	Study Design	Patients	Procedure	Survival	Adverse Events
Ghezzi 2007(95)	Historically controlled trial 2003; LPT controls from 1997-2003 Median follow-up (months): LPS 16 (4-33) LPT 60 (32-108)	Women with apparent early-stage OC N=34 LPS group: n=15 Age: 55 (range, 13-70) LPT group: n=19 Age: 61 (range, 44-71)	Laparoscopy (LPS) vs. Laparotomy (LPT)	Disease-free survival: LPS: 100% LPT: 92% p value not reported Overall survival: LPS: 100% LPT: 100%	Postoperative complications: LPS: 6.7% LPT: 42.1% p=0.047
Lecuru 2006(96)	Retrospective cohort study Jan 1, 1985 to Dec 31, 2001 Mean follow-up of 40 months (range, 1.7-182)	Pt with proven stage I EOC N=178 Group 1: laparoscopy n=34 Group 2: laparotomy n=114 Group 3: laparoscopy converted to laparotomy n=30 Age: Grp 1: 43±12 Grp 2: 51±16 Grp 3: 41±15	Laparoscopy vs. laparotomy	Survival: Grp 1: 100% Grp 2: 97% Grp 3: 100% p=0.06	Intraoperative tumour rupture: Grp 1: 31% Grp 2: 16% Grp 3: 22% p=0.19
Park 2008(97)	Retrospective cohort study Jan 2001 to Aug 2006 Median follow-up: 17 months (range, 5-61)	Pt with apparent stage I EOC N=36 LPT group: n=19 LPS group: n=17 Age: LPT=48.9±10.8 LPS=43.2±12.3	laparoscopy vs. laparotomy	Reported no difference in DFS and OS between the 2 groups (p=0.123 and 0.280, respectively)	Intraoperative complications: Laparotomy: 5.3% Laparoscopy: 11.8% p value not reported

Study	Study Design	Patients	Procedure	Survival	Adverse Events
Romagnolo 2006(98)	Retrospective cohort study Jan 1992 to June 2004 Median follow-up (months): Laparotomy: 40 Laparoscopy: 47	LMP tumours N=113 LPT: n=61 LPS: n=52 Age: mean 44 (range 20-88)	Laparotomy vs. Laparoscopy	Disease-free survival rate: Laparotomy: 90.2% Laparoscopy: 86.5% p=not significant	Intraoperative tumour rupture: Laparotomy: 6.6% Laparoscopy: 34.6% p<0.0001
Park 2009(99)	Retrospective cohort study April 1989 to May 2008 Mean follow-up: 70 months (range, 3-216)	LMP tumours N=360 Laparotomy: n=289 Laparoscopy: n=71 Age: Radical sx: 51.8±12.7 Fertility-sparing sx: 29.6±11.5	Laparotomy vs. Laparoscopy	Disease-free survival OR: 1.17 (95% CI: 0.33-4.08) p=0.808	NR
Fertility-sparing vs. more extensive (radical) surgery					
Park 2009(99)	Retrospective cohort study April 1989 to May 2008 Mean follow-up: 70 months (3-216)	LMP tumours N=360 Radical sx: n=176 Fertility-sparing sx; n=184 Age: Radical sx: 51.8±12.7 Fertility-sparing sx: 29.6±11.5	Fertility-sparing sx vs. radical sx	10 yr DFS rates Radical sx: 92% Fertility-sparing sx: 95% OR: 0.38 (95% CI: 0.09-1.46) p=0.157 10-yr OS rates Radical sx: 97% Fertility-sparing sx: 98%	Intraoperative tumour rupture: Radical sx: 13.1% Fertility-sparing sx: 6.0% p=0.021
Yinon 2007(100)	Retrospective cohort study 1979-2004 Mean follow-up:	LMP tumours N=62 USO: n=40 Cystectomy: n=22	Cystectomy vs. unilateral salpingo-oophorectomy (USO)	Disease-free survival (mean): Cystectomy: 23.6 months USO: 41 months p=0.2	Reoperation: Cystectomy: 27% USO: 37.5% p value not reported

Study	Study Design	Patients	Procedure	Survival	Adverse Events
	Cystectomy: 65 months (range, 6-180) USO: 101 months (range, 6-30)	Mean age: 28 (range, 13-44)			

Abbreviations: BOT = borderline ovarian tumours; EOC = epithelial ovarian cancer; OC = ovarian cancer; RD = residual disease;LMP = low malignant potential; LPS = laparoscopy; LPT = laparotomy; DFS = disease-free survival; OS = overall survival; OR = odds ratio; sx= surgery; NR = not reported.

Survival

A randomized controlled trial (RCT) (89) of patients undergoing a systematic aortic and pelvic lymphadenectomy versus those undergoing lymph node sampling reported no statistically significant difference in five-year progression-free (p=0.16) or overall survival (p=0.56). However, this RCT was underpowered to detect an effect of systematic lymphadenectomy on survival. The sample size calculation in this study was undertaken to detect a difference in prevalence of lymph node positivity, the study's primary outcome. The targeted sample size required to detect an effect of lymphadenectomy on survival, the secondary outcome, was deemed unattainable by the researchers. Despite the reduced power to detect a statistical difference in the secondary outcomes, the study reported a trend favouring lymphadenectomy in terms of progression-free (HR, 0.72; 95% CI, 0.46 to 1.14) and overall (HR, 0.85; 95% CI, 0.49 to 1.47) survival.

Eight additional studies (85-88,90-93) investigated the survival impact of comprehensive surgical staging in women diagnosed with early-stage ovarian cancer. In two large population-based studies (85-86), consisting exclusively of over 6600 early-stage epithelial ovarian cancer patients, Chan et al found that surgical staging with lymphadenectomy was associated with improved three-year (p<0.001) (86) and five-year disease-specific survival (p<0.001) (85) compared to staging procedures without lymphadenectomy. Similarly, Oksefjell et al reported a statistically significant improvement in 5-year overall survival rates in patients undergoing a lymphadenectomy versus those that did not (87% vs. 64%, respectively; p=0.02). Survival analyses performed by both Skirnisdottir et al (91) and Hornung et al (88) also demonstrated a statistically significant benefit in disease-free survival (p=0.004 and p=0.0007, respectively) for patients undergoing a lymphadenectomy versus those that did not. Hornung and colleagues also considered overall survival and reported a statistically significant difference (p=0.0008) in the two patient groups in favour of the lymphadenectomy group. Conversely, in 205 patients with pTI-IIb clear cell carcinoma, Suzuki et al (92) found that patients who underwent systemic lymphadenectomy did not show a significant improvement in disease-free (p=0.353) or overall survival (p=0.645) compared to those that did not. Similarly, Cho et al (87) report no significant difference between groups in progression-free and overall survival rates in patients with stage I mucinous epithelial ovarian tumours undergoing complete staging versus those whose staging was incomplete.

In an attempt to determine the benefit of surgically staging ovarian low malignant potential (LMP) tumours, Wong and colleagues (93) retrospectively reviewed the records of 247 patients with tumours of borderline malignancy and found no statistically significant difference in rates of recurrence or mortality between patients surgically staged and those who were unstaged.

Six studies (94-99, 137) compared laparoscopy versus laparotomy for the surgical management of women with apparent early ovarian cancer or borderline tumours. Patient sample sizes in these studies ranged from 34 to 360. In the three studies (95-97) that considered patients with early epithelial ovarian cancer (EOC), no statistical difference in survival rates was detected between patients undergoing a laparoscopy versus laparotomy. Similarly, in the management of patients with early borderline ovarian tumours, Romagnolo et al (98), Park et al (99), and Desfeux et al (94) found that a laparoscopic versus laparotomic surgical approach did not appear to influence survival rates, although Desfeux et al acknowledged that the number of events was too small to allow for proper statistical testing.

Fertility-preserving treatments are often desirable for women of reproductive age who are diagnosed with borderline ovarian tumours (BOT). Two studies compared the impact of conservative fertility-sparing surgeries versus more extensive surgical approaches. Yinon et al (100) specifically compared rates of recurrence in 40 patients who underwent unilateral

salpingo-oophorectomy versus 22 patients who underwent cystectomy only. No statistical difference in recurrence rates was detected (27.5% vs. 22.7%, p=0.8). Similarly, in a larger study of 360 women with BOT, Park et al (99) found no difference in disease-free survival between patients that underwent radical or fertility-sparing surgery (p=0.651).

Adverse Events

While the surgical technique did not appear to impact patient survival, there were differences detected in surgical outcomes and complication rates. Cho et al (87) found a statistically significant difference in complication rates, with 12.9% experiencing a complication in the completely staged group versus 1.0% in those incompletely staged ($p<0.001$). Ghezzi et al (95) reported a statistically significant difference in the rates of minor postoperative complications, with 6.7% of patients in the laparoscopy group experiencing such an event compared to 42.1% of patients in the laparotomy group (p=0.047). Romagnolo et al (98) reported a difference in the cases of tumour rupture or spilling during surgery, with 34.6% ruptures recorded in the laparoscopic group compared to 6.6% in patients undergoing a laparotomy ($p<0.0001$). Similarly, Lecuru et al (96) found 31% of laparoscopic patients experienced intraoperative tumour rupture versus 16% in the laparotomy group. However, this difference did not reach statistically significance. In patients with borderline tumours, the difference in the occurrence of intraoperative tumour rupture was found not to be statistically associated with the surgical approach according to Desfeux et al. (94)

DISCUSSION

The basic diagnostic work-up of patients with a suspicious adnexal mass involves a gynecological exam, ultrasound imaging and testing of serum tumour markers. While this approach is often sufficient in detecting advanced disease, the diagnosis of early-staged ovarian cancer is more challenging. In an attempt to determine the best method for the identification and diagnosis of a suspicious adnexal mass, we systematically reviewed existing guidelines and the literature.

The intention and necessity of preoperative diagnosis of an adnexal mass is to both triage the patient and better define the surgical options. Recent Canadian guidelines on the evaluation and referral of ovarian masses (138) recommend that patients with a high-level risk of underlying malignancy be reviewed in consultation with a gynecologic oncologist. Indeed, studies considering the impact of physician's specialty on the survival of patients with early stage ovarian cancer have shown a trend towards improved survival when gynecologic oncologists perform the surgery (139-140). Along with triaging appropriate patients to subspecialists, the preoperative diagnosis of an adnexal mass provides the anatomic details necessary to inform the surgical options. For patients with clearly benign masses, or likely benign masses with health issues, observation only is the usual management. Such preoperative diagnostic information can also help to distinguish patients who require extensive surgery over those for whom a more conservative surgical approach is adequate.

Statistical comparisons between diagnostic techniques could not be performed in this review. Instead, the assessment of various modalities in differentiating benign from malignant masses is based on inspection of the summary data obtained from the meta-analyses conducted. These results suggest that 3D ultrasonography has both a higher sensitivity and specificity when compared to 2D ultrasound. Established morphological scoring systems also performed with respectable sensitivity and specificity, each system with equivalent diagnostic competence. Explicit scoring systems did not appear to perform as well as other diagnostic testing methods. Not only was external validation of the models lacking in many cases, where results were available, but performance was often poorer upon subsequent replication. Assessment of an adnexal mass by colour Doppler technology, using the RI, PI, and PSV

indices, was neither as sensitive nor as specific as simple ultrasonography. Furthermore, because of overlap of vascular parameters between malignant and benign masses, a firm diagnosis based on Doppler evaluation alone can be problematic (141). Summary estimates from studies considering combined morphology and Doppler assessment were higher than the estimates for either modality alone. Both sensitivity and specificity of this combined approach were high. Of the three imaging modalities considered, MRI appeared to perform the best, although results were not statistically different from CT as determined by overlapping confidence intervals. PET did not appear to perform as well as either MRI or CT, although only three studies were considered in the analysis. The measurement of the CA-125 tumour marker appears to be less reliable than other available assessment methods; however, results were not stratified by menopausal status. It is widely reported that differences exist in CA-125 levels between premenopausal and postmenopausal women, even with the same histological diagnosis (75). Finally, frozen-section analysis has both a high sensitivity and an especially high specificity in the assessment of adnexal masses.

In the treatment of ovarian cancer, the importance of surgical management is universally recognized. Clearly, complete surgical resection is required to improve a patient's prognosis. However, it is not clear how aggressive a surgical approach is necessary in early-staged ovarian cancer. The included evidence suggests that systematic lymphadenectomy improves survival, as does proper surgical staging. There is an exception to this benefit, however, with tumours of low malignant potential. In such tumours, conservative fertility-preserving surgical approaches appear to have no adverse survival effect (99).

The accuracy and adequacy of surgical staging by laparotomy or laparoscopic approaches appears to be comparable, with neither approach conferring a survival advantage compared to the other. In spite of this, many surgeons are unwilling to perform laparoscopic surgical staging in early-staged ovarian cancer due to the potential risks of intraoperative tumour rupture, port-site metastasis, and dissemination of the tumour (97). Intraoperative tumour rupture was indeed reported to occur more frequently in patients undergoing laparoscopy versus laparotomy in two retrospective cohort studies (96, 98). However, unequivocal prospective comparative data supporting the existence of an increased occurrence of intraoperative tumour rupture in ovarian cancer patients managed by laparoscopy is still lacking (95).

The evidence included in this systematic review is not without significant limitations. In the identification of an adnexal mass literature, the population of patients often included women with a suspicious mass undergoing surgery. This inclusion of operative patients could escalate the prevalence of malignancy in comparison to what would be expected in a primary care population of women presenting with a suspicious adnexal mass. This has implications for the generalizability of the results, and it can inflate the test's sensitivity. Furthermore, the included studies did not often allow for stratification by menopausal status. Given the lower likelihood of ovarian malignancy in premenopausal patients, the accuracy of a diagnostic test can be reduced in a sample that includes a high proportion of premenopausal patients.

In addition to these methodological issues, the quality assessment of the included literature revealed several further shortcomings, especially in study design and reporting. Blinding is a crucial issue in diagnostic accuracy studies, as it is necessary to prevent information bias (142). While the index test always preceded surgery in the included studies, thereby by nature blinding the outcome, the reference test should also be interpreted without knowledge of the index test results (142). The vast majority of studies considered in the identification of a suspicious adnexal mass did not report such blinding.

One third of the diagnostic studies had data collection that occurred retrospectively or was not reported. Retrospective study designs are inherently more prone to bias than are prospective studies and can be more difficult to interpret, especially if the sampling did not

include consecutive patients. The reporting of consecutive patient sampling occurred in less than half the included studies.

Many of the surgical studies were also prone to biases inherent in their retrospective designs. Moreover, the small sample of included patients meant many studies were underpowered to detect statistically significant results. There is an obvious need for improvement in the quality of primary research in this area. Ideally, future research would consist of randomized clinical trials with a non-inferiority design where survival outcomes are considered.

Despite these limitations, the best available evidence with respect to the questions posed was collected and included. A rigorous systematic review and meta-analysis, planned a priori, provided an abundant evidentiary base and the context and direction for the development of recommendations.

CONCLUSIONS

There are numerous methodologies that have been considered in the preoperative identification of adnexal masses suspicious for malignancy. Results suggest that 3D ultrasonography has both a higher sensitivity and specificity when compared to 2D ultrasound. Established morphological scoring systems also performed with respectable sensitivity and specificity, each morphological system with equivalent diagnostic competence. Explicit scoring systems did not perform as well as other diagnostic testing methods. Assessment of an adnexal mass by colour Doppler technology, using the RI, PI, and PSV indices, was neither as sensitive nor as specific as simple ultrasonography. Summary estimates from studies considering combined morphology and Doppler assessment were higher than the estimates for either modality alone. Both sensitivity and specificity of this combined approach were high. Of the three imaging modalities considered, MRI appeared to perform the best, although results were not statistically different from CT. PET did not perform as well as either MRI or CT. The measurement of the CA-125 tumour marker appears to be less reliable than do other available assessment methods. Finally, frozen section analysis has both a high sensitivity and especially high specificity in the assessment of adnexal masses.

The evidence suggests that systematic lymphadenectomy and proper surgical staging improve survival. Conservative fertility-preserving surgical approaches are an acceptable option in women with low malignant potential tumours. The accuracy and adequacy of surgical staging by laparotomy or laparoscopic approaches appears to be comparable, with neither approach conferring a survival advantage.

CONFLICT OF INTEREST

No conflicts of interest were declared for J Dodge, A Covens, C Lacchetti, L Elit, T Le, M Devries-Aboud, or M Fung Kee Fung.

JOURNAL REFERENCE

The following systematic reviews and meta-analysis has been published in *Gynecologic Oncology* (© 2012 Elsevier Inc.;http://www.journals.elsevier.com/gynecologic-oncology/):

- Covens AL, Dodge JE, Lacchetti C, Elit LM, Le T, Devries-Aboud M, et al. Surgical management of a suspicious adnexal mass: a systematic review. Gynecol Oncol. 2012 Jul;126(1):149-56. doi: 10.1016/j.ygyno.2012.04.018.
- Dodge JE, Covens AL, Lacchetti C, Elit LM, Le T, Devries-Aboud M, et al. Preoperative identification of a suspicious adnexal mass: a systematic review and meta-analysis. Gynecol Oncol. 2012 Jul;126(1):157-66. doi: 10.1016/j.ygyno.2012.03.048.

The following clinical practice guideline has been published in *Current Oncology* (© 2012 Multimed Inc.):

- Dodge JE, Covens AL, Lacchetti C, Elit LM, Le T, Devries-Aboud M, et al. Management of a suspicious adnexal mass: a clinical practice guideline. Curr Oncol. 2012;19(4):e244-57. doi: 10.3747/co.19.980.

ACKNOWLEDGEMENTS

The Gynecology Cancer Disease Site Group would like to thank J Dodge, A Covens, C Lacchetti, L Elit, T Le, M Devries-Aboud, and M Fung Kee Fung for taking the lead in drafting this systematic review with meta-analysis.

For a complete list of the Gynecology Cancer DSG members,
please visit the PEBC section of the CCO website at http://www.cancercare.on.ca/

PEBC Contact Information

For further information about this series and all PEBC reports, including previous versions, please contact the PEBC office at:

Phone: 905-527-4322, ext. 42842 Fax: 905-526-6775 E-mail: ccopgi@mcmaster.ca

REFERENCES

1. Canadian Cancer Society's Steering Committee. Canadian Cancer Statistics 2010 Toronto: Canadian Cancer Society; 2010.
2. Scottish Intercollegiate Guidelines Network (SIGN). Epithelial ovarian cancer, a national clinical guideline. Edinburgh (Scotland): SIGN; 2003 Oct. Report No.: 75
3. Myers E, Bastian L, Havrilesky L, Kulasingam S, Terplan M, Cline K, et al. Management of Adnexal Mass. Rockville (MD): Agency for Healthcare Research and Quality; 2006 Feb. Report No.: Evidence Report/Technology Assessment No.: 130. AHRQ Publication No.: 06-E004. Contract No.: 290-02-0025.
4. Fung Kee Fung M, Bryson P, Johnston M, Chambers A, members of the Gynecology Cancer Disease Site Group. Screening postmenopausal women for ovarian cancer. Program in Evidence-based Care Evidence Summary Report No.: 4-6a. Toronto: Cancer Care Ontario; 2004 Jan.
5. Browman G, Levine M, Mohide E, Hayward R, Pritchard K, Gafni A. The practice guidelines development cycle: a conceptual tool for practice guidelines development and implementation. J Clin Oncol. 1995;13:502-12. .
6. Australian Cancer Network, National Breast Cancer Centre. Clinical practice guidelines for the management of women with epithelial ovarian cancer. Camperdown, Australia: National Breast Cancer Centre; 2004.
7. van Houwelingen H, Arends LR, Stijnen T. Advanced methods in meta-analysis: multivariate approach and meta-regression. Stat Med. 2002;21(4):589-624.
8. Reitsma JB, Glas AS, Rutjes AWS, Scholten RJPM, Bossuyt PM, AH Z. Bivariate analysis of sensitivity and specificity produces information summary measures in diagnostic reviews. J Clin Epidemiol. 2005;58:982-90.
9. Arends LR, Hamza TH, van Houwelingen JC, Heijenbrok-Kal MH, Hunink MGM, Stijnen T. Bivariate random effects meta-analysis of ROC curves. Med Decis Making. 2008;28:621-38.
10. Dinnes J, Deeks J, Kirby J, P. R. A methodological review of how heterogeneity has been examined in systematic reviews of diagnostic test accuracy. Health Tech Assess. 2005;9(12):1-113.
11. Statacorp. Stata Statistical Software: Release 11. College Station, TX: StataCorp LP; 2009.
12. The AGREE Collaboration, Group: W, Cluzeau FA, Burgers JS, Brouwers M, Grol R, et al. Development and validation of an international appraisal instrument for assessing the quality of clinical practice guidelines: the AGREE project. Qual Safe Health Care. 2003;12(1):18-23.
13. Shea B, Grimshaw JM, Wells GA, Boers M, Andersson N, Hamel C, et al. Development of AMSTAR: a measurement tool to assess the methodological quality of systematic reviews. BMC Med Res Method. 2007;7:10.
14. Geomini P, Bremer G, Kruitwagen R, Mol BWJ. Diagnostic accuracy of frozen section diagnosis of the adnexal mass: a metaanalysis. Gynecol Oncol. 2005 Jan;96(1):1-9.
15. Liu J, Xu Y, Wang J. Ultrasonography, computed tomography and magnetic resonance imaging for diagnosis of ovarian carcinoma. Eur J Radiol. 2007;62(3):328-34.
16. Geomini P, Kruitwagen R, Bremer GL, Cnossen J, Mol BWJ. The accuracy of risk scores in predicting ovarian malignancy: a systematic review. Obstet Gynecol. 2009 Feb;113(2 Pt 1):384-94.
17. Medeiros L, Rosa D, Edelweiss M, Stein A, Bozzetti M, Zelmanowicz A, et al. Accuracy of frozen-section analysis in the diagnosis of ovarian tumours: a systematic review. Int J Gynecol Cancer. 2005;15:192-202.

18. Alcazar J, Castillo G. Comparison of 2-dimensional and 3-dimensional power-Doppler imaging in complex adnexal masses for the prediction of ovarian cancer. Am J Obstet Gynecol. 2005;807(12):192(3).
19. Bazot M, Nassar-Slaba J, Thomassin-Naggara I, Cortez A, Uzan S, Darai E. MR imaging compared with intraoperative frozen-section examination for the diagnosis of adnexal tumors; correlation with final histology. Eur Radiol. 2006 Dec;16(12):2687-99.
20. Benjapibal M, Neungton C. Pre-operative prediction of serum CA125 level in women with ovarian masses. J Med Assoc Thai. 2007 Oct;90(10):1986-91.
21. Bensaid C, Le Frere Belda MA, Metzger U, Larousserie F, Clement D, Chatellier G, et al. Performance of laparoscopy in identifying malignant ovarian cysts. Surg Endosc. 2006 Sep;20(9):1410-4.
22. Booth SJ, Turnbull LW, Poole DR, Richmond I. The accurate staging of ovarian cancer using 3T magnetic resonance imaging--a realistic option. BJOG: Int J Obstet Gynaecol. 2008 Jun;115(7):894-901.
23. Boriboonhirunsarn D, Sermboon A. Accuracy of frozen section in the diagnosis of malignant ovarian tumor. J Obstet Gynaecol Res. 2004 Oct;30(5):394-9.
24. Brun J-L, Cortez A, Rouzier R, Callard P, Bazot M, Uzan S, et al. Factors influencing the use and accuracy of frozen section diagnosis of epithelial ovarian tumors. Am J Obstet Gynecol. 2008 Sep;199(3):244.e1-7.
25. Canis M, Mashiach R, Wattiez A, Botchorishvili R, Rabischong B, Jardon K, et al. Frozen section in laparoscopic management of macroscopically suspicious ovarian masses. J Am Assoc Gynecol Laparosc. 2004 Aug;11(3):365-9.
26. Chen M, Wang WC, Zhou C, Zhou NN, Cai K, Yang ZH, et al. Differentiation between malignant and benign ovarian tumors by magnetic resonance imaging. Chin Med Sci J. 2006 Dec;21(4):270-5.
27. Choudhury S, Mohiuddin ASM, Ahmed AU, Ahsan S. Preoperative discrimination of benign and malignant ovarian tumors using color Doppler sonography and its correlation with histopathology. Bangladesh Med Res Counc Bull. 2005 Apr;31(1):21-6.
28. Dai S-Y, Hata K, Inubashiri E, Kanenishi K, Shiota A, Ohno M, et al. Does three-dimensional power Doppler ultrasound improve the diagnostic accuracy for the prediction of adnexal malignancy? J Obstet Gynaecol Res. 2008 Jun;34(3):364-70.
29. Daponte A, Stergioti E, Messinis I. Risk scoring for adnexal masses and endoscopic management. Int J Gynaecol Obstet. 2007 Jan;96(1):42-3.
30. Dearking AC, Aletti GD, McGree ME, Weaver AL, Sommerfield MK, Cliby WA. How relevant are ACOG and SGO guidelines for referral of adnexal mass? Obstet Gynecol. 2007 Oct;110(4):841-8.
31. El-Shalakany A, Abou-Talib Y, Shalaby HS, Sallam M. Preoperative serum inhibin levels in patients with ovarian tumors. J Obstet Gynaecol Res. 2004 Apr;30(2):155-61.
32. Engelen MJA, Bongaerts AHH, Sluiter WJ, de Haan HH, Bogchelman DH, Tenvergert EM, et al. Distinguishing benign and malignant pelvic masses: the value of different diagnostic methods in everyday clinical practice. Eur J Obstet Gynecol Reprod Biol. 2008 Jan;136(1):94-101.
33. Erdogan N, Ozcelik B, Serin IS, Akgun M, Ozturk F. Doppler ultrasound assessment and serum cancer antigen 125 in the diagnosis of ovarian tumors. Int J Gynaecol Obstet. 2005 Nov;91(2):146-50.
34. Exacoustos C, Romanini ME, Rinaldo D, Amoroso C, Szabolcs B, Zupi E, et al. Preoperative sonographic features of borderline ovarian tumors. Ultrasound Obstet Gynecol. 2005;25(1):50-9.

35. Fanfani F, Zannoni GF, Fagotti A, Gagliardi ML, Masciullo V, Testa AC, et al. Importance of a specialized pathologist for the examination of frozen sections of adnexal masses. Int J Gynecol Cancer. 2007 Sep-Oct;17(5):1034-9.
36. Ferrazzi E, Lissoni AA, Dordoni D, Trio D, Redaelli L, Rusconi C, et al. Differentiation of small adnexal masses based on morphologic characteristics of transvaginal sonographic imaging: a multicenter study. J Ultrasound Med. 2005;24 (11):1467-73.
37. Geomini P, Zuurendonk LD, Bremer GL, de Graaff J, Kruitwagen R, Mol BW. The impact of size of the adnexal mass on the accuracy of frozen section diagnosis. Gynecol Oncol. 2005 Nov;99(2):362-6.
38. Geomini P, Kluivers K, Moret E, Bremer GL, Kruitwagen R, Mol B. Evaluation of adnexal masses with three-dimensional ultrasonography. Obstet Gynecol. 2006 Nov;108(5):1167-75.
39. Ghaemmaghami F, Fakour F, Karimi Zarchi M, Behtash N, Modares Gilani M, Mousavi A, et al. Clinical assessment, gross examination, frozen section of ovarian masses: do patients benefit? Arch Gynecol Obstet. 2008 Sep;278(3):209-13.
40. Guerra A, Cunha TM, Felix A. Magnetic resonance evaluation of adnexal masses. Acta Radiol. 2008 Jul;49(6):700-9.
41. Guerriero S, Ajossa S, Piras S, Gerada M, Floris S, Garau N, et al. Three-dimensional quantification of tumor vascularity as a tertiary test after B-mode and power Doppler evaluation for detection of ovarian cancer. J Ultrasound Med. 2007 Oct;26(10):1271-8.
42. Hata T, Yanagihara T, Hayashi K, Yamashiro C, Ohnishi Y, Akiyama M, et al. Three-dimensional ultrasonographic evaluation of ovarian tumours: a preliminary study. Hum Reprod. 1999;14(3):858-61.
43. Ilvan S, Ramazanoglu R, Ulker Akyildiz E, Calay Z, Bese T, Oruc N. The accuracy of frozen section (intraoperative consultation) in the diagnosis of ovarian masses. Gynecol Oncol. 2005;97(2):395-9.
44. Im SS, Gordon AN, Buttin BM, Leath CA, 3rd, Gostout BS, Shah C, et al. Validation of referral guidelines for women with pelvic masses. Obstet Gynecol. 2005 Jan;105(1):35-41.
45. Jokubkiene L, Sladkevicius P, Valentin L. Does three-dimensional power Doppler ultrasound help in discrimination between benign and malignant ovarian masses? Ultrasound Obstet Gynecol. 2007 Feb;29(2):215-25.
46. Kitajima K, Murakami K, Yamasaki E, Kaji Y, Fukasawa I, Inaba N, et al. Diagnostic accuracy of integrated FDG-PET/contrast-enhanced CT in staging ovarian cancer: comparison with enhanced CT. Eur J Nucl Med Mol Imaging. 2008 Oct;35(10):1912-20.
47. Kurjak A, Kupesic S, Sparac V, Kosuta D. Three-dimensional ultrasonographic and power Doppler characterization of ovarian lesions. Ultrasound Obstet Gynecol. 2000 Sep;16(4):365-71.
48. Laban M, Metawee H, Elyan A, Kamal M, Kamel M, Mansour G. Three-dimensional ultrasound and three-dimensional power Doppler in the assessment of ovarian tumors. Int J Gynaecol Obstet. 2007 Dec;99(3):201-5.
49. Lee TS, Kim JW, Park NH, Song YS, Kang SB, Lee HP. Assessing clinical performance of gynecology residents: sonographic evaluation of adnexal masses based on morphological scoring systems. Ultrasound Obstet Gynecol. 2005 Dec;26(7):776-9.
50. Leelahakorn S, Tangjitgamol S, Manusirivithaya S, Thongsuksai P, Jaroenchainon P, Jivangkul C. Comparison of ultrasound score, CA125, menopausal status, and risk of malignancy index in differentiating between benign and borderline or malignant ovarian tumors. J Med Assoc Thai. 2005 Oct;88 Suppl 2:S22-30.

51. Marchesini ACdS, Magrio FAA, Berezowski AT, Neto OBP, Nogueira AA, Candido dos Reis FJ. A critical analysis of Doppler velocimetry in the differential diagnosis of malignant and benign ovarian masses. J Womens Health (Larchmt). 2008 Jan-Feb;17(1):97-102.
52. Marret H, Sauget, S., Giraudeau, B., Body, G., Tranquart, F. Power Doppler vascularity index for predicting malignancy of adnexal masses. Ultrasound Obstet Gynecol. 2005;25:508-13.
53. Milojkovic M, Hrgovic Z, Hrgovic I, Jonat W, Maass N, Bukovic D. Significance of CA 125 serum level in discrimination between benign and malignant masses in the pelvis. Arch Gynecol Obstet. 2004 Mar;269(3):176-80.
54. Moszynski R, Szpurek D, Smolen A, Sajdak S. Comparison of diagnostic usefulness of predictive models in preliminary differentiation of adnexal masses. Int J Gynecol Cancer. 2006 Jan-Feb;16(1):45-51.
55. Mousavi AS, Borna S, Moeinoddini S. Estimation of probability of malignancy using a logistic model combining color Doppler ultrasonography, serum CA125 level in women with a pelvic mass. Int J Gynecol Cancer. 2006 Jan-Feb;16 Suppl 1:92-8.
56. Naik R, Cross P, Lopes A, Godfrey K, Hatem MH. "True" versus "apparent" stage I epithelial ovarian cancer: value of frozen section analysis. Int J Gynecol Cancer. 2006 Jan-Feb;16 Suppl 1:41-6.
57. Nakae M, Iwamoto I, Fujino T, Maehata Y, Togami S-i, Yoshinaga M, et al. Preoperative plasma osteopontin level as a biomarker complementary to carbohydrate antigen 125 in predicting ovarian cancer. J Obstet Gynaecol Res. 2006 Jun;32(3):309-14.
58. Romagnolo C, Trivella G, Bonacina M, Fornale M, Maggino T, Ferrazzi E. Preoperative diagnosis of 221 consecutive ovarian masses: scoring system and expert evaluation. Eur J Gynaecol Oncol. 2006;27(5):487-9.
59. Sladkevicius P, Jokubkiene L, Valentin L. Contribution of morphological assessment of the vessel tree by three-dimensional ultrasound to a correct diagnosis of malignancy in ovarian masses. Ultrasound Obstet Gynecol. 2007 Nov;30(6):874-82.
60. Sohaib SA, Mills TD, Sahdev A, Webb JA, Vantrappen PO, Jacobs IJ, et al. The role of magnetic resonance imaging and ultrasound in patients with adnexal masses. Clin Radiol. 2005 Mar;60(3):340-8.
61. Stewart CJR, Brennan BA, Hammond IG, Leung YC, McCartney AJ. Intraoperative assessment of ovarian tumors: a 5-year review with assessment of discrepant diagnostic cases. Int J Gynecol Pathol. 2006 Jul;25(3):216-22.
62. Szpurek D, Moszynski R, Zietkowiak W, Spaczynski M, Sajdak S. An ultrasonographic morphological index for prediction of ovarian tumor malignancy. Eur J Gynaecol Oncol. 2005;26(1):51-4.
63. Szpurek D, Moszynski R, Smolen A, Sajdak S. Artificial neural network computer prediction of ovarian malignancy in women with adnexal masses. Int J Gynecol Obstet. 2005;89(2):108-13.
64. Tan PL, Willatt JM, Lindsell D. The ability of ultrasound to detect gynaecological neoplasms and their ultrasound morphological features. Australas Radiol. 2007 Jun;51(3):260-6.
65. Tangjitgamol S, Jesadapatrakul S, Manusirivithaya S, Sheanakul C. Accuracy of frozen section in diagnosis of ovarian mass. Int J Gynecol Cancer. 2004;14(2):212-9.
66. Taskiran C, Erdem O, Onan A, Bozkurt N, Yaman-Tunc S, Ataoglu O, et al. The role of frozen section evaluation in the diagnosis of adnexal mass. Int J Gynecol Cancer. 2008 Mar-Apr;18(2):235-40.
67. Tempe A, Singh S, Wadhwa L, Garg A. Conventional and color Doppler sonography in preoperative assessment of ovarian tumors. Int J Gynaecol Obstet. 2006 Jan;92(1):64-8.

68. Testa AC, Timmerman D, Exacoustos C, Fruscella E, Van Holsbeke C, Bokor D, et al. The role of CnTI-SonoVue in the diagnosis of ovarian masses with papillary projections: a preliminary study. Ultrasound Obstet Gynecol. 2007 May;29(5):512-6.
69. Timmerman D, Testa AC, Bourne T, Ferrazzi E, Ameye L, Konstantinovic ML, et al. 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 Dec 1;23(34):8794-801.
70. Tongsong T, Wanapirak C, Sukpan K, Khunamornpong S, Pathumbal A. Subjective sonographic assessment for differentiation between malignant and benign adnexal masses. Asian Pac J Cancer Prev. 2007 Jan-Mar;8(1):124-6.
71. Topuz S, Saygili H, Akhan S, Yavuz E, Turfanda A, Berkman S. Differentiation of benign and malignant adnexal masses: value of a morphologic scoring system. Eur J Gynaecol Oncol. 2005;26(2):209-12.
72. Tsili AC, Tsampoulas C, Argyropoulou M, Navrozoglou I, Alamanos Y, Paraskevaidis E, et al. Comparative evaluation of multidetector CT and MR imaging in the differentiation of adnexal masses. Eur Radiol. 2008 May;18(5):1049-57.
73. Umemoto M, Shiota M, Shimono T, Hoshiai H. Preoperative diagnosis of ovarian tumors, focusing on the solid area based on diagnostic imaging. J Obstet Gynaecol Res. 2006 Apr;32(2):195-201.
74. Valentin L, Ameye L, Jurkovic D, Metzger U, Lecuru F, Van Huffel S, et al. 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 Apr;27(4):438-44.
75. Van Calster B, Timmerman D, Bourne T, Testa AC, Van Holsbeke C, Domali E, et al. Discrimination between benign and malignant adnexal masses by specialist ultrasound examination versus serum CA-125. J Natl Cancer Inst. 2007 Nov 21;99(22):1706-14.
76. Van Calster B, Timmerman D, Lu C, Suykens JAK, Valentin L, Van Holsbeke C, et al. Preoperative diagnosis of ovarian tumors using Bayesian kernel-based methods. Ultrasound Obstet Gynecol. 2007 May;29(5):496-504.
77. Van Holsbeke C, Van Calster B, Valentin L, Testa AC, Ferrazzi E, Dimou I, et al. External validation of mathematical models to distinguish between benign and malignant adnexal tumors: a multicenter study by the International Ovarian Tumor Analysis Group. Clin Cancer Res. 2007 Aug 1;13(15 Pt 1):4440-7.
78. Wanapirak C, Srisupundit K, Tongsong T. Sonographic morphology scores (SMS) for differentiation between benign and malignant adnexal masses. Asian Pac J Cancer Prev. 2006 Jul-Sep;7(3):407-10.
79. Wasinghon P, Suthippintawong C, Tuipae S. The accuracy of intraoperative frozen sections in the diagnosis of ovarian tumors. J Med Assoc Thai. 2008 Dec;91(12):1791-5.
80. Wilson WD, Valet AS, Andreotti RF, Green-Jarvis B, Lyshchik A, Fleischer AC. Sonographic quantification of ovarian tumor vascularity. J Ultrasound Med. 2006 Dec;25(12):1577-81.
81. Wootipoom V, Dechsukhum C, Hanprasertpong J, Lim A. Accuracy of intraoperative frozen section in diagnosis of ovarian tumors. J Med Assoc Thai. 2006 May;89(5):577-82.
82. Yarandi F, Eftekhar Z, Izadi-Mood N, Shojaei H. Accuracy of intraoperative frozen section in the diagnosis of ovarian tumors. Aust NZ J Obstet Gynaecol. 2008 Aug;48(4):438-41.
83. Yazbek J, Raju KS, Ben-Nagi J, Holland T, Hillaby K, Jurkovic D. Accuracy of ultrasound subjective 'pattern recognition' for the diagnosis of borderline ovarian tumors. Ultrasound Obstet Gynecol. 2007 May;29(5):489-95.

84. Zhang M, Cheung MK, Shin JY, Kapp DS, Husain A, Teng NN, et al. Prognostic factors responsible for survival in sex cord stromal tumors of the ovary--an analysis of 376 women. Gynecol Oncol. 2007 Feb;104(2):396-400.
85. Chan JK, Munro EG, Cheung MK, Husain A, Teng NN, Berek JS, et al. Association of lymphadenectomy and survival in stage I ovarian cancer patients. Obstet Gynecol. 2007 Jan;109(1):12-9.
86. Chan J, Fuh K, Shin J, Cheung M, Powell C, Chen L, et al. The treatment and outcomes of early-stage epithelial ovarian cancer: have we made any progress? Br J Cancer. 2008 Apr 8;98(7):1191-6.
87. Cho Y-H, Kim D-Y, Kim J-H, Kim Y-M, Kim K-R, Kim Y-T, et al. Is complete surgical staging necessary in patients with stage I mucinous epithelial ovarian tumors? Gynecol Oncol. 2006 Dec;103(3):878-82.
88. Hornung R, Urs E, Serenella E, Edward W, Ursula S, Urs H, et al. Analysis of potential prognostic factors in 111 patients with ovarian cancer. Cancer Lett. 2004 Mar 31;206(1):97-106.
89. Maggioni A, Benedetti Panici P, Dell'Anna T, Landoni F, Lissoni A, Pellegrino A, et al. Randomised study of systematic lymphadenectomy in patients with epithelial ovarian cancer macroscopically confined to the pelvis. Br J Cancer. 2006 Sep 18;95(6):699-704.
90. Oksefjell H, Sandstad B, Trope C. Is the watch and wait approach adequate after comprehensive surgical staging in invasive stage I epithelial ovarian cancer? The Norwegian Radium Hospital experience. Eur J Gynaecol Oncol. 2008;29(6):583-9.
91. Skirnisdottir I, Sorbe B. Lymph node sampling is of prognostic value in early stage epithelial ovarian carcinoma. Eur J Gynaecol Oncol. 2005;26(2):181-5.
92. Suzuki S, Kajiyama H, Shibata K, Ino K, Nawa A, Sakakibara K, et al. Is there any association between retroperitoneal lymphadenectomy and survival benefit in ovarian clear cell carcinoma patients? Ann Oncol. 2008 Jul;19(7):1284-7.
93. Wong HF, Low JJH, Chua Y, Busmanis I, Tay EH, Ho TH. Ovarian tumors of borderline malignancy: a review of 247 patients from 1991 to 2004. Int J Gynecol Cancer. 2007 Mar-Apr;17(2):342-9.
94. Desfeux P, Camatte S, Chatellier G, Blanc B, Querleu D, Lecuru F. Impact of surgical approach on the management of macroscopic early ovarian borderline tumors. Gynecol Oncol. 2005 Sep;98(3):390-5.
95. Ghezzi F, Cromi A, Uccella S, Bergamini V, Tomera S, Franchi M, et al. Laparoscopy versus laparotomy for the surgical management of apparent early stage ovarian cancer. Gynecol Oncol. 2007 May;105(2):409-13.
96. Lecuru F, Desfeux P, Camatte S, Bissery A, Blanc B, Querleu D. Impact of initial surgical access on staging and survival of patients with stage I ovarian cancer. Int J Gynecol Cancer. 2006 Jan-Feb;16(1):87-94.
97. Park JY, Bae J, Lim MC, Lim SY, Seo SS, Kang S, et al. Laparoscopic and laparotomic staging in stage I epithelial ovarian cancer: a comparison of feasibility and safety. Int J Gynecol Cancer. 2008 Nov-Dec;18(6):1202-9.
98. Romagnolo C, Gadducci A, Sartori E, Zola P, Maggino T. Management of borderline ovarian tumors: results of an Italian multicenter study. Gynecol Oncol. 2006 May;101(2):255-60.
99. Park JY, Kim DY, Kim JH, Kim YM, Kim YT, Nam JH. Surgical management of borderline ovarian tumors: The role of fertility-sparing surgery. Gynecol Oncol. 2009 Apr;113(1):75-82.
100. Yinon Y, Beiner ME, Gotlieb WH, Korach Y, Perri T, Ben-Baruch G. Clinical outcome of cystectomy compared with unilateral salpingo-oophorectomy as fertility-sparing treatment of borderline ovarian tumors. Fertil Steril. 2007 Aug;88(2):479-84.

101. Bossuyt PM, Leeflang MM. Developing criteria for including studies. In: Cochrane handbook for systematic reviews of diagnostic test accuracy. Version 0.4. The Cochrane Collaboration; 2008 [updated 2008 Sep].
102. Reeves BC, Deeks JJ, Higgins JPT, Wells GA. Chapter 13: Including non-randomized studies. In: Cochrane handbook for systematic reviews of interventions. Version 5.0.2. The Cochrane Collaboration; 2009 [updated 2009 Sep].
103. Valentin L. Prospective cross-validation of Doppler ultrasound examination and gray-scale ultrasound imaging for discrimination of benign and malignant pelvic masses. Ultrasound Obstet Gynecol 1999;14(4):273-83.
104. Timmerman D, Schwärzler P, Collins WP, Claerhout F, Coenen M, Amant F, et al. Subjective assessment of adnexal masses with the use of ultrasonography: an analysis of interobserver variability and experience. Ultrasound Obstet Gynecol 1999 13(1):11-6.
105. 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 Jul;78(1):70-6.
106. Alcazar JL, Merce LT, Laparte C, Jurado M, Lopez-Garcia G. A new scoring system to differentiate benign from malignant adnexal masses. Am J Obstet Gynecol. 2003 Mar;188(3):685-92.
107. Schneider VL, Schneider A, Reed KL, Hatch KD. Comparison of Doppler with two-dimensional sonography and CA 125 for prediction of malignancy of pelvic masses. Obstet Gynecol. 1993 Jun;81(6):983-8.
108. Buckshee K, Temsu I, Bhatla N, Deka D. Pelvic examination, transvaginal ultrasound and transvaginal color Doppler sonography as predictors of ovarian cancer. Int J Gynaecol Obstet. 1998 Apr;61(1):51-7.
109. 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 Sep;10(3):192-7.
110. Rehn M, Lohmann K, Rempen A. Transvaginal ultrasonography of pelvic masses: evaluation of B-mode technique and Doppler ultrasonography. Am J Obstet Gynecol. 1996 Jul;175(1):97-104.
111. Alcazar JL, Lopez-Garcia G. Transvaginal color Doppler assessment of venous flow in adnexal masses. Ultrasound Obstet Gynecol. 2001 May;17(5):434-8.
112. Alcazar JL, Ruiz-Perez ML, T E. Transvaginal color Doppler sonography in adnexal masses: which parameter performs best? Ultrasound Obstet Gynecol. 1996;8(2):114-9.
113. Leeners B, Schild RL, Funk A, Hauptmann S, Kemp B, Schroder W, et al. Colour Doppler sonography improves the pre-operative diagnosis of ovarian tumours made using conventional transvaginal sonography. Eur J Obstet Gynecol Reprod Biol. 1996 Jan;64(1):79-85.
114. Timor-Tritsch LE, Lerner JP, Monteagudo A, Santos R. Transvaginal ultrasonographic characterization of ovarian masses by means of color flow-directed Doppler measurements and a morphologic scoring system. Am J Obstet Gynecol. 1993 Mar;168(3 Pt 1):909-13.
115. Mol BW, Boll D, De Kanter M, Heintz AP, Sijmons EA, Oei SG, et al. Distinguishing the benign and malignant adnexal mass: an external validation of prognostic models. Gynecol Oncol. 2001 Feb;80(2):162-7.
116. Alcázar JL, Errasti T, Laparte C, Jurado M, G L-G. Assessment of a new logistic model in the preoperative evaluation of adnexal masses. J Ultrasound Med. 2001;20(8):841-8.
117. Botta G, Zarcone R. Trans-vaginal ultrasound examination of ovarian masses in premenopausal women. Eur J Obstet Gynecol Reprod Biol. 1995;62(1):37-41.

118. 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 Jan;170(1 Pt 1):81-5.
119. DePriest PD, Shenson D, Fried A, Hunter JE, Andrews SJ, Gallion HH, et al. A morphology index based on sonographic findings in ovarian cancer. Gynecol Oncol. 1993 Oct;51(1):7-11.
120. Finkler NJ, Benacerraf B, Lavin PT, Wojciechowski C, Knapp RC. Comparison of serum CA 125, clinical impression, and ultrasound in the preoperative evaluation of ovarian masses. Obstet Gynecol. 1988 Oct;72(4):659-64.
121. Schutter EM, Kenemans P, Sohn C, Kristen P, Crombach G, Westermann R, et al. Diagnostic value of pelvic examination, ultrasound, and serum CA 125 in postmenopausal women with a pelvic mass. An international multicenter study. Cancer. 1994 Aug 15;74(4):1398-406.
122. Schutter EM, Sohn C, Kristen P, Mobus V, Crombach G, Kaufmann M, et al. Estimation of probability of malignancy using a logistic model combining physical examination, ultrasound, serum CA 125, and serum CA 72-4 in postmenopausal women with a pelvic mass: an international multicenter study. Gynecol Oncol. 1998 Apr;69(1):56-63.
123. Bromley B, Goodman H, Benacerraf BR. Comparison between sonographic morphology and Doppler waveform for the diagnosis of ovarian malignancy. Obstet Gynecol. 1994 Mar;83(3):434-7.
124. Benacerraf BR, Finkler NJ, Wojciechowski C, RC K. Sonographic accuracy in the diagnosis of ovarian masses. J Reprod Med. 1990;35(5):491-5.
125. Jacobs I, Oram D, Fairbanks J, Turner J, Frost C, Grudzinskas JG. A risk of malignancy index incorporating CA 125, ultrasound and menopausal status for the accurate preoperative diagnosis of ovarian cancer. Br J Obstet Gynaecol. 1990 Oct;97(10):922-9.
126. Tingulstad S, Hagen B, Skjeldestad FE, Onsrud M, Kiserud T, Halvorsen T, et al. Evaluation of a risk of malignancy index based on serum CA125, ultrasound findings and menopausal status in the pre-operative diagnosis of pelvic masses. Br J Obstet Gynaecol. 1996 Aug;103(8):826-31.
127. Tingulstad S, Hagen B, Skjeldestad FE, Halvorsen T, Nustad K, Onsrud M. The risk-of-malignancy index to evaluate potential ovarian cancers in local hospitals. Obstet Gynecol. 1999 Mar;93(3):448-52.
128. Manjunath AP, Pratapkumar, Sujatha K, Vani R. Comparison of three risk of malignancy indices in evaluation of pelvic masses. Gynecol Oncol. 2001 May;81(2):225-9.
129. Tailor A, Jurkovic D, Bourne TH, Collins WP, Campbell S. Sonographic prediction of malignancy in adnexal masses using an artificial neural network. Br J Obstet Gynaecol. 1999 Jan;106(1):21-30.
130. Timmerman D, Verrelst H, Bourne TH, De Moor B, Collins WP, Vergote I, et al. Artificial neural network models for the preoperative discrimination between malignant and benign adnexal masses. Ultrasound Obstet Gynecol. 1999 Jan;13(1):17-25.
131. Timmerman D, Bourne TH, Tailor A, Collins WP, Verrelst H, Vandenberghe K, et al. A comparison of methods for preoperative discrimination between malignant and benign adnexal masses: the development of a new logistic regression model. Am J Obstet Gynecol. 1999 Jul;181(1):57-65.
132. 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):357-65.

133. Kupesic S, Plavsic B. Early ovarian cancer: 3-D power Doppler. Abdom Imaging. 2006;31(5):613-9.
134. DePriest PD, Varner E, Powell J, Fried A, Puls L, Higgins R, et al. The efficacy of a sonographic morphology index in identifying ovarian cancer: a multi-institutional investigation. Gynecol Oncol. 1994 Nov;55(2):174-8.
135. Zhang Z, Yu Y, Xu F, Berchuck A, van Haaften-Day C, Havrilesky LJ, et al. Combining multiple serum tumor markers improves detection of stage I epithelial ovarian cancer. Gynecol Oncol. 2007 Dec;107(3):526-31.
136. American College of Obstetricians and Gynecologists Committee on Gynecologic Practice. The role of the generalist obstetrician-gynecologist in the early detection of ovarian cancer. Gynecol Oncol. 2002 Dec;87(3):237-9.
137. Park JY, Kim DY, Suh DS, Kim JH, Kim YM, Kim YT, et al. Comparison of laparoscopy and laparotomy in surgical staging of early-stage ovarian and fallopian tubal cancer. Ann Surg Oncol. 2008 Jul;15(7):2012-9.
138. Le T, Giede C, Salem S, Lefebvre G, Rosen B, Bentley J, et al. Initial evaluation and referral guidelines for management of pelvic/ovarian masses. J Obstet Gynaecol Can. 2009 Jul;31(7):668-80.
139. Nguyen H, Averette H, Hoskins W, Penalver M, Sevin B, Steren A. National survey of ovarian carcinoma. Part V. The impact of physician's specialty on patients' survival. Cancer. 1993;72(12).
140. Mayer AR CS, Graves E, Holm C, Tseng PC, Nelson BE, Schwartz PE. Ovarian cancer staging: does it require a gynecologic oncologist? Gynecol Oncol 1992;47(2).
141. 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 Gynecol Reprod Biol. 1997 Mar;72(1):63-72.
142. Leeflang MM, Deeks JJ, Gatsonis C, Bossuyt PM, Group CDTAW. Systematic reviews of diagnostic test accuracy. Ann Intern Med. 2008;149:889-97.

Appendix 1. Environmental scan.

Websites reviewed

Scottish Intercollegiate Guidelines Network (SIGN)
National Institute for Clinical Excellence (NICE)
Agency for Healthcare Research and Quality (AHRQ)
Cancer Society of New Zealand
Cancer Care Ontario (CCO) clinical practice guidelines – Gynecology Disease Site group
British Columbia Cancer Agency
Nova Scotia Cancer Agency
Cochrane Reviews
American Society of Clinical Oncology (ASCO)
American Society for Therapeutic Radiology and Oncology (ASTRO)
American Cancer Society
National Guidelines Clearinghouse
National Comprehensive Cancer Network (NCCN)

Appendix 2. Updated literature search.

LITERATURE SEARCH STRATEGIES:

Question 1: Identification of an adnexal mass suspicious for malignancy

Update of AHRQ report literature search:

A) Pelvic exam performance

1) Pelvic exam.mp.
2) (bimanual adj pelvic).mp.
3) (physical exam and pelvis).mp.
4) (diagnostic techniques, obstetrical and gynecological"/ or culdoscopy/ or laparoscopy/ or physical examination/
5) Physical examination/
6) Ovarian cysts/ or ovarian neoplasms/ or genital neoplasms, Female/ or adnexal diseases/ or adnexal mass.mp.
7) Exp ovarian cysts/ or exp ovarian neoplasms/ or genital neoplasms, female/ or adnexal diseases/ or adnexal mass.mp.
8) Exp fallopian tube diseases/
9) 5 and (7 or 8)
10) (or/1-3) and (or/7-8)
11) 9 and 10
12) "diagnostic techniques, obstetrical and gynecological"/ and (or/7-8)
13) Culdoscopy/ and (or/7-8)
14) Or/1-3, 9-10
15) Limit 14 to (human and English language and yr=2004-2009)

B) Test performance

1) (vagin$ adj ultraso$).mp.
2) (adnex$ adj2 mas$).mp.
3) (pelvi$ adj mas$).mp.
4) (ovar$ adj mas$).mp.
5) Or/2-4
6) "sensitivity and specificity"/
7) 6 and 1
8) 6 and 5
9) 7 or 8
10) Limit 9 to (human and English language)
11) (ovar$ adj tumo$).mp.
12) 11 and 6
13) ROC curve/
14) 12 and 13
15) 14 not 10
16) 10 or 15
17) Limit 16 to yr=2004-2009

C) Predictive Models

1) (vagin$ adj ultraso$).mp.
2) (adnex$ adj2 mas$).mp.

3) (pelvi$ adj mas$).mp.
4) (ovar$ adj mas$).mp.
5) Or/2-4
6) "sensitivity and specificity"/
7) 6 and 1
8) 6 and 5
9) 7 or 8
10) Limit 9 to (human and English language)
11) Predictive value of tests/
12) Risk assessment/
13) ROC curve/
14) "multivariate analysis"/
15) Or/11-14
16) 15 and 5
17) 16 not 9
18) Limit 17 to (human and English language and yr=2004-2009)

Question 2: Surgical management of an adnexal mass suspicious for malignancy

Update of ACN literature search (approximated) supplemented with keywords from WG:

1) Pelvic mass.tw,ti.
2) Adnexal mass.tw,ti.
3) Exp pelvic neoplasms/
4) (ovary and (cancer or carcinoma or neoplasm?)).tw,ti.
5) (ovarian and (cancer or carcinoma or neoplasm?)).tw,ti.
6) Exp ovarian neoplasms/
7) Borderline ovarian tumo?r$.tw,ti.
8) Tumo?r$ of low malignant potential.tw,ti.
9) Or/1-8
10) Intraoperative pathological examination$.tw,ti.
11) Exp frozen section/
12) Frozen section$.tw,ti.
13) Debulking surgery.tw,ti.
14) Exp surgical procedures, operative/
15) Cytoreduction$.tw,ti.
16) Fertility conservation,tw,ti.
17) Fertility sparing.tw,ti.
18) Surgical management.tw,ti.
19) Secondary cytoreduction.tw,ti.
20) Interval cytoreduction.tw,ti.
21) Surgical staging.tw,ti.
22) BSO.tw,ti.
23) Bilateral salpingo-oophorectomy.tw,ti.
24) Bilateral salpingo oophorectomy.tw,ti.
25) Surgical stage.tw,ti.
26) Total hysterectomy.tw,ti.
27) Hysterectomy.tw,ti.
28) ((node or nodal) and dissection).tw,ti.

29) Exp laparotomy/
30) Exp laparoscopy/
31) Or/10-30
32) 31 and 9
33) Limit 32 to English
34) Limit 33 to humans
35) Limit 34 to yr="2004-current"
36) Letter.pt.
37) Comment.pt.
38) Editorial.pt.
39) News.pt.
40) Case report$.pt.
41) Or/36-40
42) 35 not 41

Evidence-Based Series 4-15 Version 2: Section 3

Management of a Suspicious Adnexal Mass: EBS Development Methods and External Review Process

J Dodge, A Covens, C Lacchetti, L Elit, T Le, M Devries-Aboud, M Fung Kee Fung and the Gynecology Cancer Disease Site Group

A Quality Initiative of the
Program in Evidence-Based Care (PEBC), Cancer Care Ontario (CCO)

Report Date: July 7, 2011

These guideline recommendations have been ENDORSED, which means that the recommendations are still current and relevant for decision making.

Please see Section 4: Document Review Summary and Tool for a summary of updated evidence published between 2001 and 2013, and for details on how this Clinical Practice Guideline was ENDORSED

THE PROGRAM IN EVIDENCE-BASED CARE

The Program in Evidence-based Care (PEBC) is an initiative of the Ontario provincial cancer system, Cancer Care Ontario (CCO) (1). The PEBC mandate is to improve the lives of Ontarians affected by cancer, through the development, dissemination, implementation, and evaluation of evidence-based products designed to facilitate clinical, planning, and policy decisions about cancer care.

The PEBC supports a network of disease-specific panels, termed Disease Site Groups (DSGs), as well as other groups or panels called together for a specific topic, all mandated to develop the PEBC products. These panels are comprised of clinicians, other health care providers and decision makers, methodologists, and community representatives from across the province.

The PEBC is well known for producing evidence-based guidelines, known as Evidence-based Series (EBS) reports, using the methods of the Practice Guidelines Development Cycle (1, 2). The EBS report consists of an evidentiary base (typically a systematic review), an interpretation of and consensus agreement on that evidence by our Groups or Panels, the resulting recommendations, and an external review by Ontario clinicians and other stakeholders in the province for whom the topic is relevant. The PEBC has a formal

standardized process to ensure the currency of each document, through the periodic review and evaluation of the scientific literature and, where appropriate, the integration of that literature with the original guideline information.

The Evidence-Based Series

Each EBS is comprised of three sections:

- *Section 1: Guideline Recommendations*. Contains the clinical recommendations derived from a systematic review of the clinical and scientific literature and its interpretation by the Group or Panel involved and a formalized external review in Ontario by review participants.
- *Section 2: Evidentiary Base*. Presents the comprehensive evidentiary/systematic review of the clinical and scientific research on the topic and the conclusions reached by the Group or Panel.
- *Section 3: EBS Development Methods and External Review Process*. Summarizes the EBS development process and the results of the formal external review of the draft version of Section 1: Guideline Recommendations and Section 2: Evidentiary Base.

DEVELOPMENT OF THIS EVIDENCE-BASED SERIES

Development and Internal Review

This EBS was developed by the Gynecology DSG of the CCO PEBC. The series is a convenient and up-to-date source of the best available evidence on management of a suspicious adnexal mass, developed through review of the evidentiary base, evidence synthesis, and input from external review participants in Ontario.

Report Approval Panel

Prior to the submission of this EBS draft report for external review, the report was reviewed and approved by the PEBC Report Approval Panel, which consists of two members, including an oncologist, with expertise in clinical and methodology issues. The key issues raised by the Report Approval Panel and the modifications made by the Gynecology DSG are below with the DSG responses bulleted below the comments of the panel members:

- If pathology is still the gold standard, as suggested in your last qualifying statement, what is the role of the other diagnostic technologies?
    - The need to have clarity in the preoperative diagnosis is to 1) triage the patients to a specialist, and 2) better define the surgical options (e.g., conservative vs. radical). We have now added a paragraph to the discussion regarding the intention and necessity of preoperative diagnosis of an adnexal mass.
- The authors' first recommendation, and the one for which the most detailed analysis exists, concludes that 3D US, CT and MRI are "all recommended" with considerations of more "local factors" then suggested as determinants of the modality of choice. The authors should reconsider whether they have missed an opportunity to make a more definitive recommendation that accounts for the "equality" in diagnostic efficacy and what can be reasonably assumed about cost, access, harm (e.g., radiation exposure) and patient inconvenience.
    - We thank the reviewer for his comments. We have taken this opportunity to add further recommendations based on the Working Group's expert consensus opinion that takes into consideration availability, access, and harm.
- The authors consider various diagnostic tools separately (e.g., imaging, CA-125). Is there a risk that, in practice, these modalities are used in combination and in doing so, diagnostic properties are changed? Related to this theme, are there important

differences in the eligibility of patients included in any analysis of a single modality, in which a second modality criterion was required for inclusion? The bolded text was removed from the recommendations section, and it was further delineated that the recommendations were based on expert opinion. Furthermore, the recommendation on PET was moved to indicate that the role of PET in cervical cancer has not been prospectively evaluated.
    - In many circumstances, we have evaluated the combination of diagnostic tests such as ultrasound and CA-125
- In contrast to the diagnostic efficacy section, the section that deals with "therapy" (*What is the most appropriate surgical procedure for a woman who presents with an adnexal mass suspicious for malignancy?*) does not include conventional guideline methodology or reporting. As this question is a largely a therapeutic question, standard methodologies would move down a hierarchy of acceptable guidelines, meta-analysis of randomized controlled trials (RCTs), RCTs, non-randomized comparisons, etc. The authors do not provide background in their Methods section about their approach to this question (including literature search criteria). The Results section largely is divided into positive vs. negative studies as opposed to a systematic presentation of the literature. As a result, the juxtaposition of data and conclusions is associated with ambiguity and comes across as potentially representing "expert opinion" as opposed to an evidenced-based recommendation. The authors should reconsider their approach to this question.
    - We thank the reviewer for these comments. We did indeed use standard methodologies in the development of this section of the guideline but acknowledge that this was not clear in the reporting. As such, we have explained more explicitly in the Methods section what was done. Furthermore, we rearranged the reporting of the results so that the one RCT is discussed first.
- The authors might wish to clarify whether the post-diagnostic therapeutic pathway includes multiple modalities that require systematic review in order to assess linkage. Again, while this comment may reflect lack of content expertise, if these patients should be considered for adjuvant therapy using another modality, does the use of that subsequent modality interact with the choice of the surgical approach? In addressing the surgical approach, and especially if best therapy involves sequential use of multiple modalities, the authors may need to carefully dissect out the prognostic importance of patients being able to undergo a more extensive surgical procedure vs. the therapeutic benefit of that intervention.
    - There is indeed a link between surgery and adjuvant therapy. However, the Working Group viewed these two interventions separately, both in terms of prognosis and as therapeutic approaches. As such, a discussion on adjuvant therapy in patients with a suspicious adnexal mass was deemed beyond the scope of this document.

External Review by Ontario Clinicians and Other Experts

The PEBC external review process is two-pronged and includes a targeted peer review that is intended to obtain direct feedback on the draft report from a small number of specified content experts and a professional consultation that is intended to facilitate dissemination of the final guidance report to Ontario practitioners.

Following the review and discussion of Section 1: Recommendations and Section 2: Evidentiary Base of this EBS and the review and approval of the report by the PEBC Report Approval Panel, the Gynecology Cancer DSG circulated Sections 1 and 2 to external review

participants for review and feedback. Box 1 summarizes the draft recommendations and supporting evidence developed by the working group.

BOX 1:
DRAFT RECOMMENDATIONS (approved for external review 2011-04-08)
QUESTIONS

1. What is the optimal strategy for preoperative identification of the adnexal mass suspicious for ovarian cancer?
2. What is the most appropriate surgical procedure for a woman who presents with an adnexal mass suspicious for ovarian cancer?

TARGET POPULATION

The target population of this guideline is adult women presenting with a suspicious adnexal mass, either symptomatic or asymptomatic.

INTENDED USERS

This guideline is targeted for clinicians managing the care of women with a suspicious adnexal mass, specifically general gynecologists and gynecological oncologists.

RECOMMENDATIONS
Identification of an Adnexal Mass Suspicious for Ovarian Cancer

- **Sonography, particularly three-dimensional (3D) sonography, magnetic resonance imaging (MRI), and computerized tomography (CT) imaging are each recommended for differentiating malignant from benign ovarian masses. However, the working group offers the following further recommendations, based on their expert consensus opinion and the consideration of availability, access, and harm:**
  - **Transvaginal sonography should be the first modality of choice, where technically feasible, in patients with a suspicious, isolated ovarian mass.**
  - **MRI is the most appropriate test to help clarify the malignant potential in patients where ultrasound may be unreliable.**
  - **CT is most useful in cases where extra ovarian disease is suspected or needs to be ruled out.**

  Key Evidence
  - The diagnostic performance of each diagnostic technology was compared and contrasted based on the summary data on sensitivity and specificity obtained from the meta-analysis.
  - A meta-analysis of six cohort studies that investigated 3D sonography (1-6) indicated an enhanced sensitivity of 93.5% and specificity of 91.5% with 3D technology (Section 2, Figure 2A).
  - A meta-analysis of 22 cohort studies with 24 data sets that investigated the effectiveness of MRI in the diagnosis of adnexal masses (7-28) found an overall sensitivity of 91.9% and specificity of 88.4% (Section 2, Figure 2A).
  - A meta-analysis of seven studies with eight data sets considering CT technology (2,10,12,14,22,29,30) yielded an overall sensitivity of 87.2% and specificity of 84.0% (Section 2, Figure 2A).

- **Evaluation of an adnexal mass by Doppler technology alone is not recommended. Doppler technology should be combined with a morphological assessment.**

Key Evidence

- This recommendation is based on the results of several meta-analyses on Doppler indices but not direct comparisons between them. Rather, the summary data from these meta-analyses were inspected, and reasonable sensitivities and specificities were noted.
- A meta-analysis of the resistance index (RI) included 35 cohort studies (2,5,17,30-61) with 42 data sets and yielded an overall sensitivity of 77.2% and specificity of 89.8% (Section 2, Figure 2C).
- A meta-analysis of 21 cohort studies with 22 data sets that evaluated the Pulsatility Index (PI) found an overall sensitivity of 80.6% and specificity of 79.9% (Section 2, Figure 2C).
- A meta-analysis of the peak systolic velocity (PSV) included seven cohort studies (32,33,37,42,50,51,62) and found an overall sensitivity of 80.0% and specificity of 84.2% (Section 2, Figure 2C).

Qualifying Statement

- Assessment of an adnexal mass by colour Doppler technology, using the RI, PI and PSV indices, was neither as sensitive nor specific as simple ultrasonography. Furthermore, because of the overlap of vascular parameters between malignant and benign masses, a firm diagnosis based on Doppler evaluation alone can be problematic.

➢ **Morphological scoring systems can be used to differentiate benign from malignant adnexal masses. The choice between the scoring systems should be made based on demonstrated accuracy and clinician preference.**

Key Evidence

- The diagnostic performance of each morphological scoring system was compared and contrasted based on the summary data on sensitivity and specificity obtained from the meta-analysis.
- A meta-analysis of 17 cohort studies (32,33,46,52,57,63-74) assessing the Sassone model at a cutoff of 9, found an overall sensitivity of 88.6% and specificity of 77.5% (Section 2, Figure 2B).
- A meta-analysis of nine studies (33,35,63,65,70,75-78) considering the Ferrazzi scoring system reported an overall sensitivity and specificity of 85.2% and 85.9%, respectively (Section 2, Figure 2B).
- A meta-analysis of five studies (36,79-82) measuring the performance of the Finkler scoring system found an overall sensitivity of 83.5% and specificity of 78.2% (Section 2, Figure 2B).
- A meta-analysis of data from 13 RMI studies (70,83-94), with 15 data sets, employing a cutoff of 200 to be indicative of malignancy, reported the summary sensitivity and specificity were 79.2% and 91.7%, respectively (Section 2, Figure 2B).

➢ **As a stand-alone modality, serum CA-125 is not recommended for distinguishing between benign and malignant adnexal masses.**

Key Evidence

- This recommendation is based on a meta-analysis of 49 cohort studies (17,31,35,39,52, 62,75,79-81,85,88,92,93,95-129) and two case-control studies (130,131) with a total of 52 data sets that found, at a threshold of 35 U/mL, an

overall sensitivity of 78.7% and specificity of 77.9% (Section 2, Figure 2D).

Qualifying Statement

- CA-125 values are of limited use in premenopausal women and elevated in only 50% of early-stage ovarian cancers. Caution should be used in interpreting values in such patients.

➢ **Frozen section for the intraoperative diagnosis of a suspicious adnexal mass is recommended in settings where availability and patient preferences allow.**

Key Evidence

- This recommendation is based on a meta-analysis of frozen section diagnoses that included 15 cohort studies (7,132-145) and yielded an overall sensitivity of 89.2% and specificity of 97.9% (Section 2, Figure 2D).

Surgical Procedures for an Adnexal Mass Suspicious for Malignancy

➢ **Comprehensive surgical staging with lymphadenectomy is recommended for the surgical management of patients with early-stage ovarian cancer to improve survival.**

Key Evidence

- This recommendation is based on the results of five retrospective cohort studies (146-150).
- Two large population-based studies (146,147) found that surgical staging with lymphadenectomy was associated with improved three-year ($p<0.001$) (147) and five-year disease-specific survival ($p<0.001$) (146) compared to staging procedures without lymphadenectomy.
- Oksefjell et al (148) reported a statistically significant improvement in five-year overall survival rates in patients undergoing a lymphadenectomy versus those that did not (87% versus (vs.) 64%, respectively; $p=0.02$).
- Survival analyses performed by both Skirnisdottir et al (149) and Hornung et al (150) also demonstrated a statistically significant benefit in disease-free survival ($p=0.004$ and $p=0.0007$, respectively) for patients undergoing a lymphadenectomy versus those that did not.
- Hornung and colleagues (150) also considered overall survival and reported a statistically significant difference ($p=0.0008$) in the two patient groups in favour of the patients undergoing a lymphadenectomy.
- One randomized controlled trial (RCT) (151) was identified and reported no statistically significant effect of lymphadenectomy on progression-free (hazard ratio (HR), 0.72; 95% confidence interval (CI), 0.46 to 1.14) or overall (HR, 0.85; 95% CI, 0.49 to 1.47)) survival. However, this study was underpowered to detect a difference in survival, the study's secondary outcome. Rather, the sample-size calculation for this RCT was undertaken to detect a difference in prevalence of lymph node positivity. It was deemed inadequate to inform the recommendation.

➢ **Laparoscopy is a reasonable alternative to laparotomy provided appropriate surgery and/or staging can be done. The choice between laparoscopy and laparotomy should be based on patient and clinician preferences. Discussion with a gynecologic oncologist is recommended.**

Key Evidence

- This recommendation is based on the results of six retrospective cohort studies (152-157).
- In the three studies (152,153,156) that considered patients with early epithelial ovarian cancer, no statistical difference in survival rates was detected between patients undergoing a laparoscopy versus laparotomy.
- In the management of patients with early borderline ovarian tumours, Romangnolo et al (155), Park et al (157), and Desfeux et al (154) found that a laparoscopic versus laparotomic surgical approach did not appear to influence survival rates.

➢ **Fertility-preserving surgery is an acceptable alternative to more extensive surgery in patients with low-malignant potential (LMP) tumours and those with well-differentiated surgically staged 1 ovarian cancer. Discussion with a gynecologic oncologist is recommended.**

Key Evidence

- This recommendation is based on two cohort studies that compared the impact of conservative fertility-sparing surgeries versus more radical surgical approaches. Yinon et al (158) specifically compared rates of recurrence in 40 patients who underwent unilateral salpingo-oophorectomy versus 22 patients who underwent cystectomy only. No statistical difference in recurrence rates was detected (27.5% vs. 22.7%, respectively; p=0.8). Similarly, in a larger study of 360 women with LMP tumours, Park et al (157) found no difference in disease-free survival between patients who underwent radical or fertility-sparing surgery (p=0.651).

QUALIFYING STATEMENTS

The Gynecology Cancer Disease Site Group (DSG) acknowledges that, despite definitions and criteria, it is unrealistic to expect that 100% of ovarian cancers will be identified as suspicious preoperatively. Pathology remains the gold standard.

Methods

Targeted Peer Review: During the guideline development process, two targeted peer reviewers from Ontario and one from the USA considered to be clinical and/or methodological experts on the topic were identified by the working group. Several weeks prior to completion of the draft report, the nominees were contacted by email and asked to serve as reviewers. Three reviewers agreed, and the draft report and a questionnaire were sent via email for their review. The questionnaire consisted of items evaluating the methods, results, and interpretive summary used to inform the draft recommendations and whether the draft recommendations should be approved as a guideline. Written comments were invited. The questionnaire and draft document were sent out on April 8, 2011. Follow-up reminders were sent at two weeks (email) and at four weeks (telephone call).

Professional Consultation: Feedback was obtained through a brief online survey of health care professionals who are the intended users of the guideline. Gynecologists and gynecologic oncologists in the PEBC database were contacted by email to inform them of the survey. Participants were asked to rate the overall quality of the guideline (Section 1) and whether they would use and/or recommend it. Written comments were invited. Participants were contacted by email and directed to the survey website where they were provided with access

to the survey, the guideline recommendations (Section 1) and the evidentiary base (Section 2). The notification email was sent on April 13, 2011. The consultation period ended on June 10, 2011. The working group reviewed the results of the survey.

Results

Targeted Peer Review: one reviewer out of the three that were invited provided a response. This reviewer's responses are summarized in Table 1.

Table 1. Responses to nine items on the targeted peer reviewer questionnaire.

	Reviewer Ratings (N=1)				
Question	Lowest Quality (1)	(2)	(3)	(4)	Highest Quality (5)
1. Rate the guideline development methods.					1
2. Rate the guideline presentation.					1
3. Rate the guideline recommendations.					1
4. Rate the completeness of reporting.					1
5. Does this document provide sufficient information to inform your decisions? If not, what areas are missing?					1
6. Rate the overall quality of the guideline report.					1
	Strongly Disagree (1)	(2)	Neutral (3)	(4)	Strongly Agree (5)
7. I would make use of this guideline in my professional decisions.					1
8. I would recommend this guideline for use in practice.					1

9. What are the barriers or enablers to the implementation of this guideline report?
No significant barriers or enablers were reported.

Summary of Written Comments

The reviewer advised that references by L. Cohen and A. Fleischer be added to the evidence base.

Modifications/Actions

The authors were not able to gather more information from this reviewer regarding exactly which publications had been missed. The authors examined whether references by Cohen and/or Fleischer had been considered at any time during the guideline development process. Cohen (2001) was considered by the AHRQ review and reported in Section 2 under "Other Scoring Systems". An additional paper by Cohen and Fleischer was excluded by the AHRQ. One Fleischer paper was included in the evidence base for the guideline. In the end, no modifications to the evidence base were made on the basis of this reviewer comment.

Professional Consultation: Sixty responses were received. Key results of the feedback survey are summarized in Table 2.

Table 2. Responses to four items on the professional consultation survey.

	Number (%)				
General Questions: Overall Guideline Assessment	Lowest Quality (1)	(2)	(3)	(4)	Highest Quality (5)
1. Rate the overall quality of the guideline report.	0 (0)	**0**(0)	6(11)	31(54)	23(41)
	Strongly Disagree (1)	(2)	(3)	(21)	Strongly Agree (5)
2. I would make use of this guideline in my professional decisions.	2(4)	0(0)	6(11)	21(38)	31(55)
3. I would recommend this guideline for use in practice.	0(0)	0(0)	7(13)	22(39)	31(55)

4. What are the barriers or enablers to the implementation of this guideline report?

Practitioners listed barriers to the implementation of this guideline and suggested ways that uptake could be enhanced or enabled. The feedback included the following:

1. Awareness

Some practitioners mentioned that lack of awareness of the guideline would be a barrier to uptake. They recommended making primary care physicians aware of the guideline, and one reviewer suggested presentation of the guideline at group rounds to make the entire team aware of current literature and recommendations.

2. Resources

Many practitioners reported that lack of resources would be a barrier to implementation of the guideline, especially in smaller communities. Specific resources that were noted as lacking included:

- Gynecologic oncologists (e.g., for staging using lymphadectomy).
- MRI only available in larger communities and delays with MRI booking.
- CT access.
- Patients may not always have easy access to an ultrasound department and quality of reporting is variable.
- Pathologists with experience in gynecologic pathology for frozen section diagnoses.
- Regional Minimally Invasive Surgery programs in the community.

Other

In order to enable the implementation of the guideline, practitioners also suggested:

- A summary table outlining the sensitivities and specificities of the various diagnostic procedures.
- Guidance for referral of patients to a gynecologist or gynecologic oncologist for non-gynecologist clinicians who identify an adnexal mass.

- Educating radiologists about appropriate use of ultrasound.
- Boundaries to indicate whether the risk for a specific woman is low or high after being investigated.
- A recommendations regarding CA-125 use for post-op women.

Modifications/Actions/Response

- A summary table outlining the sensitivities and specificities of the various diagnostic procedures is presented in Section 2 of the report.
- Referral criteria are considered to be outside of the scope of this review and guideline. For more information on referral criteria, readers may refer to a guideline published in 2009 by the Society of Obstetricians and Gynaecologists of Canada: *Initial evaluation and referral guidelines for management of pelvic/ovarian masses* (138).
- Cutoffs for the scoring systems included in the report were added to Appendix 2 in Section 1.
- CA-125 use post-op is outside of the scope of this guideline. In the event that the reviewer meant that we should clarify the recommendation regarding CA-125 use as a stand-alone modality in post*menopausal* women, we amended the CA-125 recommendation so that it no longer refers only to premenopausal women, but to all women, as there is no evidence for its efficacy as a stand-alone modality for any age group.

Summary of Written Comments

Twenty of the sixty responders to professional consultation provided additional written comments. The majority indicated that the document was of high quality and would be of use to practitioners. Suggestions for improvements or additions to the document included:

1. Mention of specific suspicious features may prove useful for the radiologist colleague accessing this guideline.
2. There is clear operative requirement of staging and lymphadectomy, but the role of tertiary referral is not clear.
3. Several comments were made relating the scoring systems described in the report. The feedback generally indicated that there are many practitioners in the province who are not aware of these scoring systems. A direct link from the recommendations to the scoring systems was requested. It was also suggested that the guideline recommend one scoring system that would be the most reliable. Other comments related to scoring systems include:
    a. Many practitioners are using the Risk of Malignancy Index (RMI) and it should be available as an appendix to this guideline. Also related to RMI:
        i. Recommendation on the use of RMI II and how the general gynecologist should use it for guidance when considering referral to a tertiary care centre.
        ii. Reference to RMI cutoff of 200 instead of 400 which is what is commonly used.
4. There was also a request for an appendix for ultrasound features of malignancy and definitions of resistance index (RI) pulsatility index (PI), and Peak Systologic (sic) Velocity (PSV).

Modifications/Actions/Response

1. As specific suspicious features are described in Section 2 of the report, no changes were made based on this comment.
2. The topic of referral is beyond the scope of this guideline and was therefore not considered for inclusion. As mentioned above, readers may refer to the Society of

Obstetricians and Gynaecologists of Canada guideline *Initial evaluation and referral guidelines for management of pelvic/ovarian masses* (138).

3. Response to comments related to the scoring systems:
    a. An appendix was added to Section 1 with descriptions of the Sassone, DePriest, Ferrazzi, Lerner, Finkler, and RMI scoring systems. References to the original publications were also added to Section 2 of the report. Because the sensitivity and specificity are comparable for these scoring systems, it is not possible to recommend one of them as most reliable. Therefore, the wording of the recommendation was adjusted to make this clear and to indicate that the choice of scoring system should be based on clinician preference.
    b. The recommendation was also reworded to state that choice of version of the RMI should be based on clinician preference as they all have comparable sensitivity and specificity.
    c. The cutoff value of 200 for the RMI was based on the cutoff used in the initial report by Jacobs et al. The sensitivity and specificity values reported here are based on this cutoff. This was also reported in the AHRQ report as the most common cutoff value, therefore, the group decided to not to include reference to a cutoff of 400.
    d. Other definitions such as RI, PI, and PSV can be found in Section 2 of the document.

Conclusion

This EBS report reflects the integration of feedback obtained through the external review process with final approval given by the Gynecology Cancer Disease Site Group, and the Report Approval Panel of the PEBC. Updates of the report will be conducted as new evidence informing the question of interest emerges.

REFERENCES

1. Browman GP, Levine MN, Mohide EA, Hayward RSA, Pritchard KI, Gafni A, et al. The practice guidelines development cycle: a conceptual tool for practice guidelines development and implementation. J Clin Oncol. 1995;13:502-12. Comment in: Ann Oncol. 2002 Sep;13(9):1507-9; author reply: 1509.
2. Browman GP, Newman TE, Mohide EA, Graham ID, Levine MN, Pritchard KI, et al. Progress of clinical oncology guidelines development using the practice guidelines development cycle: the role of practitioner feedback. J Clin Oncol. 1998;16(3):1226-31.

Ontario
Cancer Care Ontario
Action Cancer Ontario

Evidence-Based Series 4-15 Version 2: Section 4

Management of a Suspicious Adnexal Mass: Document Assessment and Review

J Dodge, A Covens, C Lacchetti, L Elit, T Le, M Devries-Aboud, M Fung Kee Fung, and the Gynecology Cancer Disease Site Group

A Quality Initiative of the
Program in Evidence-Based Care (PEBC), Cancer Care Ontario (CCO)

Review Date: September 9, 2016

The 2011 Guideline recommendations are

ENDORSED

This means that the recommendations are still current and relevant for decision making.

The original version of this guidance document was released by Cancer Care Ontario's Program in Evidence-based Care in 2011.

In September 2015, this document was assessed in accordance with the PEBC Document Assessment and Review Protocol and was determined to require a review. As part of the review, a PEBC methodologist conducted an updated search of the literature. A clinical expert (TL) reviewed and interpreted the new eligible evidence and proposed the existing recommendations could be endorsed. The Gynecology Cancer Disease Site Group (DSG) endorsed the recommendations found in Section 1 (Clinical Practice Guideline) on September 9, 2016.

DOCUMENT ASSESSMENT AND REVIEW RESULTS

Questions Considered

1. What is the optimal strategy for preoperative identification of the adnexal mass suspicious for ovarian cancer?

2. What is the most appropriate surgical procedure for a woman who presents with an adnexal mass suspicious for ovarian cancer

Literature Search and New Evidence
The new search (January 2011 to January 2016) yielded 9386 papers. Three hundred and eighty six were retained for full text review. After the full text review, 93 studies were retained. Brief results of these searches are shown in the Document Review Tool.

Impact on Guidelines and Its Recommendations
The new data supports existing recommendations. Hence, the Gynecology Cancer DSG ENDORSED the 2011 recommendations on management of a suspicious adnexal mass.

A new qualifyling statement was added to the recommendation on ultrasound-based morphological scoring systems. "Ultrasound diagnostic criteria using a set of simple rules to distinguish between benign and malignant masses, and the IOTA (International Ovarian Tumour Analysis) predictive adnexal model had been extensively studied with acceptable sensitivity and specificity. This can serve as potential alternative diagnostic strategy to the RMI score." The following paragraphs are the justification for the inclusion of this qualifying statement.

In 2008 the International Ovarian Tumour Analysis (IOTA) group proposed ultrasound-based criteria for diagnosing benign and malignant adnexal masses (the Simple Rules). These are based on a set of 5 ultrasound features suggestive of benign tumor (B-features) and 5 ultrasound features indicative of a malignancy (M-features). The B-features include: (1) unilocular cyst; (2) presence of solid components where the largest solid component is < 7 mm in largest diameter; (3) presence of acoustic shadows; (4) smooth multilocular tumor less than 100 mm in largest diameter; and (5) no detectable internal blood flow on Doppler examination. The M-features include: (1) irregular solid tumour; (2) ascites; (3) at least four papillary structures;(4) irregular multilocular solid tumour with a largest diameter of at least 100 mm; and (5) very high color content on color Doppler examination. Adnexal masses are classified as benign if only B-features are observed and as malignant if only M-features are observed on ultrasound. If no features are observed or if conflicting/mixed features are present, the Simple Rules cannot classify the tumour as benign or malignant (inconclusive result). Inconclusive result can be classified using subjective assessment by an experienced ultrasound operator. This strategy has been recently validated in 4848 patients with adnexal masses against histopathologic diagnosis with a sensitivity of 99.7% and specificity of 33.7%.
In addition, the IOTA group also constructed an Assessment of Different Neoplasias in the adneXa (ADNEX) model. This is a risk prediction model to differentiate between benign, borderline tumours, stage I invasive, stage II-IV invasive ovarian cancer and secondary metastatic cancer. The ADNEX model consists of three clinical predictors and six ultrasound predictors. The clinical predictors are age (years), serum CA-125 (U/mL) and type of center to which the patient has been referred for ultrasound examination. Type of center was divided into oncology centers versus other hospitals. The ultrasound predictors are the maximal diameter of the adnexal mass (mm), proportion of solid tissue (%), number of papillary projections (0, 1, 2, 3, > 3), presence of more than 10 cyst locules (yes/no), acoustic shadows (yes/no), and presence of ascites (yes/no). The model performed well in differentiating between benign and malignant masses with the area under the receiver operating characteristic curves (AUC) of the ADNEX model of 0.954 (95% confidence interval 0.947 to 0.961) on the development data and 0.943 (0.934 to 0.952) on the validation data. The sensitivity for diagnosis of malignancy was 96.5% and specificity was 71.3% on the validation data set.

DOCUMENT REVIEW TOOL

Number and title of document under review	4-15 **Management of a Suspicious Adnexal Mass**
Current Report Date	July 7 2011
Clinical Expert	Dr. T. Le
Research Coordinator	Nadia Coakley
Date Assessed	July 6, 2016
Approval Date and Review Outcome (once completed)	September 9, 2016 ENDORSED

Original Question(s):

3. What is the optimal strategy for preoperative identification of the adnexal mass suspicious for ovarian cancer?
4. What is the most appropriate surgical procedure for a woman who presents with an adnexal mass suspicious for ovarian cancer?

Target Population:

The target population of this guideline is adult women presenting with a suspicious adnexal mass, either symptomatic or asymptomatic.

Study Section Criteria:

Articles were eligible for inclusion in this systematic review if they were systematic reviews, meta-analyses, clinical practice guidelines, randomized trials, or comparative cohort studies. Studies identified in the update of the AHRQ report literature search were included based on the same inclusion criteria put forth in the AHRQ report (3).

For studies investigating single modality identification of an adnexal mass, the inclusion criteria were:

1) comparison of the test (e.g., bimanual pelvic exam or ultrasound, to histology or negative surgery
2) greater than 20 patients included in study
3) able to construct a 2-by-2 table, which compares the results of the diagnostic test with the definitive histological diagnosis.

For studies investigating the use of multi-modality scoring systems (i.e., RMI), the inclusion criteria were:

1) patients with suspicion of cancer
2) studies with scoring, risk score, combined modality approach
3) assesses predictive value of two or more variables using multivariable model
4) greater than 50 patients included in study.

Studies identified in the update of the Australian Cancer Network (6) guideline were based on the following selection criteria:

1) greater than 20 patients included in study
2) patients with an adnexal mass suspicious for early stage (I-II) malignancy,
3) two-armed (or greater) study design with a comparison of surgical

procedures/techniques/approaches
4) report on at least one of the following outcomes: optimal surgery, overall survival, progression-free or disease-free survival, reduction in the number of surgeries, morbidity, adverse events, quality of life.

Search Details:

LITERATURE SEARCH STRATEGIES:

Question 1: Identification of an adnexal mass suspicious for malignancy

Update of AHRQ report literature search:

Pelvic exam performance
Pelvic exam.mp.
(bimanual adj pelvic).mp.
(physical exam and pelvis).mp.
(diagnostic techniques, obstetrical and gynecological"/ or culdoscopy/ or laparoscopy/ or physical examination/
Physical examination/
Ovarian cysts/ or ovarian neoplasms/ or genital neoplasms, Female/ or adnexal diseases/ or adnexal mass.mp.
Exp ovarian cysts/ or exp ovarian neoplasms/ or genital neoplasms, female/ or adnexal diseases/ or adnexal mass.mp.
Exp fallopian tube diseases/
5 and (7 or 8)
(or/1-3) and (or/7-8)
9 and 10
"diagnostic techniques, obstetrical and gynecological"/ and (or/7-8)
Culdoscopy/ and (or/7-8)
Or/1-3, 9-10
Limit 14 to (human and English language and yr=2004-2009)

Test performance
(vagin$ adj ultraso$).mp.
(adnex$ adj2 mas$).mp.
(pelvi$ adj mas$).mp.
(ovar$ adj mas$).mp.
Or/2-4
"sensitivity and specificity"/
6 and 1
6 and 5
7 or 8
Limit 9 to (human and English language)
(ovar$ adj tumo$).mp.
11 and 6
ROC curve/
12 and 13
14 not 10
10 or 15
Limit 16 to yr=2004-2009

Predictive Models
(vagin$ adj ultraso$).mp.
(adnex$ adj2 mas$).mp.
(pelvi$ adj mas$).mp.
(ovar$ adj mas$).mp.
Or/2-4
"sensitivity and specificity"/
6 and 1
6 and 5
7 or 8
Limit 9 to (human and English language)
Predictive value of tests/
Risk assessment/
ROC curve/
"multivariate analysis"/
Or/11-14
15 and 5
16 not 9
Limit 17 to (human and English language and yr=2004-2009)

Question 2: Surgical management of an adnexal mass suspicious for malignancy

Update of ACN literature search (approximated) supplemented with keywords from WG:

Pelvic mass.tw,ti.
Adnexal mass.tw,ti.
Exp pelvic neoplasms/
(ovary and (cancer or carcinoma or neoplasm?)).tw,ti.
(ovarian and (cancer or carcinoma or neoplasm?)).tw,ti.
Exp ovarian neoplasms/
Borderline ovarian tumo?r$.tw,ti.
Tumo?r$ of low malignant potential.tw,ti.
Or/1-8
Intraoperative pathological examination$.tw,ti.
Exp frozen section/
Frozen section$.tw,ti.
Debulking surgery.tw,ti.
Exp surgical procedures, operative/
Cytoreduction$.tw,ti.
Fertility conservation,tw,ti.
Fertility sparing.tw,ti.
Surgical management.tw,ti.
Secondary cytoreduction.tw,ti.
Interval cytoreduction.tw,ti.
Surgical staging.tw,ti.
BSO.tw,ti.
Bilateral salpingo-oophorectomy.tw,ti.
Bilateral salpingo oophorectomy.tw,ti.

Surgical stage.tw,ti.
Total hysterectomy.tw,ti.
Hysterectomy.tw,ti.
((node or nodal) and dissection).tw,ti.
Exp laparotomy/
Exp laparoscopy/
Or/10-30
31 and 9
Limit 32 to English
Limit 33 to humans
Limit 34 to yr="2004-current"
Letter.pt.
Comment.pt.
Editorial.pt.
News.pt.
Case report$.pt.
Or/36-40
35 not 41

Brief Summary/Discussion of New Evidence:
Medline, EMBASE and Cochrane were searched from 2009 to July 11, 2016. Abstracts of conferences were not searched separately, but were picked up from the EMBASE search. Four separate searches were done which yielded a total of 10412 hits. Four hundred and one were retained for full text review. After the full text review, 104 studies were retained to be abstracted and can be seen below.

Clinical Expert Interest Declaration: None

Instructions. Instructions. For each document, please respond YES or NO to all the questions below. Provide an explanation of each answer as necessary.	
1. Does any of the newly identified evidence, on initial review, contradict the current recommendations, such that the current recommendations may cause harm or lead to unnecessary or improper treatment if followed?	No
2. On initial review, a. Does the newly identified evidence support the existing recommendations? b. Do the current recommendations cover all relevant subjects addressed by the evidence, such that no new	Yes No

recommendations are necessary?	
3. Is there a good reason (e.g., new stronger evidence will be published soon, changes to current recommendations are trivial or address very limited situations) to postpone updating the guideline? Answer Yes or No, and explain if necessary:	No
4. Do the PEBC and the DSG/GDG responsible for this document have the resources available to write a full update of this document within the next year?	Yes

Review Outcome	Endorse
DSG/GDG Approval Date	September 9, 2016
DSG/GDG Commentary	Might need to incorporate the discussion of simple rule diagnostic strategy and IOTA adnex model

Data Tables 4-15

Author	Study type	Comparison	N	Results
Imaging				
Haggerty 2014 ABSTRACT	Retrospective	MRI vs pathology	N=237	Pelvic MRI had a sensitivity of 95.0% and specificity of 94.1% for diagnosis of malignancy when compared to available pathology (n =88). The predicted specific histologic subtype by MRI (n= 84) was accurate in 56/57 women (98.25%) with an anticipated benign diagnosis and in 23/27 women (85.19%) with an anticipated malignancy. The correlation between a benign diagnosis from MRI and benign final surgical pathology was 0.74 (P < 0.001).
Goodrich 2014 ABSTRACT	Prospective	To investigate the relationship between imaging and the multivariate index assay (MIA) OVA1 in predicting the likelihood of ovarian malignancy before surgery	N=1110	Of the 1110 women enrolled with an adnexal mass on imaging, 1024 were evaluable. There were 255 malignancies and 769 benign tumors. High-risk findings were present in 59% of 1232 imaging tests and 61% of 1024 MIA tests. The risk of malignancy increased with MIA score; similarly, the likelihood of malignancy was higher for high-risk compared to low-risk imaging. Sensitivity and specificity for predicting malignancy was 98%

Author	Study type	Comparison	N	Results
				(95% CI 96-99) and 27% (95% CI 24-30) for imaging "OR" MIA and 80% (95% CI 74-84) and 72% (95% CI 69-76) for imaging "AND" MIA, respectively. Only 1.6% of ovarian tumors were malignant when both tests were low risk. Logistic regression analysis revealed the following combined test performance: sensitivity 93%, specificity 54%, positive and negative predictive value 40% and 96%, respectively.
Anthoulakis 2014	Systematic review (2002 to 2012) PubMed, EMBASE, Scopus, Evidence Based Medicine Reviews Cochrane, and Google Scholar.	Unenhanced and/or contrast enhanced MRI versus Doppler ultrasound, CT, and PET/CT (18F-FDG PET/CT).	Outcomes: Sensitivity (rate) and (Specificity (rate) in the detection of ovarian cancer	37 citations of which 6 papers (5 prospective and 1 retrospective) were included in the systematic review. Conclusions: MRI with intravenous contrast administration provides the highest post-test probability of ovarian cancer detection. However, the preponderant contribution of MRI in adnexal mass evaluation is its specificity because it provides confident diagnosis of many benign adnexal lesion
Dasgupta 2010	Prospective	Assessment of benign and malignant ovarian tumours by colour doppler ultrasonography compared to hisopathological examination	N=56	Pulsatility index - In 83.78% cases of benign tumour it is equal or greater than 1, it was less than 1 in 84.21% in malignant ovarian tumours. The sensitivity, specificity, PPV and NPV were respectively 84.21%, 83.78%, 72.72% and 91.11%. Resistance index in benign tumours were equal to or more than 0.4 in 81.08% and less than 0.4 in 68.42% in case of malignant tumours. The sensitivity, specificity, PPV and NPV were respectively 68.42%, 81.08%, 65% and 83.33%. Timed average maximal velocity - the sensitivity, specificity, PPV and NPV were respectively 89.45%, 89.19%, 80.95% and 94.28%, and also considering septal/central localisation of vessels as a criterion for malignancy, it was found the sensitivity, specificity, PPV and NPV were 89.47%, 62.16%, 54.84% and 92% respectively. Considering absence of dicrotic notch for malignant tumours we found sensitivity, specificity, PPV and NPV were 89.47%, 81.08%, 70.83% and 93.75% respectively.
Hafeez 2013	Retrospective	Ultrasound	N=86	Sensitivity and specificity of ultrasound to be 90.7%, 95%CI (0.77, 0.97) and 91.4%, 95%CI (0.76, 0.98) respectively. Positive predictive value was 93%, 95%CI (0.79, 0.98) and negative predictive value was 89%, 95%CI (0.73, 0.96). A total of 78 ovarian masses were detected, out of which 42 were malignant and 36 were benign.

Author	Study type	Comparison	N	Results
Firoozabadi 2011	Prospective	CT scan, physical examination Ultrasound and pathological findings	N=139	The sensitivity and specificity of sonography-physical examination were 51.9% and 87.9% respectively and the sensitivity and specificity of CT scan images were 79.2% and 91.6% respectively.
Chaudhari 2013	Prospective	Transabdominal ultrasonography and color Doppler vs histopathological examination	N-60	Efficacy of Sassone scoring system for diagnosing malignant tumors sensitivity 75%, specificity 90.91%, positive predictive value 75%, negative predictive value 90.91% and an accuracy of 86.67%. Efficacy of De Priest scoring system sensitivity 66.67%, specificity 100%, positive predictive value 100%, negative predictive value 92.31% and an accuracy of 93.33%. Efficacy of Ferrazzi scoring system sensitivity 75%, specificity 100%, positive predictive value 100%, negative predictive value 91.67%, and an accuracy of 93.33%. Efficacy of alcazar scoring system sensitivity 100%, specificity 100%, Positive predictive value 100%, negative predictive value 100%, and an accuracy of 100%.
Takeuchi 2010	Retrospective	Diffusion weighted (DWI) MRI vs. apparent diffusion coefficient (ADC)	N=49	The solid portion of all 39 malignant tumors showed homogeneous or heterogeneous high intensity on DWI, whereas only 3 of the 10 benign tumors (3 thecomas) showed high intensity. The mean (SD) ADC value in the 39 malignant tumors (1.03 [0.19]) was significantly lower than that in 10 benign tumors (1.38 [0.30]). A relatively low ADC (1.08-1.20) in the 3 thecomas may reflect their abundant cellular nature, and the presence of low intensity on T2-weighted images was suggestive for benign fibrous tumor. Low intensity on DWI with high ADC may suggest benign lesions; however, it may be occasionally difficult to differentiate benign and malignant lesions only on the basis of DWI
Mubarak 2011	Retrospective	multidetector 64-slice computed tomography (MDCT) vs. histopathology and surgical findings	N=100	MDCT was found to have 97% sensitivity, 91% specificity, and an accuracy of 96% in the differentiation of benign and malignant ovarian masses, while PPV and NPV were 97% and 91%, respectively.
Alcazar 2009	Not stated	3-dimensional (3D) power Doppler (PD) sonography to discriminate between benign and malignant cystic-solid and solid vascularized adnexal masses and to define cutoff values for 3D	N=143	A total of 113 masses (74%) were malignant, and 39 (26%) were benign. Morphologic evaluation revealed 30 unilocular solid masses (19.7%), 43 multilocular solid masses (28.3%), and 79 mostly solid masses (52%). The mean VI (9.365% versus 3.3%; P< .001), FI (34.318 versus 28.794; P< .001), and VFI (3.233 versus 1.15; P<0.01) were significantly higher in malignant tumors. No differences were found in the resistive index, pulsatility

Author	Study type	Comparison	N	Results
		PD indices to be used in a clinical setting.		index, and peak systolic velocity. Receiver operating characteristic analysis revealed an area under the curve of 0.77 (95% confidence interval, 0.69-0.85), 0.71 (0.60-0.81), and 0.75 (0.66-0.83) for the VI, FI and VFI, respectively. For reducing the false-positive rate by almost one-third, sensitivity values for the VI (cutoff, 1.556%), FI (25.212), and VFI (0.323) were 92%, 95%, and 93%, respectively.
Thomassin-Naggara 2009	Not stated	To assess the contribution of diffusion-weighted Magnetic Resonance MR imaging (DWI) for characterizing complex adnexal masses	N=77	All masses that displayed simultaneously low signal intensity within the solid component on T2-weighted and on b(1,000) diffusion-weighted images were benign. Alternatively, the presence of a solid component with intermediate T2 signal and high b(1,000) signal intensity was associated with a PLR of 4.5 for a malignant adnexal tumour.
Tsuboyama 2014	Retrospective	to correlate fluorodeoxyglucose uptake in ovarian masses on positron emission tomography/computed tomography (PET/CT) with pathological grades of malignancy and subtypes and to determine the appropriate approach for combining PET/CT and contrast-enhanced magnetic resonance imaging (CE-MRI) to characterize ovarian masses.	N=127	The SUVmax of malignant tumors was significantly higher than that of benign and borderline lesions (mean, 7.8, 1.7, 2.4; P < 0.05). Among malignant tumors, SUVmax was significantly lower in mucinous adenocarcinomas compared with nonmucinous malignant tumors (mean, 3.3, 8.4; P < 0.05) and lower in clear cell adenocarcinomas compared with other subtypes of nonmucinous malignant tumors (mean, 6.0, 9.4; P < 0.05). The SUVmax cutoff that best differentiated malignant lesions from benign/borderline lesions was 2.4 for mucinous and 4.0 for nonmucinous tumors. These cutoffs correctly classified lesions as malignant or not in 88.2% of cases (112/127). When PET/CT was combined with CE-MRI, the readers correctly classified 85% (34/40) and 86.5% (32/37) of indeterminate lesions on CE-MRI. However, PET/CT was not useful for classifying determinate lesions on CE-MRI, particularly because PET/CT correctly classified only 70.1% (12/17) of clear cell adenocarcinomas, whereas CE-MRI alone correctly classified 94.1% (1617). Thus, compared with CE-MRI alone, the diagnostic accuracy of CE-MRI + PET/CT when PET/CT was added only for indeterminate lesions on CE-MRI was significantly higher for both readers for differentiating between benign and borderline/malignant (P < 0.05), as well as between benign/borderline and malignant (P < 0.01).
Tsuboyama 2012 ABSTRACT	Not stated	to correlate FDG uptake in such ovarian masses on PET/CT with their morphology on contrast-enhanced MR (CEMR) and pathological grade of malignancy, and to evaluate the value	N=69	Malignant tumours showed significantly higher FDG uptake than benign and borderline lesions (mean, 8.1, 1.9, 2.2, P < 0.01), and of malignant tumours, significant high FDG uptake was detected on solid group (mean 9.3), but not on septate group (mean 2.6). Sensitivity, specificity, and accuracy for benign/borderline lesions by CEMR were 42.1, 100, 67.6%, respectively, those by PET/CT using cut-off SUVmax of 2.7 were

Author	Study type	Comparison	N	Results
		of adding PET/CT to CEMR in differentiating benign/borderline lesions from malignancies		81.6, 93.3, 86.8%, respectively, and those by CEMR + PET/CT using cut off SUVmax of 2.5 for septate and 4.0 for solid group were 94.7, 93.3, 94.1%, respectively. PET/CT and PET/ CT + CEMR showed significantly higher sensitivity and accuracy than CEMR (P < 0.05, <0.01, respectively).
Medeiros 2009	Systematic review	MEDLINE, CANCERLIT, LILACS, COCHRANE and EMBASE databases from January 1990 to December 2007 were seared using the terms "ovarian neoplasm" and "transvaginal ultrasound with color Doppler" were combined with "sensitivity and specificity". Reference lists of all available primary studies were reviewed to identify additional relevant citations.	To estimate the accuracy ultrasonogra phy with color Doppler in the diagnosis of ovarian tumors.	Twelve studies were analyzed, which included 2398 women. The pooled sensitivity was 0.87 (95% confidence interval [CI], 0.84-0.90); and the specificity was 0.92 (95% CI 0.87-0.90). The diagnostic odds ratio for ovarian cancer and borderline lesions vs benign lesions was 125 (95% CI, 55-283). Summary receiver operating characteristic curves were constructed because of heterogeneity in the diagnostic odds ratio. For malignant ovarian cancer and borderline versus benign lesions, the area under the curve was 0.9577. In conclusion, ultrasonography with color Doppler is a useful preoperative test for predicting the diagnosis of pelvic masses
Thomassin-Naggara 2012	Prospective	Dynamic contrast-enhanced magnetic resonance imaging (DCE-MRI) to differentiate malignant from benign adnexal tumours	N=56	Malignant tumours displayed higher F(T), Vb, rAUC and lower Ve than benign tumours (P < 0.0001, P = 0.0006, P = 0.04 and P = 0.0002, respectively). F(T) was the most relevant factor for discriminating malignant from benign tumours (AUROC = 0.86). Primary ovarian invasive tumours displayed higher F(T) and shorter Dt than borderline tumours. Malignant adnexal tumours with associated peritoneal carcinomatosis at surgery displayed a shorter Dt than those without peritoneal carcinomatosis at surgery (P = 0.01).
Kitajima 2011	Prospective	To evaluate the diagnostic value of integrated 18F-fluorodeoxyglucose (FDG) positron emission tomography and computed tomography (PET/CT) to discriminate malignant from benign ovarian tumors.	N=108	The SUV max of benign (n=26), borderline (n=12) and malignant (n=73) lesions was 2.00 +/- 1.02, 2.72 +/- 1.04, and 7.55 +/- 4.29, respectively. Although there were significant differences between benign and malignant, and borderline and malignant lesions (P<0.0001), there was no significant difference between benign and borderline lesions. Using an SUV max cutoff of 2.55, the sensitivity, specificity and accuracy of FDG-PET/CT scanning to detect malignant or borderline tumors were 82.4, 76.9, and 81.1%, respectively. The SUV max of stage I (n=35), stage II (n=8), stage III (n=34) and stage IV (n=8) was 3.59 +/- 2.32, 5.18 +/- 1.34, 8.72 +/- 2.69, and 15.05 +/- 3.77, respectively, and significant differences were observed between SUV max values and the

Author	Study type	Comparison	N	Results
				various International Federation of Gynecology and Obstetrics stage (P<0.0001).
Haggerty 2014	Retrospective	MRI vs final pathology	N=237	Data from 237 female patients who underwent pelvic MRI were included, and 41.35% underwent surgical intervention for the adnexal mass. Pelvic MRI (n = 88) was determined to have a sensitivity of 95.0% and specificity of 94.1%. The predicted specific histologic subtype by MRI (n = 84) was accurate in 56 (98.25%) of 57 women with an anticipated benign diagnosis and in 23 (85.19%) of 27 women with an anticipated malignancy. The agreement between a benign diagnosis from MRI and benign final surgical pathology was 0.85 (95% confidence interval, 0.716-0.976).
Mansour 2009	Prospective	RMI + Three-dimensional power Doppler (3-DPD)	N=400	3-DPD was added to RMI for prediction of malignancy in 400 cases of ovarian masses. Sensitivity of RMI for prediction of malignancy was 88%, with a cutoff value of 202.5 at 95% confidence interval. Sensitivity of 3-DPD for prediction of malignancy was 75%, adding 3-DPD to RMI increased its sensitivity to 99%. Considering the pilot nature of the study, further studies are needed to corroborate such findings.
Lucidarme 2010	Prospective	Preoperative three-dimensional (3D)-TVS was performed and the voxel data were analysed by the new technology. The findings at 3D-TVS, serum CA125 levels and the TVS-based diagnosis were compared with histology.	N=264	Among 375 removed ovaries, 141 cancers (83 adenocarcinomas, 24 borderline, 16 cases of carcinomatosis, nine of metastases and nine others) and 234 non-cancerous ovaries (107 normal, 127 benign tumours) were histologically diagnosed. The new computer-aided technology correctly identified 138/141 malignant lesions and 206/234 non-malignant tissues (98% sensitivity, 88% specificity). There were no false-negative results among the 47 FIGO stage I/II ovarian lesions. Standard TVS and CA125 had sensitivities/specificities of 94%/66% and 89%/75%, respectively. Combining standard TVS and the new technology in parallel significantly improved TVS specificity from 66% to 92% (p < 0.0001).
Kuyumcuoglu 2011	Retrospective	to evaluate the predictive value of PET/CT in benign and malignant ovarian tumors and compare with computerized tomography and post-operative pathology	N=69	The ROCs and AUCs values four predictors were shown in Figure 1. The AUCs (95 % CI) values calculated for CA 125, ultrasonography (USG), PET/CT and CT were as follows: 0.855 (0.752-0.958), 0.703 (0.540-0.866), 0.681 (0.514-0.848) and 0.631 (0.463-0.799) respectively CA 125 has the highest AUC value in order to predict the malignant potential of the patient. USG has the highest AUC value between the imaging techniques, following PET/CT and CT CONCLUSION:

Author	Study type	Comparison	N	Results
				According to this study among four modalities that distinguish malignant potential preoperatively; CA 125 is the best parameter USG and PET provide similar benefits in detecting malignant ovarian masses preoperatively. Both of these parameters are superior to CT Combination of CA 125, USG and PET/CT may be useful in detecting malignant ovarian masses preoperatively resulting in less invasive surgeries.
Thomassin-Naggara 2013	Retrospective	preliminary validation of an MR imaging scoring system vs. surgical pathological findings	N=394	There was almost perfect agreement (kappa > 0.80) for each MR imaging feature except for grouped septa (kappa = 0.558) and thickened regular septa (kappa = 0.555). The classification was accurate in both the training set (area under the receiver operating characteristic [ROC] curve [AUC] = 0.981 for reader 1 and 0.961 for reader 2) and the validating set (AUC = 0.964 for reader 1 and 0.943 for reader 2). ROC curve analysis demonstrated that the optimal cutoff point was an ADNEX MR score of 3; an ADNEX MR score of 4 or higher was associated with malignancy with a sensitivity of 93.5% (58 of 62) and a specificity of 96.6% (258 of 267).
Mari-Hualde 2015 ABSTRACT	Not stated	To evaluate the accuracy of 18F-FDGPET / CT (PET/CT) in the preoperative characterization of undetermined ovarian masses	N=29	22/29 cases PET / CT were classified as malignant (75.86%), presenting 2 False Positive (mature teratoma and endometriosic cyst). 7/29p were classified as benign (29.14%), with 1 false negative lesion (a borderline clear cell carcinoma). The accuracy of PET/CT was S: 95.2% (IC 95%77.3-99.2%), SP: 75% (IC 95% 40.9-92.9%), PPV: 89.4% (IC 95% 71.8-96.6%) and NPV: 87.6% (IC 95% 50.1-98%). In 21/29 cases (72.41%) (mean age: 51.9 years) tumour pathology was confirmed after surgery: 1 intestinal GIST, 2 fallopian tube carcinomas and 18 ovarian carcinomas. PET/CT classified gynaecologic tumours according to FIGO as: 8p(40%) stage I (SUVmax:6.4), 2p(10%) stage II (SUVmax:5.9), 3p(15%) stage III (SUVmax: 15.9) and 7p(35%) stage IV(SUVmax:12.2) . These findings were histologically confirmed: 16 serous adenocarcinomas (80%), 2 clear cell carcinomas (10%), 1 mucinous carcinoma and 1endometrioid carcinoma. In 8/29p (mean age: 40.9 years and SUVmax:3.2) malignancy was ruled out by surgery or follow up: 4 cystic endometriosis, 2 mature teratomas and 2 simple cysts.
Aramendia-Vidaurreta	Not stated	Calculates seven different types of	N=145	On evaluation of the classifier, an accuracy of 98.78%, sensitivity of 98.50%, specificity of

Author	Study type	Comparison	N	Results
		characteristics (local binary pattern, fractal dimension, entropy, invariant moments, gray level co-occurrence matrix, law texture energy and Gabor wavelet) from ultrasound images of the ovary, from which several features are extracted and collected together with patient age.		98.90% and area under the curve of 0.997 were calculated.
FNAC				
Pal 2015	Prospective	Ultrasound guided fine needle aspiration vs. histology	N-70	On histopathology, 62 cases were concordant with cytology. Sensitivity and specificity were 95.23 and 95.83% respectively. Diagnostic accuracy was 93.94% in respect to the correct diagnosis
Serum tests				
Macedo 2014	Systematic review / Meta Analysis	A systematic review was performed to estimate the accuracy of human epididymis protein 4 (HE4) assay in the diagnosis of ovarian tumors. Studies that evaluated HE4 levels for the diagnosis of ovarian tumors and compared them with paraffin-embedded sections as the diagnostic standard were included.	MEDLINE (PubMed), EMBASE, Cochrane, IBECS, BIOSIS, Web of Science, SCOPUS, congress abstracts, and Grey literature from January 1990 to April 2013	Forty-five studies were analyzed, which included 10,671 women and 3946 ovarian cancer cases. The pooled sensitivity for the diagnosis of borderline tumors or ovarian cancer was 78% (95% confidence interval, 77%-79%), and the specificity was 86% (95% confidence interval, 85%-87%). Summary receiver operating characteristic curves were constructed. For malignant and borderline ovarian tumors versus benign lesions, the area under the curve was 0.916. Besides the overall analysis, stratification was performed in premenopause and postmenopause, early and late stages, and for accuracy by enzyme-linked immunosorbent assay and chemiluminescence microparticle immuno assay. A HE4 level is a useful preoperative test for predicting the benign or malignant nature of pelvic masses.
Moszynsk 2013	Retrospective	Evaluation of HE4 usefulness as a test in assessment of ovarian tumors which are suspicious and difficult to classify correctly via subjective ultrasound examination.	N=245	Within the analyzed group 85 (58.6%) benign and 60 (41.4%) malignant tumors were confirmed histopathologically. The comparison of HE4 with subjective ultrasound assessment showed lowered NRI in the entire analyzed group as well as in the groups of tumors classified as "probably benign" or "probably malignant" (NRI = -0.16; P = 0.0139 and NRI = -0.133; P = 0.0489, respectively). The analysis of logistic regression model confirmed that biomarkers do not improve diagnostic accuracy. The difference between

<table>
<tr><th>Author</th><th>Study type</th><th>Comparison</th><th>N</th><th>Results</th></tr>
<tr><td></td><td></td><td></td><td></td><td>areas under ROC for HE4 (0.891) and CA125 (0.902) was not statistically significant (P = 0.760).</td></tr>
<tr><td>Terzic 2014</td><td>Prospective</td><td>Tumor marker levels (Ca 125, CEA, HE 4, Ca 19.9, and Ca 15.3), taken from all women on admission, were compared with postoperative histopathological findings of extracted tumors</td><td>N=358</td><td>Women with malignant tumors had the highest levels of Ca 125, CEA, and HE 4 (p<0.01). Mucinous adenocarcinoma produced the highest amounts of Ca 19.9 and CEA. Ca 15.3 was the highest in women with endometrioid carcinoma. There were no significant differences in the levels of all examined tumor markers (p>0.05) between women with benign and borderline tumors. Ca 125, HE 4, and Ca 15.3 can discriminate the malignant from other tumor types well. The highest sensitivity, specificity, positive and negative predictive values (91.04%, 87.6%, 67.9%, 77.2%, respectively) were achieved for the combination of Ca 125 and HE 4.</td></tr>
<tr><td>Worasethsin 2013</td><td>Prospective</td><td>D-dimer vs. CA-125</td><td>N=200</td><td><table><tr><td></td><td>D-dimer</td><td>CA-125</td></tr><tr><td>Sensitivity</td><td>91.8%</td><td>75.4%</td></tr><tr><td>Specificity</td><td>71.9%</td><td>73.0%</td></tr><tr><td>Positive predictive value</td><td>58.9%</td><td>59.7%</td></tr><tr><td>Negative predictive value</td><td>95.2%</td><td>84.8%</td></tr></table>The likelihood ratio to be negative of the D-dimer test is high at 0.11. In patients with epithelial ovarian cancer, D-dimer is increased in 83% of early stage (stage1) ovarian cancer while only 39% of early stage patients have a CA-125 level above the cut off.</td></tr>
<tr><td>Sagi-Dain 2015</td><td>Retrospective</td><td>CA19-9 and CA 125 compared to CA 125 alone</td><td>N=503</td><td>A non significant effect on sensitivity 86.9% vs. 88.9% p=0.54 and specificity 79.9% vs. 73.5% p=0.1 in differentiating malignant from benign masses.</td></tr>
<tr><td>Bozkurt 2013</td><td>Prospective</td><td>Serum levels of CA-125, CA15-3, CA19-9, carcinoembryonic antigen (CEA), and alpha-fetoprotein (AFP) compared with histopathologically confirmed diagnosis</td><td>N=168</td><td><table><tr><td></td><td>CA125 35U/ml</td><td>CA125 65U/ml</td><td>CA19-9</td></tr><tr><td>Sensitivity</td><td>78.9%</td><td>65.7%</td><td>18.4</td></tr><tr><td>Specificity</td><td>86.9%</td><td>95.3%</td><td>93%</td></tr><tr><td>PPV</td><td>63.8%</td><td>80.6%</td><td>43.7%</td></tr><tr><td>NPC</td><td>93.3%</td><td>90.5%</td><td>79.6%</td></tr></table><table><tr><td></td><td>CEA</td><td>AFP</td><td>CA15-3</td></tr><tr><td>Sensitivity</td><td>16%</td><td>2.6%</td><td>26.3%</td></tr><tr><td>Specificity</td><td>93%</td><td>98%</td><td>96.1%</td></tr><tr><td>PPV</td><td>37%</td><td>33.3%</td><td>66.6%</td></tr><tr><td>NPC</td><td>83%</td><td>77.5%</td><td>81.6%</td></tr></table></td></tr>
<tr><td>Terzic 2014</td><td>Prospective</td><td>Ca 125, CEA, HE 4,</td><td>N= 358</td><td>Women with malignant tumors had the</td></tr>
</table>

Author	Study type	Comparison	N	Results
		Ca 19.9 and Ca 15.3 vs histopathological results		highest levels of Ca 125, CEA and HE 4 ($p<0.01$). Mucinous adenocarcinoma produced the highest amounts of Ca 19.9 and CEA. Ca 15.3 was the highest in women with endometrioid carcinoma. There were no significant differences in the levels of all serum tumor markers between women with benign and borderline tumors ($p>0.05$). Malignant forms of tumors were well indicated by Ca 125, HE 4 and Ca 15.3 levels. The combination of Ca 125 and HE 4 resulted in the highest sensitivity, specificity, and positive or negative predictive value (91.04%, 87.6%, 67.9%, 77.2%, respectively).
Partheen 2011	Prospective	We hypothesized that measurement of the biomarkers HE4 and CA-125 preoperatively would improve the assignment of these patients to the correct level of care	N=394	Receiver operator characteristic (ROC) area under the curve (AUC) in the benign versus malignant cohorts was 86.8% for CA-125 and 84.4% for HE4. Negative predictive value was 91.7% when at least one of the biomarkers was positive, with only early stage epithelial ovarian cancer showing false negative results. Sensitivity at set specificity (75%) was 87% for risk of ovarian malignancy algorithm (ROMA) in the postmenopausal cohort (cut-off point, 26.0%) and 81% in the premenopausal cohort (cut-off point, 17.3%). ROC AUC in the benign versus stage I epithelial ovarian cancer was only 72% for HE4 and 76% for CA-125
Van Gorp 2011	prospective	HE4 and CA125 as a diagnostic test in ovarian cancer: prospective validation of the Risk of Ovarian Malignancy	N= 389	When all malignant tumours were included, ROMA (receiver operator characteristic (ROC)-area under curve (AUC)=0.898) and HE4 (ROC-AUC)=0.857) did not perform significantly better than CA125 alone (ROC-AUC=0.877). Using a cutoff for ROMA of 12.5% for pre-menopausal patients, the test had a sensitivity of 67.5% and a specificity of 87.9%. With a cutoff of 14.4% for post-menopausal patients, the test had a sensitivity of 90.8% and a specificity of 66.3%. For EOC vs benign disease, the ROC-AUC of ROMA increased to 0.913 and for invasive EOC vs benign disease to 0.957.
Wu 2012	Systematic review and meta analysis	The use of HE4 in the diagnosis of ovarian cancer in patients with pelvic or gynecological masses. We also evaluated the diagnostic performance of HE4 for differentiating between patients	Pubmed – Jan 1990 – Aug 2011	A total of 9 studies involving 1807 women were included. When the control group was composed of healthy women, the pooled sensitivity and specificity for HE4 in diagnosing ovarian cancer were 83% (95% confidence interval [CI], 77%-88%) and 90% (95% CI, 87%-92%), respectively. The area under the SROC curve was 0.9271. When the control group was composed of women with benign disease, the pooled sensitivity and specificity for HE4 were 74% (95% CI, 69%-

Author	Study type	Comparison	N	Results
		with benign gynecological disease and those with ovarian cancer.		78%) and 90% (95% CI, 87%-92%). The area under the SROC curve was 0.8853
Macuks 2012	Case control??	To analyze biomarker concentrations included in the two novel ovarian tumor differential diagnostic tests (risk of ovarian malignancy algorithm and OVA1) approved by food and drug administration in patients with ovarian tumors and to establish a new ovarian cancer risk assessment algorithm in conjunction with ultrasound score and menopausal status.	83 ovarian cancer patients, 76 patients with benign ovarian tumors, and 79 healthy control subjects in the control group	Mean serum concentrations of cancer antigen 125 (CA125), human epididymis secretory protein 4 (HE4), and beta-2-microglobulin were upregulated, but apolipoprotein A1, transferrin, and transthyretin were downregulated among ovarian cancer patients. When only one biomarker was introduced in the logistic regression analysis, together with ultrasonographic score and menopausal status, HE4 (area under the curve (AUC) = 0.930; 95 % confidence interval (CI) 0.891-0.969) was more accurate than CA125 (AUC = 0.902; 95 % CI 0.855-0.949) in ovarian cancer diagnostic, but when both biomarkers were included in the logistic regression analyses, ovarian cancer diagnostic accuracy was increased (AUC = 0.939; 95 % CI 0.902-0.977).
Bristow 2013	Prospective	Pre-operative serum samples and physician assessment of ovarian cancer risk were correlated with final surgical pathology.	N=494	For all ovarian malignancies, the sensitivity of the multivariate index assay was 95.7% (95%CI=89.3-98.3) when combined with clinical impression. The multivariate index assay correctly predicted ovarian malignancy in 91.4% (95%CI=77.6-97.0) of cases of early-stage disease, compared to 65.7% (95%CI=49.2-79.2) for CA125-II. The multivariate index assay correctly identified 83.3% malignancies missed by clinical impression and 70.8% cases missed by CA125-II. Multivariate index assay was superior in predicting the absence of an ovarian malignancy, with a negative predictive value of 98.1% (95%CI=95.2-99.2). Both clinical impression and CA125-II were more accurate at identifying benign disease. The multivariate index assay correctly predicted benign pathology in 204 patients (50.7%, 95%CI=45.9-55.6) when combined with clinical impression.
Wang Y.Q. 2016	Retrospective	To compare a novel prognostic inflammation score (PIS) based on preoperative serum albumin and neutrophil to	N=143	Both univariate and multivariate analyses showed that NLR and albumin were independent prognostic factors for overall survival (OS) and progression-free survival (PFS). An inverse correlation was observed between the NLR and serum albumin concentration. The novel prognostic

Author	Study type	Comparison	N	Results
		lymphocyte ratio (NLR)		inflammation score (PIS) was shown to be a significant predictor for OS and PFS (both P < 0.001) according to multivariate analysis. Additionally, low PIS was associated with advanced tumor stage (P < 0.001), metastasis (P < 0.001) and preoperative high PLR (P < 0.001).
Wang C. 2016	Prospective	To evaluate HE4, CA125, progesterone (Prog), and estradiol (E2) for differentiating pelvic masses in postmenopausal women and aimed to build a multi-marker model which may improve the diagnostic value.	N=149	After comparing by Z score statistics, HE4 + CA125 + E2 model was chosen as the best multi-marker model. In the training group, the area under curve of the HE4 + CA125 + E2 model was 0.97 (0.93, 1.00), sensitivities of the model for distinguishing BPM from EOC, from early EOC, and from advanced EOC were 90.16, 86.21, and 95.65 %; specificities were 92.11, 92.11, and 92.11 %. In the validation group, sensitivities of HE4 + CA125 + E2 model for distinguishing BPM from EOC, from early EOC, and from advanced EOC were 96.77, 100.00, and 87.50 %, specificities were 84.21, 100.00, and 84.21 %. The multi-marker model showed significant improvement when compared to CA125 or HE4, and it might be an effective pelvic mass differentiation method.
Prediction Models				
Meys 2015 ABSTRACT	Meta analysis	Eligible were prospective diagnostic studies designed to predict ovarian cancer in a preoperative setting in women with an adnexal mass.	MEDLINE, EMBASE and the Cochrane were searched (January 1990-August 2014)	We analyzed 42 articles, enrolling 18,446 adnexal tumors; 13,067 (70.8%) benign and 5,379 (29.2%) malignant. Subjective assessment, simple rules, LR2 and RMI were prospectively validated in 22, 7, 3 and 18 studies, respectively. Subjective assessment by experts performed best with a pooled sensitivity of 0.93 [95%CI 0.91-0.94] and specificity of 0.89 [95%CI 0.85-0.95] (Figure 1). Simple rules (classifying inconclusives as malignant) (sensitivity 0.92 [95%CI 0.88-0.95] and specificity 0.81 [95%CI 0.77-0.85]) and LR2 (sensitivity 0.93 [95%CI 0.89-0.95] and specificity 0.84 [95%CI 0.78-0.89]) outperformed RMI (sensitivity 0.75 [95%CI 0.72-0.79], specificity 0.92 [95%CI 0.88 to-0.94]). A two-step strategy using simple rules with subjective assessment in case simple rules are inconclusive, matched test performance of subjective assessment by an expert (sensitivity 0.90 [95%CI 0.87- 0.93] and specificity 0.93 [95%CI 0.90-0.95]).
Murala 2014	Retrospective	Performance of IOTA simple rules and RMI in preoperative classification of adnexal lesions	N=51	The IOTA simple rules yielded a conclusive result in 31 cases (65%), which resulted in a sensitivity of 94% (95% confidence interval of +/- 10%), a of specificity of 80% (+/- 20%), a positive predictive value(PPV) of 85% (+/- 16%), a negative predictive value(NPV) of

Author	Study type	Comparison	N	Results
				92%(+/- 14%). The corresponding sensitivity, specificity, positive and negative predictive value of using RMI >250 as a cut off for malignancy is 72%(+/- 18%), 80%(+/- 18%) 82%(+/- 16%), 70%(+/- 19%) respectively. Combining the simple rules and RMI resulted in sensitivity, specificity, PPV and NPV of 93%, 83%. 87% and 92% respectively.
Kaijser 2013 ABSTRACT	Prospective	OTA simple descriptors (SD) or simple rules (SR) as a triage test in patients with ovarian tumours: Subsequent value of CA125, HE4 or ROMA in clinical reality	N=360	The final database comprised 360 adnexal tumours. Malignancy rate was 40%. SD were applicable in 49% of all cases with a sensitivity of 100% and specificity of 97.1%. SR were applicable in 81% and had 99.1% sensitivity and 92.9% specificity. Sensitivity and specificity of four different strategies using ROMA, HE4 and CA125 are displayed in Figures 1 and 2, respectively. In contrast a strategy using subjective assessment of expert examiners as second stage test after SR yielded the best overall test performance (sensitivity 96.5%, specificity 90.3%).
Imperial 2015 ABSTRACT	Retrospective	To determine the accuracy of Risk of Malignancy Index in distinguishing benign adnexal masses and Epithelial Ovarian Cancer	N=121	A total of 121 patients with adnexal masses underwent surgery from October 2009 to December 2013 showing that there is a higher incidence rate of ovarian malignancy in nulligravid women ages from 41 to 60 years old. Risk of Malignancy Index using the standard value of >200 had 48.10% sensitivity, specificity of 67.20% and accuracy of 58.70%, Sassone score had the higher accuracy rate of 75% in predicting ovarian malignancy and serum CA125 had 67.6%. However, based on area under receiver operating characteristics (ROC), a cutoff value of >273 of RMI showed significant increase in sensitivity to 70%, higher specificity of 80.60% and increased accuracy rate of 65.80%. RMI >273, Sassone score >10 and serum CA125 > 60 mIU/mL that is computed based ROC yield a more accurate result in predicting ovarian carcinoma.
Franchi 2011 ABSTRACT	Prospective	To compare the pre-surgical ability of a multivariate predictive algorithm combining CA 125, HE4 and menopausal status (ROMA) vs ultrasound (US) imaging performed by an experienced examiner, for estimation of the risk of malignancy in	N= 173	Using a cut off for ROMA of 7.4% for pre-menopausal patients, and a cut off of 25.3% form post-menopausal patients. The different test performed as follow: US (Sens 96.7%, Spec 87.5%, ROC-AUC = 0.95 95%CI : 0.91,0.98), ROMA (SENS 82.7%, Spec 83.8%, ROC-AUC = 0.89 95%CI: 0.84,0.94), HE4 (Sens 74.5%, Spec 92.0%, ROC-AUC = 0.87 95%CI: 0.82,0.93) CA 125 (Sens 90.8%, Spec 66.7%, ROC-AUC = 0.89 95% CI: 0.84,0.94). US expertise opinion remains superior in discriminating malignant masses compared to ROMA algorithm, HE4 and CA125 alone.

Author	Study type	Comparison	N	Results
		patients with adnexal masses.		However, combination of biomarkers could offer an aid to less experienced sonographers in the preoperative triage of adnexal masses.
Ashrafganjooei 2011 ABSTRACT	Not stated	The aim of this study was to evaluate the use of a Risk of Malignancy Index (RMI) based on a serum CA125 level, ultrasound findings and menopausal status vs final diagnosis	N=150	The RMI identified malignant cases more accurately than any individual criterion in diagnosing ovarian cancer. Using a cut-off level of 238 to indicate malignancy, the RMI showed a sensitivity of 89.5%, a specificity of 96.2%, a PPV of 77.3%, a NPV of 98.4% and an accuracy of 95.4%.
Montagnana 2011	Prospective	ROMA	N=104	The median CA125 and HE4 serum concentrations were significantly higher among EOC patients than in healthy females (both $p<0.05$) and those with a benign mass (both $p<0.05$). The pre-menopausal group included 36 benign cases (29 of which were classified by ROMA as low-risk with a specificity of 80.6%; 95% CI: 64.0%-91.8%), and 15 EOC (eight of which were classified by ROMA as high-risk, with a sensitivity of 53.3%; 95% CI: 26.6%-78.7%). The post-menopausal group enclosed 13 benign cases (11 of which were classified by ROMA as low-risk with a specificity of 84.6%; 95% CI: 54.6%-98.0%), and 40 EOC (33 of which were classified by ROMA as high-risk with a sensitivity of 82.5%; 95% CI: 67.2%-92.7%). In the pre-menopausal group, the AUC was 0.64 ($p=0.12$, 95% CI: 0.44-0.83) for CA125, 0.77 ($p=0.003$, 95% CI: 0.62-0.92) for HE4 and 0.77 ($p=0.002$, 95% CI: 0.63-0.92) for ROMA. In the post-menopausal group, the AUC was 0.84 ($p=0.0003$, 95% CI: 0.73-0.94) for CA125, 0.94 ($p<0.0001$, 95% CI: 0.88-0.99) for HE4 and 0.92 ($p<0.0001$, 95% CI: 0.85-0.99) for ROMA.
Alcazar 2013	Prospective	International Ovarian Tumor Analysis (IOTA) 'simple' rules for discriminating between benign and malignant adnexal masses vs. final histology	N=340	Of the tumors, 55 (16.2%) were malignant and 285 (83.8%) were benign. The IOTA simple rules could be applied in 270 (79.4%) cases. In these cases, sensitivity was 87.9% (95% CI, 72.4-95.2), specificity 97.5% (95% CI, 94.6-98.8), LR+ 34.7 (95% CI, 15.6-77.3) and LR- 0.12 (95% CI, 0.05-0.31).
Fujiwara 2015	Prospective	HE4 levels and ROMA as diagnostic tools of type I and type II EOC in Japanese women.	N= 319 women (131 benign, 19 borderline, 75 malignant, and 94	CA125, HE4, and ROMA were evaluated for sensitivity and by receiver operating characteristics (ROC) in type I and type II EOC. The results showed that, at 75% specificity, the sensitivity of CA125 and HE4 for type II was 92.1% for both markers and for type I was 51.5% and 78.8%, respectively. The sensitivities of ROMA (type I, 84.8% and type

Author	Study type	Comparison	N	Results
			healthy controls	II, 97.4%) were better than those of CA125 and HE4. CA125, HE4, and ROMA were all highly accurate markers for type II. For type I, HE4 and ROMA showed better sensitivity than CA125. ROMA displayed the best diagnostic power for type I and type II including for the early stage of type I. In conclusion, HE4, CA125, and ROMA are valuable markers for type II EOC diagnosis. HE4 and ROMA analyses may improve differentiation between type I EOC and a benign mass. Measurement of combined HE4 and CA125 levels provides a more accurate method for EOC diagnosis.
Kaijser 2013	Existing dataset from prospective study	Diagnostic accuracy between LR2 and ROMA vs final histology	N= 360	216 women had benign disease and 144 a malignancy. Overall test performance of LR2 (AUC 0.952) with 94% sensitivity and 82% specificity was significantly better than ROMA (AUC 0.893) with 84% sensitivity and 80% specificity. Difference in AUC was 0.059 (95% CI: 0.026-0.091; P-value 0.0004). Similar results were obtained when stratified for menopausal status.
Moore 2014	Prospective	ROMA + Initial Clinical Risk Assessment (ICRA)	N=461	There were 375 benign tumors, 48 EOC, 18 LMP tumors and 20 non-ovarian malignancies. For detection of ovarian cancer alone, ICRA had a sensitivity of 85.4%, a specificity of 84.3%, and a NPV of 97.8%. Adding ROMA to ICRA produced a significant improvement of 8.4% in sensitivity, achieving a sensitivity of 93.8% with a specificity of 67.2% and a NPV of 98.8%. Examination of all malignancies (ovarian & non-ovarian) provided a sensitivity of 89.7% for ROMA+ICRA in comparison to 77.9% for ICRA alone, a significant increase in sensitivity of 11.8%. The NPV also significantly increased from 95.5% to 97.3%. Overall, ROMA detected 13 additional malignancies missed by ICRA
Kadija 2012	Prospective cross- sectional	HE4), the combination of HE4+CA125, and the Risk of Ovarian Malignancy Algorithm (ROMA) for patients with pelvic mass, particularly in differentiating endometriosis from carcinoma.	N=108	The level of HE4 and CA125 was significantly higher among the patients with malignant tumors, compared with patients with nonmalignant disease. At the predefined specificity of 95%, HE4 and CA125 showed sensitivity of 65.5% and 58.6%, respectively, whereas the combination of HE4+CA125 reached 68.9% at the same specificity. Importantly, the level of HE4 did not differ significantly between the patients with endometriosis and with other nonmalignant diseases (which was not the case with CA125). Risk of Ovarian Malignancy Algorithm classified 96% of benign
Moszynski 2014	Prospective	to externally	N= 268	The subjective ultrasonographic assessment

<table>
<tr><th>Author</th><th>Study type</th><th>Comparison</th><th>N</th><th>Results</th></tr>
<tr><td></td><td></td><td>validate the diagnostic performance of the International Ovarian Tumor Analysis logistic regression models (LR1 and LR2, 2005) and the Timmerman logistic regression model (1999), the Alcazar model (2003), the risk of malignancy index (RMI, 1990), and the risk of malignancy algorithm (ROMA, 2009).</td><td></td><td>and all of the studied predictive models achieved similar diagnostic performance in the whole study population. However significant differences were observed when pre- and postmenopausal patients were analyzed separately In the subgroup of premenopausal patients, the highest area under the curve (AUC) was achieved by subjective ultrasonographic assessment (0.931), the Alcazar model (0.912), and LR1 (0.909). Alternatively in the group of postmenopausal patients, the highest AUC was noted for the Timmerman model (0.973), ROMA (0.951), and RMI (0.938).</td></tr>
<tr><td>Rossi 2014</td><td>Retrospective</td><td>Comparison of Pelvic Masses Score (PMS) and Risk of Malignancy Index (RMI)</td><td>N=55</td><td><table><tr><td></td><td>PMS</td><td>RMI</td></tr><tr><td>Sensitivity</td><td>100%</td><td>85%</td></tr><tr><td>Specificity</td><td>93.8%</td><td>91%</td></tr><tr><td>PPV</td><td>70%</td><td>60%</td></tr><tr><td>NPV</td><td>100%</td><td>97.8%</td></tr></table></td></tr>
<tr><td>Klangsin 2012</td><td>Prospective</td><td>Comparison of Sassone, DePriest, Lerner, Vera and Kawai and Valentin scoring systems for prediction of malignant ovarian tumours</td><td>N=146</td><td><table><tr><td></td><td>Sensitivity</td><td>Specificity</td></tr><tr><td>Sassone</td><td>75%</td><td>79.3%</td></tr><tr><td>DePriest</td><td>89.1%</td><td>73.2%</td></tr><tr><td>Lerner</td><td>82.8%</td><td>68.3%</td></tr><tr><td>Vera & Kawai</td><td>79.7%</td><td>82.9%</td></tr><tr><td>Valentin</td><td>82.8%</td><td>85.4%</td></tr></table></td></tr>
<tr><td>Akturk 2015</td><td>Prospective</td><td>Sassone, Lerner, Ferrazzi, DePriest, Finkler, Maggino, Granberg, Szpurek, and Uelan</td><td>N=322</td><td>The sensitivity, specificity, and positive (PPV) and negative (NPV) predictive values, and diagnostic accuracy (DA) of nine scoring systems in all cases.
Sassone [3] 84.5 63.6 33.8 94.9 67
Lerner [5] 70.7 74.6 38 92.1 74
Ferrazzi [7] 72.4 66.7 32.3 91.7 67.7
DePriest [4] 87.9 67.8 37.5 96.2 71
Finkler [1] 75.9 75.4 40.4 93.4 75
Maggino [6] 75.9 68.9 34.9 92.9 70
Granberg [2] 77.6 60.2 30 92.4 63
Szpurek [9] 86.2 80.3 49 96.4 81
Ueland [8] 84.5 72 39.8 95.5 74</td></tr>
<tr><td>Kader Ali Mohan 2010</td><td>Retrospective</td><td>RMI 1, 2 and 3 comparisons</td><td>N=260</td><td>The RMI indices for distinguishing between benign and malignant patients pre-operatively had a sensitivity of 71% for RMI 1 and 3 and 79% for RMI 2. Specificity for RMI 1 and 3 was 89% and for RMI 2 was 88%. PPV for RMI 1 and 3 was 74% and was 74% for RMI 2. NPV for RMI 1 and 3 was 87% and for RMI 2 was 91%</td></tr>
<tr><td>Harry 2009</td><td>Not stated</td><td>RMI</td><td>N= 142</td><td>The sensitivity of the RMI for diagnosing</td></tr>
</table>

Author	Study type	Comparison	N	Results
				malignant ovarian disease was 94% (32/34) while the specificity was 70% (76/108).
Li 2012	Meta Analysis	(MEDLINE/PUBMED, EMBASE, Web of Science, Google Scholar, the Cochrane Library and ClinicalTrials.gov) and full texts bibliography were searched for relevant abstracts.	All studies assessed with the QUADAS-2 (Quality Assessment of Diagnostic Accuracy Studies-2). EOC predictive value of ROMA was systematically evaluated, and comparison among the predictive performances of ROMA, HE4 and CA125 were conducted within the same population.	Data of 7792 tests were retrieved from 11 studies. The overall estimates of ROMA for EOC predicting were: sensitivity (0.89, 95% CI 0.84-0.93), specificity (0.83, 95% CI 0.77-0.88), and AUC (0.93, 95% CI 0.90-0.95). Comparison of EOC predictive value between HE4 and CA125 found, specificity: HE4 (0.93, 95% CI 0.87-0.96) > CA125 (0.84, 95% CI 0.76-0.90); AUC: CA125 (0.88, 95% CI 0.85-0.91) > HE4 (0.82, 95% CI 0.78-0.85). Comparison of OC predictive value between HE4 and CA125 found, AUC: CA125 (0.89, 95% CI 0.85-0.91) > HE4 (0.79, 95% CI 0.76-0.83). Comparison among the three tests for EOC prediction found, sensitivity: ROMA (0.86, 95%CI 0.81-0.91) > HE4 (0.80, 95% CI 0.73-0.85); specificity: HE4 (0.94, 95% CI 0.90-0.96) > ROMA (0.84, 95% CI 0.79-0.88) > CA125 (0.78, 95%CI 0.73-0.83). ROMA is helpful for distinguishing epithelial ovarian cancer from benign pelvic mass. HE4 is not better than CA125 either for EOC or OC prediction.
Van Holsbeke 2009	Prospective	Internal validation of mathematical models to predict malignancy in adnexal	N=507	All IOTA models performed very well and quite similarly, with sensitivity and specificity ranging between 92% and 96% and 74% and 84%, respectively, and AUCs between 0.945 and 0.950. A least squares support vector machine with linear kernel and a logistic regression model had the largest AUCs. For pattern recognition, the AUC was 0.963, sensitivity was 90.2%, and specificity was 92.9%. This internal validation of mathematical models to estimate the malignancy risk in adnexal tumors shows that the IOTA models had a diagnostic performance similar to that in the original data set. Pattern recognition used by an expert sonologist remains the best method, although the difference in performance between the best mathematical model is not large
Yamamoto 2009	Retrospective	RMI 1, RMI 2, RMI 3, and RMI 4), incorporating menopausal status, serum CA125 levels, and ultrasound findings, to discriminate a benign from a malignant pelvic	N=253	This study confirms that, for the diagnosis of malignancy, four malignancy risk indices were more accurate than menopausal status, serum CA125 levels, and ultrasound findings separately. The accuracy of the RMI 4 was better than RMI 1 (P=0.0013), RMI 2 (P=0.0009) and RMI 3 (P=0.0013). The RMI 4 at a cutoff level of 450 yielded a sensitivity of 86.8%, a specificity of 91.0%, a positive predictive value of 63.5%, a negative

Author	Study type	Comparison	N	Results
		mass.		predictive value of 97.5%, and an accuracy of 90.4%
Bouzari 2011	Retrospective	RMI	N= 182	The RMI with the cut-off point of 265 had a sensitivity of 91.3%, specificity of 96.2 %, PPV of 77.7% and NPV of 98.7% for diagnosis of malignant masses
Yamamoto 2014	Prospective	To validate the risk-of-malignancy index (RMI) incorporating menopausal status, CA 125 levels, and imaging findings for discriminating benign from malignant pelvic masses and to evaluate the ability of 4 different RMIs	N=296	The sensitivity of RMIs 1, 2, 3, and 4 was 73.0%, 81.1%, 73.0%, and 77.0%, respectively, and the specificity was 93.7%, 89.6%, 93.7%, and 92.3%, respectively. The RMI 2 was significantly better at predicting malignancy than RMIs 1 3; however, there was no statistically significant difference in performance of RMIs 2 4
Enakpene 2009	Retrospective	To test the accuracy of risk of malignancy index (RMI) in preoperative prediction of malignancy and treatment of adnexal masses.	N= 302	The best individual performance was found in RMI at a cut-off of 250 with a sensitivity of 88.2%, specificity of 74.3%, positive predictive value of 71.3% and negative predictive value of 90%. When RMI was used to triage patient treatment, 81.5% of patients who had laparoscopy had histological diagnosis of benign ovarian tumor and 7.5% had malignant tumor. In contrast, 74.4% of patients who had laparotomy had histological diagnosis of malignant ovarian tumor and 16% had benign tumor.
Farzaneh 2014	Prospective	ROMA in pre- and post-menopause women before surgery	N=99	The only significant difference was the older age of the malignant group vs. benign group (P = 0.001) regarding demographic findings. As concerns the clinical symptoms, presence of abdominal discomfort in pre-diagnosis period was the only significant parameter in malignant group (P = 0.001). Additionally, data analysis of patients as a total group showed that specificity (96.4%), positive predictive value (PPV) (94.1%), area under the curve (AUC) (0.907), and diagnostic accuracy (DA) (86.9%) of the ROMA were higher than HE4 (91.1%, 85.7%, 0.857 and 81.8%. respectively) and CA125 (87.9%, 67.3%, 0.828 and 75.8%, respectively) alone. Besides, negative predictive value (NPV) (86.4%) and sensitivity (86.1%) of CA125 were higher than HE4 (79.7% and 69.8%, respectively). In contrast, specificity of HE4 (91.1%) was higher than CA125 (67.9%). Data analysis of patients as two groups (pre and post menopause groups) showed the same results
Alanbay 2012	Retrospective	The aim of this study	N=50	The RMI IV was the best method for

Author	Study type	Comparison	N	Results
		was to assess the prognostic values of risk of malignancy index (RMI IV), ultrasound score, menopausal status, and serum CA125 and CA19-9 level in patients with borderline ovarian tumor (BOT).		discrimination between BOTs and benign adnexal masses and was more accurate than the other parameters. When Receiver Operator Characteristic area under the curves for menopausal status was analyzed, serum CA 125 and CA19-9 level, ultrasound score, RMI IV(CA125), and RMI IV(CA19-9) were, 0.580, 0.625, 0.548, 0.694, 0.734 and 0.711, respectively. The best RMI IV cut-off was found to be 200 for discrimination of benign and BOT lesions. In the RMI formulation, replacing CA125 with CA19-9 didn't affect RMI IV sensitivity and specificity for discrimination.
Di Legge 2012	Prospective IOTA database	(IOTA) logistic regression models (LR1 and LR2), the IOTA simple rules and the risk of malignancy index (RMI). Vs histological diagnosis	N=2445	The frequency of invasive malignancy was 10% in small tumors, 19% in medium-sized tumors and 40% in large tumors; 11% of the large tumors were borderline tumors vs 3% and 4%, respectively, of the small and medium-sized tumors. The type of benign histology also differed among the three subgroups. For all methods, sensitivity with regard to malignancy was lowest in small tumors (56-84% vs 67-93% in medium-sized tumors and 74-95% in large tumors) while specificity was lowest in large tumors (60-87%vs 83-95% in medium-sized tumors and 83-96% in small tumors). The DOR and the AUC value were highest in medium-sized tumors and the AUC was largest in tumors with a largest diameter of 7-11 cm.
Chen 2015	Prospective	Serum HE4, CA125, CA153, CA199, CA211, CA724 and ROMA	N=232	The combination of HE4 and CA125 (AUC of 0.963, sensitivity of 96.6%, specificity of 65.7%) provided the best differential power in diagnosing ovarian cancer. ROMA performed better in the diagnosis of pelvic masses (AUC of 0.917, sensitivity of 82.0%, specificity of 78.8%) and uterine cancer (AUC of 0.838, sensitivity of 82.0%, specificity of 60.0%) compared with applying HE4 and CA125 individually
Gulati 2011 ABSTRACT	Not stated	Presence of Ovarian Crescent Sign (OCS) was noted and calculation of Risk of Malignancy Index (RMI) was done	N=50	RMI had sensitivity of 55.6%,negative predictive value of 90.7%,specificity and positive predictive value of 95.1% & 71.4% respectively.OCS was absent in all malignant lesions,giving a sensitivity and negative predictive value of 100%. Specificity and positive predictive value of negative OCS was 80.4% & 52.9% respectively.All masses with presence of crescent sign were benign.Combining OCS with RMI was not found to be beneficial over OCS individual
Lokich 2015 ABSTRACT	Not stated	To evaluate the use of ROMA to assist in		Initial Tumour board recommendation for patient management had a sensitivity for

Author	Study type	Comparison	N	Results
		identifying women who can safely undergo conservative management.		detecting malignancy of 100% (95% CI: 95.7-100%), specificity of 47.8% (95% CI: 42.8-52.9%), and negative predictive value (NPV) of 100% (95% CI: 98.0-100%). Actual patient management had a sensitivity of 98.8% (95% CI: 93.5-100%), specificity of 46.0% (95% CI: 41.0-51.1%) and NPV of 99.4% (95% CI: 97.0-100%). ROMA alone for the detection of EOC had a sensitivity of 95.3% (95% CI: 86.9-99.0%), specificity of 65.6% (95% CI: 60.6-70.3%), and NPV 98.8% (95% CI: 96.7-99.8%). For Stage I-II EOC, ROMA had a sensitivity of 89.3% (95% CI: 71.8-97.7%). All 84 malignancies, including 28 early-stage EOC, were recommended for surgery. Only 1 of 22 patients with an LMP tumor was assigned to observation. Clinical assessment in conjunction with ROMA identified 187 (37.4%) women for conservative management.
Martin Rodriguez 2015 ABSTRACT	Retrospective	Diagnostic accuracy of HE4, CA125 and Roma	N=62	Postmenopausal women: - HE4 had the highest specificity (94.7%) compared with CA125 (57.9%) and ROMA (57.9%) - Sensitivity of HE4 (75.0%) was lower than CA125 (87.5%) and ROMA (87.5%). HE4 had two false negatives (one mucinous OC and one granulosa cell OC). CA125 and ROMA had one false positive (granulosa OC) - HE4 also performed the best area under the curve (0.94), followed by ROMA (0.92) and CA125 (0.90) Premenopausal women: - HE4 had a high specificity (80.8%) compared with CA125 (50.0%) and ROMA (73.1%) - Sensitivity could not be assessed because all positive cases were postmenopausal women.
Thompson 2014 ABSTRACT	Retrospective	RMI	N=67	Sixty seven women were identified as having an ovarian cyst removed. There were nine cancers, all of which had ultrasound features of malignancy so the RMI score was the same. An RMI of 200 gave a sensitivity of 67% and specificity of 97%, whereas an RMI of 250 had a sensitivity of 56% and specificity of 98%. There were 20 simple cysts, all of which were benign, but eight of these women had an RMI III score of 25-250 so were given a moderate risk of malignancy. When all ovarian cancers were examined, there were two which had an RMI I score of zero, but both of these were in postmenopausal women with an abnormal CA125 level.
Yoshida 2016	Prospective	To evaluate the prediction of malignancy in women with pelvic masses using the	N=384	Of the 384 women, 224 presented a benign ovarian tumor, 32 BOT, 87 EOC, 26 non-epithelial ovarian cancer, and 15 had ovarian metastases. The best AUCs were obtained for the discrimination of EOC from benign

Author	Study type	Comparison	N	Results
		Copenhagen Index (CPH-I) and Risk of Ovarian Malignancy Algorithm (ROMA).		tumors. CPH-I performed slightly better than ROMA, and both approached 89% sensitivity and 85% specificity. When all malignant tumors (EOC, BOT, ovarian metastases and non-epithelial ovarian cancer - entire cohort) were included, the performance of CPH-I and ROMA declined to nearly 72%, although the specificity remained close to 85%
Szubert 2016	Retrospective	The external, two-center validation of the IOTA ADNEX model for differential diagnosis of adnexal tumors	N=204	ADNEX achieved high accuracy in discriminating between malignant and benign ovarian tumors in both centers (79.9% and 81.3% in Centers I and II, respectively). Multiclass accuracy was substantially lower than in binary classification (malignant vs. benign): 64.2% and 74.0% in Centers I and II, respectively. Sensitivity and specificity for the diagnosis of specific tumor types in Center I were as follows: benign tumors – 72.4% and 94.3%; borderline tumors – 33.3% and 87.0%, stage I ovarian cancers – 00.0% and 91.8%; stage II–IV ovarian cancers – 68.2% and 83.1%; and metastatic tumors – 00.0% and 99.5%. Sensitivity and specificity in Center II were as follows: benign tumors – 75.3% and 97.1%; borderline tumors – 50.0% and 88.2%, stage I ovarian cancers – 40.0% and 97.5%; stage II–IV ovarian cancers – 95.0% and 88.3%; and metastatic tumors – 20.0% and 98.3%.
Meys 2016	systematic review and meta-analysis	Our objective was to compare the diagnostic accuracy of subjective assessment, simple rules, LR2 and RMI for differentiating benign from malignant adnexal masses prior to surgery	MEDLINE, EMBASE and CENTRAL were searched (January 1990-August 2015)	We analysed 47 articles, enrolling 19,674 adnexal tumours; 13,953 (70.9%) benign and 5721 (29.1%) malignant. Subjective assessment by experts performed best with a pooled sensitivity of 0.93 (95% confidence interval [CI] 0.92-0.95) and specificity of 0.89 (95% CI 0.86-0.92). Simple rules (classifying inconclusives as malignant) (sensitivity 0.93 [95% CI 0.91-0.95] and specificity 0.80 [95% CI 0.77-0.82]) and LR2 (sensitivity 0.93 [95% CI 0.89-0.95] and specificity 0.84 [95% CI 0.78-0.89]) outperformed RMI (sensitivity 0.75 [95% CI 0.72-0.79], specificity 0.92 [95% CI 0.88-0.94]). A two-step strategy using simple rules, when inconclusive added by subjective assessment, matched test performance of subjective assessment by expert examiners (sensitivity 0.91 [95% CI 0.89-0.93] and specificity 0.91 [95% CI 0.87-0.94]).
Frozen section				
Pavlakis 2009	Retrospective	To determine accuracy of frozen section diagnosis of ovarian tumours.	N=932	Sensitivity fro frozen section diagnosis for benign, borderline and malignant epithelial tumours was 98.82%, 98.97% and 87.66%. Specificity was 98.01%, 97.06% and 100%.

Author	Study type	Comparison	N	Results
		Frozen sections were compared with final paraffin sections		27 cases had a diagnostic discrepancy.
Rakhshan 2009	Not stated	Frozen section diagnosis of adnexal masses was compared with permanent section diagnosis as the gold standard.	N=282	The overall accuracy of frozen section diagnosis was 95.7%. The sensitivity of frozen section diagnosis for benign, borderline and malignant lesions was 99, 60, and 92%, respectively. The tumor size in discrepant cases was larger than the concordant cases, however no association between mucinous histology and inaccurate diagnosis was found. The sensitivities of gross examination and clinical data in distinguishing benign from non-benign lesions were 93 and 70%, respectively.
Ouladsahebmadarek 2015		intraoperative cytology and frozen section (FS) for diagnosis of ovarian masses	N= 131	Scrape cytology for intraoperative diagnosis of benign ovarian tumors had sensitivity of 89.06% compared to 90.62% for FS. Specificity of both scrape and FS techniques for benign tumors was 94.91%. Sensitivity and specificity of scrape for malignant ovarian tumors were 94.91% and 89.06%, respectively. The related values for FS were 94.91% and 90.62%. The overall accuracy percentage of scrape and FS for diagnosis of ovarian neoplasms was 91.86% and 92.68%, respectively
Malipatil 2013	Retrospective	Compared the frozen section diagnosis of ovarian tumors with their final diagnosis in paraffin sections	N=218	Results were analyzed on two parameters: (i) status of malignancy and (ii) histological type. The overall accuracy was 95%. The sensitivity for benign, borderline and malignant tumors was 99.3%, 86.66% and 96.3%, respectively. The corresponding specificities were 92.6%, 97% and 100%. Most of the discrepant cases were of borderline category. The overall accuracy for histological diagnosis was 80.7%. The number of sections examined at frozen and paraffin had a statistically significant association with the accuracy of frozen section.
Salman 2013 ABSTRACT	Retrospective	To evaluate the accuracy of frozen section (FS) in evaluation of adnexal mass and to define clinicopathological factors associated with misdiagnosis	N=745	745 valid reports were evaluated. Of these; 507 (68.1%) had benign, 44 (5.9%) had borderline, 194 (26.0%) had malignant histological diagnosis at permanent section. In 717 of 745 (96.2%) patients, FS analysis agreed with PP. 28 of 745 cases (3.8%) were diagnosed incorrectly by FS. Univariate analysis showed that borderline histology ($p<0.0001$) and tumor size larger than 10 cm ($p= 0.005$) were associated with misdiagnosis of ovarian tumors by FS. Based on multivariate analysis, borderline histology (OR:14.4, $p<0.0001$) and tumor size larger than 10 cm (OR:2.3, $p=0.049$) were the

Author	Study type	Comparison	N	Results
				independent predictors for misdiagnosis by FS
Jhan 2016	Retrospective	Frozen section diagnosis vs final histopathology	N=54	The frozen section was accurate in 51(92.6%) cases. It had a moderately high sensitivity of 75%, high specificity of 97.6%, high positive predictive value of 90% and high negative predictive value of 93.2%. Lack of agreement was found in cases of ovarian tumours of the mucinous and borderline variety.
Ratnavelu 2016	Systematic review	Studies that used frozen section for intraoperative diagnosis of ovarian masses suspicious of malignancy, provided there was sufficient data to construct 2 x 2 tables.	MEDLINE (January 1946 to January 2015), EMBASE (January 1980 to January 2015) and relevant Cochrane registers	All studies were retrospective, and the majority reported consecutive sampling of cases. Sensitivity and specificity results were available from 38 studies involving 11,181 participants (3200 with invasive cancer, 1055 with borderline tumours and 6926 with benign tumours, determined by paraffin section as the reference standard). The median prevalence of malignancy was 29% (interquartile range (IQR) 23% to 36%, range 11% to 63%). We assessed test performance using two thresholds for the frozen section test. Firstly, we used a test threshold for frozen sections, defining positive test results as invasive cancer and negative test results as borderline and benign tumours. The average sensitivity was 90.0% (95% confidence interval (CI) 87.6% to 92.0%; with most studies typically reporting range of 71% to 100%), and average specificity was 99.5% (95% CI 99.2% to 99.7%; range 96% to 100%). Similarly, we analysed sensitivity and specificity using a second threshold for frozen section, where both invasive cancer and borderline tumours were considered test positive and benign cases were classified as negative. Average sensitivity was 96.5% (95% CI 95.5% to 97.3%; typical range 83% to 100%), and average specificity was 89.5% (95% CI 86.6% to 91.9%; typical range 58% to 99%). Results were available from the same 38 studies, including the subset of 3953 participants
Abudukadeer 2016	Retrospective	Frozen section diagnosis vs permanent section	N=804	The overall accuracy to determine the status of malignancy was 92.6%. There were 38 (7.4%) false negative and no false positive frozen section diagnoses. The sensitivity, specificity, and positive predictive and negative predictive values for benign ovarian tumors were 100.0%, 97.0%, 91.3%, and 100.0%, respectively; for borderline tumors they were 64.3%, 97.0%, 91.5%, and 94.0%, respectively, and for malignant tumors they were 90.0%, 100.0%, 100.0%, and 85.5%, respectively
Combination				

Author	Study type	Comparison	N	Results
Weinberger 2013 (abstract)	Retrospective	Sensitivity and specificity of CA 125 and "simple rules" ultrasound	N=347 tumours	Malign ovarian tumor was detected by US in 118 (sensitivity 94 %, specificity 93 %) and by CA 125 in 93 cases (sensitivity 68 %, specificity 69 %). Invasive OC was diagnosed by US in 99 (sensitivity 98 %, specificity 93 %), Borderline tumour in 19 cases (sensitivity 79 %, specificity 93 %). Invasive OC was diagnosed by CA 125 in 80 (sensitivity 74 %, specificity 69 %), BTO in 13 cases (sensitivity 46 %, specificity 69 %).
Terzic 2013	Prospective	Anamnestical data and ultrasound scans were compared with hystopathological findings of tumors	N=520	Sensitivity of anamnestical malignancy index (AMI) was 73.33%, specificity 72.87%, positive predictive value (PPV) 39.49% and negative predictive value (NPV) 91.88%. Sensitivity of characteristic malignancy index (CMI) was 92.38%, specificity 67.36%, PPV 40.59% and NPV 97.34%. Sensitivity of laboratory malignancy index (LMI) was 56.45%, specificity 90.24%, PPV 68.63%, and NPV 84.57%.
Moolthia 2009	Retrospective	RMI based on CA124 and ultrasound	N=209	Using a cut-off level of 200 to indicate malignancy, the RMI 1 gave sensitivity of 70.6%, specificity of 83.9%, PPV of 75%, and NPV of 80.6%. The RMI 2 gave sensitivity of 80%, specificity of 78.2%, PPV of 71.6%, and NPV of 85.1%. The RMI 2 was significantly better in predicting malignancy than RMI 1.
Arun-Muthuvel 2014	Prospective	RMI, CA125 and Ultrasound	N-467	RMI with a cut-off 150 had sensitivity of 84% and specificity of 97% in detecting ovarian cancer. CA-125>30 had a sensitivity of 84% and a specificity of 83%. An ultrasound score more than 2 had a sensitivity of 96% and specificity of 81%.
Sood 2010	Prospective	Ultrasound vs FNAC	57 aspirates from 50 patients	A comparison of cytological findings with the histological diagnosis was possible in 53 aspirates; in the remaining four cases (7%) the smears were acellular. On cytology, 31 lesions were diagnosed as neoplastic and 22 as non-neoplastic. The overall sensitivity of cytology in diagnosing neoplastic and non-neoplastic ovarian lesions was 93.9% and the specificity was 100%. The positive predictive value was 100% and negative predictive value 90.9%. The overall diagnostic accuracy was 96.2 %
Anton 2012	Prospective	CA125, HE4, risk ovarian malignancy algorithm (ROMA), and risk malignancy index (RMI)	N=128	The sensitivities associated with the ability of CA125, HE4, ROMA, or RMI to distinguish between malignant versus benign ovarian masses were 70.4%, 79.6%, 74.1%, and 63%, respectively. Among carcinomas, the sensitivities of CA125, HE4, ROMA (pre-and post-menopausal), and RMI were 93.5%, 87.1%, 80%, 95.2%, and 87.1%, respectively. The most accurate numerical values were

Author	Study type	Comparison	N	Results
				obtained with RMI, although the four parameters were shown to be statistically equivalent
Van Gorp 2012	Prospective	ROMA was assessed and compared to the diagnostic accuracy of the two most widely used ultrasound methods, namely the risk of malignancy index (RMI) and subjective assessment by ultrasound.	N=432	Of the 432 eligible patients, 374 could be analysed. Subjective assessment had the highest area under the receiver operator characteristic curve (AUC) (0.968, 95%CI:0.945-0.984), followed by the RMI (0.931, 95%CI:0.901-0.955). The subjective assessment and RMI both had significantly higher AUCs than the ROMA (0.893, 95%CI:0.857-0.922; P<0.0001 and P=0.0030, respectively). The pre- and postmenopausal populations generated similar results
Nunes 2014	Meta Analysis	MEDLINE, EMBASE and Cochrane from the publication of the first study in 2008. The terms used were 'simple rules', 'simple rules ovarian', 'ovar tumor' and 'ultrasound'. Quality assessment was performed using the modified Quality Assessment of the Diagnostic Accuracy of Studies (QUADAS-2) checklist.		Three hundred and three women were included in the validation study with 168 (55.4%) benign, 19 (6.3%) borderline and 116 (38.3%) malignant tumors on histological examination. The rules were applicable in 237 (78.2%) of the tumors and for these tumors, sensitivity was 96.2% (95% CI, 90.5-99.0%) and specificity was 88.6% (95% CI, 82.0-93.5%). Six of the 88 discovered studies were included in the meta-analysis along with the current validation study, which resulted in inclusion of a total of 3568 patients. When the meta-analysis was performed the pooled sensitivity (when the rules were applicable) was 93% (95% CI, 90-96%) (I(2) =32.1%) and the pooled specificity was 95% (95% CI, 93-97%) (I(2) =78.1%). Heterogeneity was observed across the studies. Sensitivity was higher and specificity lower in the study populations in which the prevalence of malignant tumors was greatest.
Al-Musalhi 2015	Retrospective	CA 125 and RMI vs histopathological diagnosis	N=361	Using the proposed cut-off 35 U/ml for CA-125 and 200 for RMI, the CA-125 test was more sensitive for detecting the majority of malignant ovarian tumors compared to the RMI (69% vs. 57%). Both tests were more sensitive in detecting epithelial ovarian cancer compared to other ovarian cancers. However, RMI was more specific in excluding benign ovarian lesions compared to CA-125 (81% vs. 68%). Additionally, RMI had a better area under the curve compared to CA-125 (0.771 vs. 0.745; p<0.005). Lowering the RMI cut-off to 150 resulted in a better sensitivity (62% vs. 57%) and had an acceptable specificity (78% vs. 81%) compared to a cut-off of 200
Valentin 2009	prospective	Whether CA 125 measurement is	N=1066	There were 242 (30%) malignancies. For 534 tumors judged to be certainly benign or

Author	Study type	Comparison	N	Results
		superior to ultrasound imaging performed by an experienced examiner for discriminating between benign and malignant adnexal lesions, and to determine whether adding CA 125 to ultrasound examination improves diagnostic performance		certainly malignant by the ultrasound examiner the sensitivity and specificity of ultrasound examination and CA 125 (> or =35 U/mL indicating malignancy) were 97% vs. 86% (95% CI of difference, 4.7-17.2) and 99% vs. 79% (95% CI of difference, 15.7-24.2); for 209 tumors judged probably benign or probably malignant, sensitivity and specificity were 81% vs. 57% (95% CI of difference, 12.3-36.0) and 91% vs. 74% (95% CI of difference, 8.5-25.7); for 66 tumors that were difficult to classify, sensitivity and specificity were 57% vs. 39% (95% CI of difference, -9.7 to 41.1) and 74% vs. 67% (95% CI of difference, -14.6 to 27.7). Diagnostic performance deteriorated when CA 125 was used as a second-stage test after ultrasound examination
Radosa 2011	Not stated	To evaluate the discriminative power of expert sonography, serum CA-125 measurement, risk malignancy index (RMI) by Jacobs, and 2 preoperative triage strategies (combination of CA-125 measurement and RMI assessment with expert sonography).	1362 surgical explorations with indication of an adnexal mass from our department were included in this study	Discriminative power of the evaluated tests differed depending on patients' menopausal state. In the premenopause, expert sonography reached the highest discriminative power with a kappa value of 0.53, a PPV of 0.45, and an NPV of 0.99. In the postmemopause, the combinations of expert sonography with CA-125 serum measurement or RMI assessment achieved the highest discriminative power: The combination of CA-125 and expert sonography reached a PPV of 0.89 and an NPV of 0.97; kappa yielded 0.84. The RMI combined with expert sonography as a triage strategy showed comparable results with a PPV of 0.89, an NPV of 0.96, and a kappa value of 0.82.
Rossi 2011	Prospective	We developed a new scoring system, named Pelvic Masses Score (PMS), that takes into account the ultrasound morphological pattern, the Doppler flowmetry of the pelvic mass, the CA125 serum level and the menopausal status.	N=160	Statistical analysis of the data obtained from the new scoring system reveals that sensitivity, specificity, positive and negative predictive values (PPV and NPV) are higher than in the case of data separately derived from the Sassone score, OTI index or RMI index.
Goyal 2014 ABSTRACT	Retrospective	Assessment of adnexal masses by color Doppler study and serum CA125 assay	N=68	Among the 68 patients, histopathological examination showed 28 (41.17%) as malignant and 40 (58.82%) benign ovarian tumors. Sensitivity, specificity, positive and negative predictive value of the diagnosis made by color Doppler and CA125 in discrimination of

Author	Study type	Comparison	N	Results
				benign and malignant ovarian tumor was calculated. With cut-off of 0.5, RI had sensitivity of 92.85% and specificity of 90%. Predictive values for positive (86.66%) and negative (94.73%) were also quite high. PI was found to be moderately sensitive with sensitivity of 71.42% with cut-off of <1 for malignancy. CA125 with a cut-off of >35 was found to detect benign cases in 90% of patients. RI was found to be more sensitive in detection of malignant cases and CA125 was more accurate for detection of benign cases
Konopacka 2012 ABSTRACT	Retrospective	To assess the diagnostic accuracy of four different modalities in the evaluation of suspicious adnexal masses (preoperative imaging, tumor markers, video-assisted laparoscopic intraoperative assessment, and frozen section) compared to the final pathology	N=131	Of the 131 patients evaluated, 44 had tumor markers examined, 127 had imaging performed (pelvic ultrasound, CT, and MRI), and 87 had specimens evaluated with frozen section. One hundred twenty-nine patients were assessed intraoperatively by an experienced gynecologic oncologist. There were 24 (18.3%) borderline/ invasive malignancies and 107 (81.7%) benign pathologies. The single most accurate test was intraoperative impression (96.9% accurate), followed by frozen section (95.4% accurate). Combined, intraoperative assessment and frozen section yielded a 91.7% sensitivity and 99.1% specificity, with an accuracy of 97.7%.
Alcazar 2011	Hospital records	Gray-scale ultrasound versus CA-125	N=1802	In total, 543 women were postmenopausal and 1259 were premenopausal. Histologically, 444 masses were malignant and 1358 were benign. Malignancy was more frequent in postmenopausal women (n=271 [49.9%]) than in premenopausal women (n=173 [13.7%]) (P<0.001). The median CA-125 level was higher in malignant (185.1 IU/mL; interquartile range, 538.1 IU/mL) than in benign (22.0 IU/mL; interquartile range, 92.8 IU/mL) tumors (P<0.001). The ROC curves indicated that the optimum CA-125 cut-off limits for accurate diagnosis of ovarian cancer would be 42.5 IU/mL for premenopausal and 30.5 IU/mL for postmenopausal women. Overall, CA-125 levels were a more accurate indicator of ovarian cancer in postmenopausal than in premenopausal women. In postmenopausal women, gray-scale ultrasound was more sensitive than CA-125 levels in the diagnosis of ovarian cancer, but CA-125 levels were more specific. In premenopausal women, gray-scale ultrasound was more sensitive and more specific than CA-125 levels
Abdalla 2015	Prospective	To compare the	N=302	Malignant disease was diagnosed in

Author	Study type	Comparison	N	Results
ABSTRACT		diagnostic performance of Gynecologic Imaging Reporting and Data System (GI-RADS) with risk of malignancy algorithm (ROMA) in the preoperative differentiation between malignant and nonmalignant adnexal masses		46(15,23%) cases. The sensitivity of GI-RADS and CA125 was 89,13% and 78,26% respectively. The specificity was 63,67% and 80,85% respectively. The positive predictive value was 44,08% and 42,35% respectively. The negative predictive value was 97,02% and 95,39% respectively.
Piovano 2015 ABSTRACT	Prospective	A preoperative ultrasound was performed and preoperative CA125 and HE4serum levels were measured. The diagnostic accuracy and the performance indices of CA125, HE4, ROMA and ultrasound (SRs+SA) and their combinations were assessed.	N=391	In the premenopausal group HE4 had the highest area under the curve (AUC) (0.698; 95% IC0.557-0.840; p<0.003), followed by ROMA (0.696; 95% IC 0.553-0.839; p<0.003) and CA125 (0.695; 95% IC0,566-0,824; p<0.003). The sensitivity(86%) and specificity (97%) of the SRs in this group was higher than the biomarkers and the ROMA algorithm. In the menopausal group ROMA had the highest AUC (0.898; 95% IC 0.846-0.950; p<0.0001), followed by CA125 (0.889; 95% IC 0.833-0.946; p<0.0001) and HE4 (0,817; 95% IC 0,740-0,880; p<0.0001). In this group, the combination of SRs and CA125 had the highest sensitivity (92%) while HE4 had the highest specificity (97%).
Myriokefalitaki 2011 ABSTRACT	Retrospective	Diagnostic performance of preoperative level of platelet count in women with ovarian masses.	N=501	362 (72.3%) women were found to have benign masses, 103 (20.6%) were diagnosed with ovarian malignancy and 36 (7.2%) borderline tumours. The mean diameter of the ovarian masses on ultrasonography was 9.77 cm (SD: 5.48 cm). The preoperative platelet count in women with ovarian cancer was higher than in women with benign ovarian mass (388 x106 vs 299 x106; p value < 0.001). We found a positive correlation (rho: 0.374) of thrombocytosis with ovarian cancer (p value < 0.001). The relative risk for ovarian cancer in the presence of preoperative thrombocytosis was 10.04. The sensitivity, specificity, positive predictive value (PPV) and negative predictive value (NPV) of thrombocytosis in differentiating between benign and malignant ovarian masses were 39 %, 96.1%, 74.1% and 84.7% respectively.
Romagnolo 2016	Not stated	HE4, CA125 and risk of ovarian malignancy algorithm (ROMA)	N=405	Good diagnostic performance in discriminating benign from EOC patients was obtained for CA125, HE4 and ROMA when calculating optimal cut-off values: premenopause, specificity (SP) > 86.6,

Author	Study type	Comparison	N	Results
				sensitivity (SN) > 82.6, area under the curves (AUC) > 0.894; postmenopause, SP > 93.2, SN > 82, AUC > 0.928. Fixing SP at 98%, performance indicators obtained for benign vs EOC patients were: premenopause, SN:65.2%, positive predictive value (+ PV): 75%, positive likelihood ratio (+ LR): 26.4 for CA125; SN:69.6%, + PV:76.2%, + LR:28.1 for HE4; SN:69.6%, + PV: 80%; + LR:35.1 for ROMA; postmenopause, SN:88%, + PV: 95.7%, + LR:38.7 for CA125; SN:78%, + PV:95.1%, + LR:34.3 for HE4; SN:88%, + PV:97.8%, + LR:77.4 for ROMA. When using routine cut-off thresholds, ROMA showed better well-balanced values of both SP and SN (premenopause, SN:87%, SP:86.1%; postmenopause, SN:90%; SP:94.3%).
Timmerman 2016	Prospective cross sectional cohort study involving 22 centres	To develop and validate a model to predict the risk of malignancy in adnexal masses using the ultrasound features in the Simple Rules	N=4848	The malignancy rate was 43% (1402/3263) in oncology centers and 17% (263/1585) in other centers. The area under the receiver operating characteristic curve on validation data was very similar in oncology centers (0.917; 95% confidence interval, 0.901-0.931) and other centers (0.916; 95% confidence interval, 0.873-0.945). Risk estimates showed good calibration. In all, 23% of patients in the validation data set had a very low estimated risk (<1%) and 48% had a high estimated risk (>30%). For the 1% risk cutoff, sensitivity was 99.7%, specificity 33.7%, LR+ 1.5, LR- 0.010, PPV 44.8%, and NPV 98.9%. For the 30% risk cutoff, sensitivity was 89.0%, specificity 84.7%, LR+ 5.8, LR- 0.13, PPV 75.4%, and NPV 93.9%.
Van Den Akker 2016	Retrospective	RMI and frozen section	N=670	Frozen sections were performed in 323 (48.2%) patients, of whom 206 (63.8%) were diagnosed with benign ovarian tumors, 55 (17.0%) with borderline tumors, and 62 (19.2%) with malignant tumors. Overall, 109 (16.3%) women had an RMI below 20, 106 (97.2%) of whom had benign histology results. Among 235 patients with an RMI over 100, 3 (1.3%) postmenopausal women had malignancies that were missed because frozen sections were not performed

REFERENCES

1. Abdalla N, Bachanek M, Timorek-Lemieszczuk A, Cendrowski K, Sawicki W. Comparison of the diagnostic value of gynecologic imaging reporting and data system and Roma in the presurgical assesment of adnexal tumors. Int J Gynecol Cancer. 2015 October;1):374.
2. Abudukadeer A, Azam S, Zunong B, Zuoremu Mutailipu A, Huijun B, Qun L. Accuracy of intra-operative frozen section and its role in the diagnostic evaluation of ovarian tumors. Eur J Gynaecol Oncol. 2016;37(2):216-20.
3. Akdeniz N, Kuyumcuoglu U, Kale A, Erdemoglu M, Caca F. Risk of malignancy index for adnexal masses. Eur J Gynaecol Oncol. 2009;30(2):178-80.
4. Akturk E, Dede M, Yenen MC, Kocyigit YK, Ergun A. Comparison of nine morphological scoring systems to detect ovarian malignancy. Eur J Gynaecol Oncol. 2015;36(3):304-8.
5. Al-Musalhi K, Al-Kindi M, Ramadhan F, Al-Rawahi T, Al-Hatali K, Mula-Abed WA. Validity of cancer antigen-125 (CA-125) and risk of malignancy index (RMI) in the diagnosis of ovarian cancer. Oman Med J. 2015;30(6):428-34.
6. Alanbay I, Akturk E, Coksuer H, Ercan M, Karasahin E, Dede M, et al. Comparison of risk of malignancy index (RMI), CA125, CA 19-9, ultrasound score, and menopausal status in borderline ovarian tumor. Gynecol Endocrinol. 2012;28(6):478-82.
7. Alcazar JL, Guerriero S. Gray-scale ultrasound versus CA-125 levels for predicting malignancy in adnexal masses. Int J Gynaecol Obstet. 2011;114(3):290-1.
8. Alcazar JL, Pascual MA, Olartecoechea B, Graupera B, Auba M, Ajossa S, et al. IOTA simple rules for discriminating between benign and malignant adnexal masses: prospective external validation. Ultrasound Obstet Gynecol. 2013;42(4):467-71.
9. Alcazar JL, Rodriguez D. Three-dimensional power Doppler vascular sonographic sampling for predicting ovarian cancer in cystic-solid and solid vascularized masses. J Ultrasound Med. 2009;28(3):275-81.
10. Anthoulakis C, Nikoloudis N. Pelvic MRI as the "gold standard" in the subsequent evaluation of ultrasound-indeterminate adnexal lesions: A systematic review. Gynecol Oncol. 2014 March;132(3):661-8.
11. Anton C, Carvalho FM, Oliveira EI, Maciel GA, Baracat EC, Carvalho JP. A comparison of CA125, HE4, risk ovarian malignancy algorithm (ROMA), and risk malignancy index (RMI) for the classification of ovarian masses. Clinics (Sao Paulo, Brazil). 2012;67(5):437-41.
12. Aramendia-Vidaurreta V, Cabeza R, Villanueva A, Navallas J, Alcazar JL. Ultrasound Image Discrimination between Benign and Malignant Adnexal Masses Based on a Neural Network Approach. Ultrasound Med Biol. 2016 01 Mar;42(3):742-52.
13. Arun-Muthuvel V, Jaya V. Pre-operative evaluation of ovarian tumors by risk of malignancy index, CA125 and ultrasound. Asian Pacific Journal of Cancer Prevention: Apjcp. 2014;15(6):2929-32.
14. Ashrafganjooei T. Risk of malignancy index in evaluation of pelvic masses. Int J Gynecol Cancer. 2011 October;3):S673.
15. Bouzari Z, Yazdani S, Kelagar ZS, Abbaszadeh N. Risk of malignancy index as an evaluation of preoperative pelvic mass. Caspian Journal of Internal Medicine. 2011;2(4):331-5.
16. Bozkurt M, Yumru AE, Aral I. Evaluation of the importance of the serum levels of CA-125, CA15-3, CA-19-9, carcinoembryonic antigen and alpha fetoprotein for distinguishing benign and malignant adnexal masses and contribution of different test combinations to diagnostic accuracy. Eur J Gynaecol Oncol. 2013;34(6):540-4.
17. Bristow RE, Smith A, Zhang Z, Chan DW, Crutcher G, Fung ET, et al. Ovarian malignancy risk stratification of the adnexal mass using a multivariate index assay. Gynecol Oncol. 2013;128(2):252-9.
18. Chaudhari SN, Deshpande PN, Gupta PR, Warty TR, Bhikane DB. Evaluation of the scoring

systems to differentiate between benign and malignant adnexal masses in a tertiary care center, Pune. Journal of SAFOG. 2013;5(3):135-8.
19. Chen X, Zhou H, Chen R, He J, Wang Y, Huang L, et al. Development of a multimarker assay for differential diagnosis of benign and malignant pelvic masses. Clin Chim Acta. 2015;440:57-63.
20. Dasgupta S, Maity SP, Sharma PP, Mukhopadhyay A, Ghosh TK. Pre-operative assessment of benign and malignant ovarian tumours using colour Doppler ultrasonography. J Indian Med Assoc. 2010 August;108(8):495-7.
21. Di Legge A, Testa AC, Ameye L, Van Calster B, Lissoni AA, Leone FP, et al. Lesion size affects diagnostic performance of IOTA logistic regression models, IOTA simple rules and risk of malignancy index in discriminating between benign and malignant adnexal masses. Ultrasound Obstet Gynecol. 2012;40(3):345-54.
22. Enakpene CA, Omigbodun AO, Goecke TW, Odukogbe AT, Beckmann MW. Preoperative evaluation and triage of women with suspicious adnexal masses using risk of malignancy index. J Obstet Gynaecol Res. 2009;35(1):131-8.
23. Farzaneh F, Honarvar Z, Yaraghi M, Yaseri M, Arab M, Hosseini M, et al. Preoperative evaluation of risk of ovarian malignancy algorithm index in prediction of malignancy of adnexal masses. Iranian Red Crescent Medical Journal. 2014;16(6).
24. Firoozabadi RD, Karimi Zarchi M, Mansurian HR, Moghadam BR, Teimoori S, Naseri A. Evaluation of diagnostic value of CT scan, physical examination and ultrasound based on pathological findings in patients with pelvic masses. Asian Pacific Journal of Cancer Prevention: Apjcp. 2011;12(7):1745-7.
25. Franchi D, Sandri M, Boveri S, Peiretti M, Radice D, Preti E, et al. Ultrasound imaging compared to a multivariate predictive algorithm combining HE4 and CA 125 (roma) in patients with adnexal masses. Int J Gynecol Cancer. 2011 October;3):S364.
26. Fujiwara H, Suzuki M, Takeshima N, Takizawa K, Kimura E, Nakanishi T, et al. Evaluation of human epididymis protein 4 (HE4) and Risk of Ovarian Malignancy Algorithm (ROMA) as diagnostic tools of type I and type II epithelial ovarian cancer in Japanese women. Tumour Biol. 2015;36(2):1045-53.
27. Goodrich S, Bristow RE, Santoso JT, Desimone CP, Miller RW, Smith A, et al. Evaluating an adnexal mass using a multivariate index assay and imaging. Gynecol Oncol. 2014 June;133:73.
28. Goyal S. Assessment of adnexal masses by color Doppler study and serum CA125 assay. BJOG. 2014 April;121:49.
29. Gulati A, Sharma A, Suneja A, Vaid NB, Sharma S, Yadav P. Comparison of ovarian crescent sign & risk of malignancy index in prediction of ovarian malignancy. Int J Gynecol Cancer. 2011 May;2):117.
30. Hafeez S, Sufian S, Beg M, Hadi Q, Jamil Y, Masroor I. Role of ultrasound in characterization of ovarian masses. Asian Pacific Journal of Cancer Prevention: Apjcp. 2013;14(1):603-6.
31. Haggerty AF, Hagemann AR, Chu C, Siegelman E, Rubin SC. Pelvic magnetic resonance imaging diagnosis correlates with pathology of adnexal masses. Gynecol Oncol. 2014 June;133:90-1.
32. Haggerty AF, Hagemann AR, Chu C, Siegelman ES, Rubin SC. Correlation of pelvic magnetic resonance imaging diagnosis with pathology for indeterminate adnexal masses. Int J Gynecol Cancer. 2014;24(7):1215-21.
33. Harry VN, Narayansingh GV, Parkin DE. The risk of malignancy index for ovarian tumours in Northeast Scotland--a population based study. Scott Med J. 2009;54(2):21-3.
34. Imperial NA, Rivera W. Risk of malignancy index in the preoperative evaluation of patients with adnexal masses. International Journal of Gynecology and Obstetrics. 2015 October;131:E412.
35. Imperial NA, Rivera W. Risk of malignancy index in the preoperative evaluation of patients with adnexal masses. BJOG. 2015 April;122:137.

36. Kader Ali Mohan GR, Jaaback K, Proietto A, Robertson R, Angstetra D. Risk Malignancy Index (RMI) in patients with abnormal pelvic mass: Comparing RMI 1, 2 and 3 in an Australian population. Aust N Z J Obstet Gynaecol. 2010;50(1):77-80.
37. Kadija S, Stefanovic A, Jeremic K, Radojevic MM, Nikolic L, Markovic I, et al. The utility of human epididymal protein 4, cancer antigen 125, and risk for malignancy algorithm in ovarian cancer and endometriosis. Int J Gynecol Cancer. 2012;22(2):238-44.
38. Kaijser J, Van Gorp T, Van Holsbeke C, Sayasneh A, Vergote I, Bourne T, et al. IOTA simple descriptors (SD) or simple rules (SR) as a triage test in patients with ovarian tumours: Subsequent value of CA125, HE4 or ROMA in clinical reality? BJOG. 2013 June;120:371.
39. Kaijser J, Van Gorp T, Van Hoorde K, Van Holsbeke C, Sayasneh A, Vergote I, et al. A comparison between an ultrasound based prediction model (LR2) and the risk of ovarian malignancy algorithm (ROMA) to assess the risk of malignancy in women with an adnexal mass. Gynecol Oncol. 2013;129(2):377-83.
40. Khan AH, Mamoon N, Usman M, Laleka US, Malik B. Accuracy of intra-operative frozen section in the diagnosis of female genital tract neoplasms. J Pak Med Assoc. 2016 February;66(2):143-6.
41. Kitajima K, Suzuki K, Senda M, Kita M, Nakamoto Y, Onishi Y, et al. FDG-PET/CT for diagnosis of primary ovarian cancer. Nucl Med Commun. 2011;32(7):549-53.
42. Klangsin S, Suntharasaj T, Suwanrath C, Kor-Anantakul O, Prasartwanakit V. Comparison of the five sonographic morphology scoring systems for the diagnosis of malignant ovarian tumors. Gynecol Obstet Invest. 2013;76(4):248-53.
43. Konopacka A, Finger T, Sternchos J, Chang-Jackson SCR, Nezhat F. Assessing adnexal masses for malignancy: A comparison of four diagnostic modalities. J Minim Invasive Gynecol. 2012 November-December;1):S12.
44. Kuyumcuoglu U, Guzel AI, Celik Y, Erdemoglu M, Komek H. 18F-FDG PET-CT and USG/CT in benign and malignant ovarian tumors with postoperative histopathological correlation. Ginekol Pol. 2011;82(8):602-6.
45. Li F, Tie R, Chang K, Wang F, Deng S, Lu W, et al. Does risk for ovarian malignancy algorithm excel human epididymis protein 4 and CA125 in predicting epithelial ovarian cancer: a meta-analysis. BMC Cancer. 2012;12 (no pagination)(258).
46. Lokich E, Palisoul M, Romano N, Stuckey AR, Robison KM, DiSilvestro PA, et al. ROMA guided conservative management for women diagnosed with an ovarian cyst or pelvic mass. Gynecol Oncol. 2015 April;137:21.
47. Lucidarme O, Akakpo JP, Granberg S, Sideri M, Levavi H, Schneider A, et al. A new computer-aided diagnostic tool for non-invasive characterisation of malignant ovarian masses: results of a multicentre validation study. Eur Radiol. 2010;20(8):1822-30.
48. Macedo AC, da Rosa MI, Lumertz S, Medeiros LR. Accuracy of serum human epididymis protein 4 in ovarian cancer diagnosis: a systematic review and meta-analysis. Int J Gynecol Cancer. 2014;24(7):1222-31.
49. Macuks R, Baidekalna I, Donina S. An ovarian cancer malignancy risk index composed of HE4, CA125, ultrasonographic score, and menopausal status: Use in differentiation of ovarian cancers and benign lesions. Tumor Biology. 2012 October;33(5):1811-7.
50. Malipatil R, Crasta JA. How accurate is intraoperative frozen section in the diagnosis of ovarian tumors? J Obstet Gynaecol Res. 2013;39(3):710-3.
51. Mansour GM, El-Lamie IK, El-Sayed HM, Ibrahim AM, Laban M, Abou-Louz SK, et al. Adnexal mass vascularity assessed by 3-dimensional power Doppler: does it add to the risk of malignancy index in prediction of ovarian malignancy?: four hundred-case study. Int J Gynecol Cancer. 2009;19(5):867-72.
52. Mari-Hualde A, Zamudio D, Pilkinton P, Castellanos T, Alonso S, Chiva L, et al. Accuracy of 18F-FDG-PET/CT in the assessment of incidental suspicious ovarian masses. Advantages and

limitations of a powerful imaging technique. Eur J Nucl Med Mol Imaging. 2015 October;1):S684-S5.
53. Martin Rodriguez S, Ascorbe Salcedo P, Jareno Blanco MS. Diagnostic accuracy of HE4, CA125 and Roma for women with suspected ovarian cancer. Clin Chem Lab Med. 2015 July;53:S424.
54. Medeiros LR, Rosa DD, da Rosa MI, Bozzetti MC. Accuracy of ultrasonography with color Doppler in ovarian tumor: a systematic quantitative review. Int J Gynecol Cancer. 2009;19(2):230-6.
55. Meys EMJ, Kaijser J, Kruitwagen RFPM, Slangen BFM, Van Calster B, Aertgeerts B, et al. Subjective assessment versus ultrasound models to diagnose ovarian cancer: A review and meta-analysis. Int J Gynecol Cancer. 2015 October;1):33-4.
56. Meys EMJ, Kaijser J, Kruitwagen RFPM, Slangen BFM, Van Calster B, Aertgeerts B, et al. Subjective assessment versus ultrasound models to diagnose ovarian cancer: A systematic review and meta-analysis. Eur J Cancer. 2016 01 May;58:17-29.
57. Montagnana M, Danese E, Ruzzenente O, Bresciani V, Nuzzo T, Gelati M, et al. The ROMA (Risk of Ovarian Malignancy Algorithm) for estimating the risk of epithelial ovarian cancer in women presenting with pelvic mass: is it really useful? Clin Chem Lab Med. 2011;49(3):521-5.
58. Moolthiya W, Yuenyao P. The risk of malignancy index (RMI) in diagnosis of ovarian malignancy. Asian Pacific Journal of Cancer Prevention: Apjcp. 2009;10(5):865-8.
59. Moore RG, Hawkins DM, Miller MC, Landrum LM, Gajewski W, Ball JJ, et al. Combining clinical assessment and the Risk of Ovarian Malignancy Algorithm for the prediction of ovarian cancer. Gynecol Oncol. 2014;135(3):547-51.
60. Moszynski R, Szubert S, Szpurek D, Michalak S, Krygowska J, Sajdak S. Usefulness of the HE4 biomarker as a second-line test in the assessment of suspicious ovarian tumors. Arch Gynecol Obstet. 2013;288(6):1377-83.
61. Moszynski R, Zywica P, Wojtowicz A, Szubert S, Sajdak S, Stachowiak A, et al. Menopausal status strongly influences the utility of predictive models in differential diagnosis of ovarian tumors: an external validation of selected diagnostic tools. Ginekol Pol. 2014;85(12):892-9.
62. Mubarak F, Alam MS, Akhtar W, Hafeez S, Nizamuddin N. Role of multidetector computed tomography (MDCT) in patients with ovarian masses. International Journal of Women's Health. 2011;3(1):123-6.
63. Murala KS, Ma K, Henry RJW, Snape S, Dwivedi R. Performance of IOTA simple rules and RMI in preoperative classification of adnexal lesions in DGH setting. BJOG. 2014 April;121:49-50.
64. Myriokefalitaki E, Burbos N, Crocker SG, Morris EP, Nieto JJ, Duncan TJ. Diagnostic performance of preoperative level of platelet count in women with ovarian masses. Int J Gynecol Cancer. 2011 October;3):S118.
65. Nunes N, Ambler G, Foo X, Naftalin J, Widschwendter M, Jurkovic D. Use of IOTA simple rules for diagnosis of ovarian cancer: meta-analysis. Ultrasound Obstet Gynecol. 2014;44(5):503-14.
66. Ouladsahebmadarek E, Tabrizi AD, Sayyah-Melli M, Jafari-Shobeiri M, Mostafa-Garabaghi P, Nazari F. Comparison of intra operative cytology and frozen section with permanent pathologic results in ovarian masses. International Journal of Women's Health and Reproduction Sciences. 2015;3(2):99-102.
67. Pal S, Chakrabarti S, Deuoghuria D, Phukan JP, Sinha A, Mondal PK. Evaluation of Ultrasound-Guided Fine-Needle Aspiration Cytology of Ovarian Masses with Histopathological Correlation. Acta Cytol. 2015;59(2):149-55.
68. Partheen K, Kristjansdottir B, Sundfeldt K. Evaluation of ovarian cancer biomarkers HE4 and CA-125 in women presenting with a suspicious cystic ovarian mass. J Gynecol Oncol. 2011;22(4):244-52.
69. Pavlakis K, Messini I, Vrekoussis T, Yiannou P, Panoskaltsis T, Voulgaris Z. Intraoperative assessment of epithelial and non-epithelial ovarian tumors: a 7-year review. Eur J Gynaecol Oncol. 2009;30(6):657-60.

70. Piovano E, Macchi C, Cavallero C, Fuso L, De Ruvo D, Viora E, et al. Preoperative management of the adnexal mass: A prospective multi-center study based on 391 patients. Int J Gynecol Cancer. 2015 October;1):487.
71. Radosa MP, Camara O, Vorwergk J, Diebolder H, Winzer H, Mothes A, et al. Preoperative multimodal strategies for risk assessment of adnexal masses: analysis of 1362 cases in a gynecologic cancer center. Int J Gynecol Cancer. 2011;21(6):1056-62.
72. Rakhshan A, Zham H, Kazempour M. Accuracy of frozen section diagnosis in ovarian masses: experience at a tertiary oncology center. Arch Gynecol Obstet. 2009;280(2):223-8.
73. Ratnavelu ND, Brown AP, Mallett S, Scholten RJ, Patel A, Founta C, et al. Intraoperative frozen section analysis for the diagnosis of early stage ovarian cancer in suspicious pelvic masses. Cochrane Database of Systematic Reviews. 2016 March 01;2016 (3) (no pagination)(CD010360).
74. Romagnolo C, Leon AE, Fabricio ASC, Taborelli M, Polesel J, Del Pup L, et al. HE4, CA125 and risk of ovarian malignancy algorithm (ROMA) as diagnostic tools for ovarian cancer in patients with a pelvic mass: An Italian multicenter study. Gynecol Oncol. 2016 01 May;141(2):303-11.
75. Rossi A, Braghin C, Soldano F, Isola M, Capodicasa V, Londero AP, et al. A proposal for a new scoring system to evaluate pelvic masses: Pelvic Masses Score (PMS). Eur J Obstet Gynecol Reprod Biol. 2011;157(1):84-8.
76. Rossi A, Forzano L, Romanello I, Ambrosini G, Iuri V, Marchesoni D. Comparison of pelvic masses score (PMS) and Risk of Malignancy Index (RMI 3) in the evaluation of pelvic masses. Eur J Gynaecol Oncol. 2014;35(4):421-4.
77. Sagi-Dain L, Lavie O, Auslander R, Sagi S. CA 19-9 in evaluation of adnexal mass: Retrospective cohort analysis and review of the literature. Int J Biol Markers. 2015 01 Jul;30(3):e333-e40.
78. Salman MC, Basaran D, Selcuk I, Boyraz G, Ozgul N, Usubutun A, et al. Accuracy of frozen section diagnosis in the evaluation of adnexal mass: Retrospective evaluation of 745 cases with multivariate regression analysis. Int J Gynecol Cancer. 2013 October;1):456-7.
79. Sood T, Handa U, Mohan H, Goel P. Evaluation of aspiration cytology of ovarian masses with histopathological correlation. Cytopathology. 2010;21(3):176-85.
80. Takeuchi M, Matsuzaki K, Nishitani H. Diffusion-weighted magnetic resonance imaging of ovarian tumors: differentiation of benign and malignant solid components of ovarian masses. J Comput Assist Tomogr. 2010;34(2):173-6.
81. Terzic M, Dotlic J, Likic I, Ladjevic N, Brndusic N, Mihailovic T, et al. Predictive factors of malignancy in patients with adnexal masses. Eur J Gynaecol Oncol. 2013;34(1):65-9.
82. Terzic M, Dotlic J, Likic I, Nikolic B, Brndusic N, Pilic I, et al. Diagnostic value of serum tumor markers for adnexal masses. Central European Journal of Medicine. 2014 June;9(3):417-23.
83. Terzic M, Dotlic J, Likic I, Nikolic B, Brndusic N, Pilic I, et al. Diagnostic value of serum tumor markers evaluation for adnexal masses. Central European Journal of Medicine. 2014 April;9(2):210-6.
84. Thomassin-Naggara I, Aubert E, Rockall A, Jalaguier-Coudray A, Rouzier R, Darai E, et al. Adnexal masses: development and preliminary validation of an MR imaging scoring system. Radiology. 2013;267(2):432-43.
85. Thomassin-Naggara I, Balvay D, Aubert E, Darai E, Rouzier R, Cuenod CA, et al. Quantitative dynamic contrast-enhanced MR imaging analysis of complex adnexal masses: a preliminary study. Eur Radiol. 2012;22(4):738-45.
86. Thomassin-Naggara I, Darai E, Cuenod CA, Fournier L, Toussaint I, Marsault C, et al. Contribution of diffusion-weighted MR imaging for predicting benignity of complex adnexal masses. Eur Radiol. 2009;19(6):1544-52.
87. Thompson R, Dempsey A, Abdel-Aty M. Which risk of malignancy index (RMI) calculation is a better predictor of malignancy, and at what level should we refer to the cancer centre? A

retrospective observational study conducted at East Lancashire Hospitals NHS Trust. BJOG. 2014 November;121:9.
88. Timmerman D, Van Calster B, Testa A, Savelli L, Fischerova D, Froyman W, et al. Predicting the risk of malignancy in adnexal masses based on the Simple Rules from the International Ovarian Tumor Analysis group. Am J Obstet Gynecol. 2016 01 Apr;214(4):424-37.
89. Timmerman D, Testa A, Bourne T, Ameye L, Jurkovic D, Van Holsbeke C, et al. Simple ultrasound-based rules for the diagnosis of ovarian cancer. Ultrasound Obstet Gynecol. 2008;31:681-690.
90. Tsuboyama T, Tatsumi M, Kim T, Onishi H, Hori M, Nakamoto A, et al. The value of combining 18F-FDG PET/CT and contrastenhanced MR in the evaluation of ovarian masses with solid component or thickened septa. J Med Imaging Radiat Oncol. 2012 August;56:150.
91. Tsuboyama T, Tatsumi M, Onishi H, Nakamoto A, Kim T, Hori M, et al. Assessment of combination of contrast-enhanced magnetic resonance imaging and positron emission tomography/computed tomography for evaluation of ovarian masses. Invest Radiol. 2014;49(8):524-31.
92. Valentin L, Jurkovic D, Van Calster B, Testa A, Van Holsbeke C, Bourne T, et al. Adding a single CA 125 measurement to ultrasound imaging performed by an experienced examiner does not improve preoperative discrimination between benign and malignant adnexal masses. Ultrasound Obstet Gynecol. 2009;34(3):345-54.
93. Van Den Akker PAJ, Zusterzeel PLM, Aalders AL, Snijders MPLM, Samlal RAK, Vollebergh JHA, et al. Use of risk of malignancy index to indicate frozen section analysis in the surgical care of women with ovarian tumors. International Journal of Gynecology and Obstetrics. 2016 01 Jun;133(3):355-8.
94. Van Calster B, Van Hoorde K, Froyman W, Kauser J, Landolfo C,Anthoulakis, et al. Practical guidance for applying the ADNEX model from the IOTA group to discriminate between different subtypes of adnexal tumors. Facts Views Vis OBGYN. 2015;7(1):32-41.
95. Van Gorp T, Cadron I, Despierre E, Daemen A, Leunen K, Amant F, et al. HE4 and CA125 as a diagnostic test in ovarian cancer: prospective validation of the Risk of Ovarian Malignancy Algorithm. Br J Cancer. 2011;104(5):863-70.
96. Van Gorp T, Veldman J, Van Calster B, Cadron I, Leunen K, Amant F, et al. Subjective assessment by ultrasound is superior to the risk of malignancy index (RMI) or the risk of ovarian malignancy algorithm (ROMA) in discriminating benign from malignant adnexal masses. Eur J Cancer. 2012;48(11):1649-56.
97. Van Holsbeke C, Van Calster B, Testa AC, Domali E, Lu C, Van Huffel S, et al. Prospective internal validation of mathematical models to predict malignancy in adnexal masses: results from the international ovarian tumor analysis study. Clin Cancer Res. 2009;15(2):684-91.
98. Wang C, Cheng L, Zhang P, Guo L, Liu W, Zhang Z, et al. Development of a multi-marker model combining HE4, CA125, progesterone, and estradiol for distinguishing benign from malignant pelvic masses in postmenopausal women. Tumor Biology. 2016 01 Feb;37(2):2183-91.
99. Wang YQ, Jin C, Zheng HM, Zhou K, Shi BB, Zhang Q, et al. A novel prognostic inflammation score predicts outcomes in patients with ovarian cancer. Clin Chim Acta. 2016 May 01;456:163-9.
100. Weinberger V, Minar L. Diagnostics of malign ovarian tumors by ultrosound and CA 125-our experience. Int J Gynecol Cancer. 2013 October;1):498.
101. Worasethsin P, Narkwichean A. D-dimer as a tumor marker in pre-operative assessment of adnexal masses. J Med Assoc Thai. 2013;96(11):1395-400.
102. Wu L, Dai ZY, Qian YH, Shi Y, Liu FJ, Yang C. Diagnostic value of serum human epididymis protein 4 (HE4) in ovarian carcinoma: a systematic review and meta-analysis. Int J Gynecol Cancer. 2012;22(7):1106-12.

103. Wu Y, Peng H, Zhao X. Diagnostic performance of contrast-enhanced ultrasound for ovarian cancer: a meta-analysis. Ultrasound Med Biol. 2015;41(4):967-74.
104. Yamamoto Y, Tsuchida A, Ushiwaka T, Nagai R, Matsumoto M, Komatsu J, et al. Comparison of 4 risk-of-malignancy indexes in the preoperative evaluation of patients with pelvic masses: A prospective study. Clinical Ovarian and other Gynecologic Cancer. 2014 01 Dec;7(1-2):8-12.
105. Yamamoto Y, Yamada R, Oguri H, Maeda N, Fukaya T. Comparison of four malignancy risk indices in the preoperative evaluation of patients with pelvic masses. Eur J Obstet Gynecol Reprod Biol. 2009;144(2):163-7.
106. Yoshida A, Derchain SF, Pitta DR, Andrade LA, Sarian LO. Comparing the Copenhagen Index (CPH-I) and Risk of Ovarian Malignancy Algorithm (ROMA): Two equivalent ways to differentiate malignant from benign ovarian tumors before surgery? Gynecol Oncol. 2016;140(3):481-5.

OUTCOMES DEFINITIONS

1. ARCHIVED – An archived document is a document that will no longer be tracked or updated but may still be useful for academic or other informational purposes. The document is moved to a separate section of the Web site and each page is watermarked with the phrase "ARCHIVED".

2. ENDORSED – An endorsed document is a document that the DSG/GDG has reviewed for currency and relevance and determined to be still useful as guidance for clinical decision making. A document may be endorsed because the DSG/GDG feels the current recommendations and evidence are sufficient, or it may be endorsed after a literature search uncovers no evidence that would alter the recommendations in any important way.

3. DELAY – A Delay means that there is reason to believe new, important evidence will be released within the next year that should be considered before taking further action.

4. UPDATE – An Update means that the DSG/GDG recognizes that there is new evidence that makes changes to the existing recommendations in the guideline necessary but these changes are more involved and significant than can be accomplished through the Document Assessment and Review process. The
DSG/GDG will rewrite the guideline at the earliest opportunity to reflect this new evidence. Until that time, the document will still be available as its existing recommendations are still of some use in clinical decision making

.