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

Dirk Timmerman, François Planchamp, Tom Bourne, Chiando Landolfo, Andreas du Bois, Luis Chiva, David Cibula, Nicole Concin, Daniela Fischerova, Wouter Froyman, Guillermo Gallardo Madueño, Birthe Lernyey, Annika Lofth, Liliana Meru, Philippe Morice, Denis Querleu, Antonio Carla Testa, Ignace Vergote, Vincent Vandecaveye, Giovanni Scambia, Christina Fotopoulou

ABSTRACT
The European Society of Gynaecological Oncology (ESGO), the International Society of Ultrasound in Obstetrics and Gynecology (ISUOG), the International Ovarian Tumour Analysis (IOTA) group, and the European Society for Gynaecological Endoscopy (ESGE) jointly developed clinically relevant and evidence-based statements on the pre-operative diagnosis of ovarian tumors, including imaging techniques, biomarkers, and prediction models. ESGO/ISUOG/IOTA/ESGE nominated a multidisciplinary international group, including expert practising clinicians and researchers who have demonstrated leadership and expertise in the pre-operative diagnosis of ovarian tumors and management of patients with ovarian cancer (19 experts across Europe). A patient representative was also included in the group. To ensure that the statements were evidence-based, the current literature was reviewed and critically appraised. Preliminary statements were drafted based on the review of the relevant literature. During a conference call, the whole group discussed each preliminary statement and a first round of voting was carried out. Statements were removed when a consensus among group members was not obtained. The voters had the opportunity to provide comments/suggestions with their votes. The statements were then revised accordingly. Another round of voting was carried out according to the same rules to allow the whole group to evaluate the revised version of the statements. The group achieved consensus on 18 statements. This Consensus Statement presents these ESGO/ISUOG/IOTA/ESGE statements on the pre-operative diagnosis of ovarian tumors and the assessment of carcinomatosis, together with a summary of the evidence supporting each statement.

INTRODUCTION
The accurate characterization of newly diagnosed adnexal lesions is of paramount importance to define appropriate treatment pathways. Patients with masses that are suspicious for malignancy should be referred to a gynaecological oncology center, in order to receive specialist care, as per the definitions of the European Society of Gynaecological Oncology (ESGO) and national and international recommendations and guidelines. For a non-gynaecological primary tumor, patients need to be referred to an appropriate specialist, while patients with benign lesions may be followed up and treated conservatively or may be suitable for less radical surgical treatment, depending on the clinical context. Treatment decision-making processes should be based on a combination of the patient’s overall clinical picture, symptoms, preferences, previous medical and surgical history, tumor markers, and clinical and radiological findings. A single diagnostic modality alone should not determine the patient’s journey.

The ESGO, the International Society of Ultrasound in Obstetrics and Gynecology (ISUOG), the International Ovarian Tumour Analysis group (IOTA), and the European Society for Gynaecological Endoscopy (ESGE) have, jointly, developed clinically relevant and evidence-based statements on the pre-operative diagnosis of ovarian tumors and assessment of disease spread, including imaging techniques, biomarkers, and predictive models. Neither screening and follow-up modalities nor economic analysis of the imaging techniques, biomarkers, and prediction models addressed herein are included within the remit of this Consensus Statement.

RESPONSIBILITIES
The present series of statements form a consensus of the authors regarding their currently accepted approaches for the pre-operative diagnosis of ovarian tumors and assessment of disease spread, based on the available literature and evidence. Any clinician applying or consulting these statements is expected to use independent medical judgment in the context of individual clinical circumstances to determine all patients’ care and treatment. These statements are presented without any warranty regarding their content, use or application, and the authors disclaim any responsibility for their application or use in any way.

METHODS
This Consensus Statement on the pre-operative diagnosis of ovarian tumors and assessment of disease spread was developed using an eight-step process, chaired by Professors Christina Fotopoulou and Dirk
Nomination of multidisciplinary international group

Identification of scientific evidence

Formulation of preliminary consensus statements

Discussion of each preliminary consensus statement

First round of voting

Revision of consensus statements (where necessary)

Second round of voting

Finalization of statements

Figure 1 Eight-step process for development of the Consensus Statement on the pre-operative diagnosis of ovarian tumors and assessment of disease spread.

Timmerman (Figure 1). Aiming to assemble a multidisciplinary international group, ESGO/ISUOG/IOTA/ESGE nominated 19 practising clinicians and researchers who have demonstrated leadership and expertise in the pre-operative diagnosis of ovarian tumors and clinical management of patients with ovarian cancer through research, administrative responsibilities, and/or committee membership (including eight members of ESGO, five members of ISUOG, four members of IOTA, and two members of ESGE). These experts included seven gynecologists with special interest in ultrasonography, two radiologists, and 10 gynecological oncologists. They did not represent the societies from which they were selected, and were asked to base their decisions on their own experience and expertise. Also included in the group was a patient representative, who is Chair of the Clinical Trial Project of the European Network of Gynaecological Cancer Advocacy Groups (ENGAGe). An initial conference call including the whole group was held to facilitate introductions, as well as to review the purpose and scope of this Consensus Statement.

To ensure that the statements were evidence-based, the current literature was reviewed and critically appraised. Thus, a systematic literature review of relevant studies published between 1 May 2015 and 1 May 2020 was carried out using the MEDLINE database (Online supplemental appendix 1). The literature search was limited to publications in the English language. Priority was given to high-quality systematic reviews, meta-analyses, and validating cohort studies, although studies with lower levels of evidence were also evaluated. The search strategy excluded editorials, letters, and case reports. The reference list of each identified article was reviewed for other potentially relevant articles. Final results of the literature search were distributed to the whole group, including electronic full-text versions of each article. One of the authors (FP) provided the methodology and medical writing support for the entire process and did not participate in voting for statements.

The chairs were responsible for drafting preliminary statements based on the review of the relevant literature. These were then sent to the multidisciplinary international group prior to a second conference call. During this conference call, the whole group discussed each preliminary statement and a first round of binary voting (agree/disagree) was carried out for each potential statement. All 20 participants took part in each vote, but they were permitted to abstain from voting if they felt they had insufficient expertise to agree/disagree with the statement or if they had a conflict of interest that could be considered to influence their vote. Statements were removed when a consensus among group members was not obtained. The voters had the opportunity to provide comments/suggestions with their votes. The chairs then discussed the results of this first round of voting and revised the statements if necessary. The voting results and the revised version of the statements were again sent to the whole group and another round of binary voting was organized, according to the same rules, to allow the whole group to evaluate the revised version of the statements. The statements were finalized based on the results of this second round of voting. The group achieved consensus on 18 statements. In this Consensus Statement, we present a summary of the supporting evidence, the finalized series of statements, and their levels of evidence and grades.

RESULTS

General remarks

Even though the test performance of any biochemical or radiological diagnostic test appears to increase after excluding borderline ovarian tumors and non-gynecological primary tumors, such as of the gastrointestinal tract or breast, we included in our literature assessment studies addressing all types of adnexal tumor, as this is a better reflection of clinical reality.

Ultrasonography

A transvaginal ultrasound examination is often regarded in clinical practice as the standard first-line imaging investigation for the assessment of adnexal pathology. The diagnostic accuracy of ultrasonography in differentiating between benign and malignant adnexal masses has been shown to relate to the expertise of the operator. The European Federation of Societies for Ultrasound in Medicine and Biology has published minimum training requirements for gynecological ultrasound practice in Europe, including standards for theoretical knowledge and practical skills. These identify three levels of training and expertise. Thus, Level III (expert) can be attributed to a practitioner who is likely to spend the majority of their time undertaking gynecological ultrasound and/or teaching, research and development in the field. A Level II practitioner should have undertaken at least 2000 gynecological ultrasound examinations. The training required to attain this level of practice would usually be gained during a period of expert ultrasound training, which may be within, or after completion of, a specialist training program. To maintain competence at Level II, practitioners should perform at least 500 examinations each year. A Level I practitioner should have performed a minimum of 300 examinations under the supervision of a Level II practitioner or an experienced Level
I practitioner with at least 2 years’ regular practical experience. To maintain Level I status, the practitioner should perform at least 300 examinations each year. A prospective randomized controlled trial to assess the effect of the quality of gynecological ultrasonography on the management of patients with suspected ovarian cancer has shown that women with a Level III (expert) ultrasound examination undergo significantly fewer unnecessary major procedures and have a shorter inpatient hospital stay compared with those having a Level II (routine) examination by a sonographer.9

Subjective assessment by expert ultrasound examiners has excellent performance to distinguish between benign and malignant ovarian tumors.5–14 In many cases, expert examiners should be able to narrow the diagnosis down further to a specific histological sub-type. The typical pathognomonic ultrasound features of some key histological types have been published in the series ‘Imaging in gynecological disease’ in Ultrasound in Obstetrics and Gynecology (https://obgyn.onlinelibrary.wiley.com/doi/toc/10.1002/(ISSN)1469-0705.IMAGINGINGYNECOLOGICALDISEASE). The most common and typical findings for each pathology are summarized in Table 1.

Risk of malignancy index (RMI) and risk of ovarian malignancy algorithm (ROMA)

Several attempts have been made to develop more objective ultrasound-based approaches for discriminating between benign and malignant adnexal tumors. These include the risk of malignancy index (RMI), a scoring system based on menopausal status, a transvaginal ultrasound score, and serum cancer antigen 125 (CA 125) level.8 Many studies have demonstrated the diagnostic performance of the RMI in classifying adnexal masses.11–13 Three variants of the RMI (RMI-II, RMI-III, RMI-IV) have been developed, but these offer no significant additional diagnostic advantage compared with the original version (RMI-I).11,12,27,28 Moore et al.30 developed an algorithm, the risk of ovarian malignancy algorithm (ROMA), based on both CA 125 and human epidymis protein 4 (HE4). Westwood et al.18 pooled data comparing the ROMA with the RMI-I to guide referral decisions for women with suspected ovarian cancer and found similar performance if women with borderline tumors and non-epithelial cancers were excluded from the analyses. More recently, another meta-analysis showed a higher specificity of the RMI-I than the ROMA in pre-menopausal women but a similar performance for detecting ovarian cancer in post-menopausal women presenting with an adnexal mass.17 Limitations of the RMI are the absence of an estimated risk of malignancy and its considerable dependency on serum CA 125, the latter resulting in a relatively low sensitivity for early-stage invasive and borderline disease, especially in pre-menopausal women31,32 (see Tumor Markers).

IOTA methods

To homogenize and standardize the quality, description, and evaluation of ultrasonography across different centers, and thereby increase diagnostic accuracy, the IOTA group first published a consensus paper on terms and definitions to describe adnexal lesions in 2000.33 Using this standardized methodology, the IOTA group has developed different prediction models based on logistic regression analysis.34–36 In a large-scale external validation study, Van Holsbeke et al.37 showed that the IOTA logistic regression models 1 (LR1, with 12 variables) and 2 (LR2, with six variables) outperformed 12 other models, including the RMI. The LR2 model was easier to use than the LR1 model. Demonstrating the standardization and reproducibility of the IOTA models, Sayasneh et al.38 showed that even less experienced sonographers are able to differentiate accurately between benign and malignant ovarian masses using the IOTA LR1 model. The IOTA group also developed ‘Simple Rules’ that may be applied to a mass based on the presence or absence of five benign and five malignant ultrasound features. These rules can be applied to about 80% of adnexal masses, with the rest being classed as inconclusive. They have now been broadly accepted and are widely used in clinical practice.39–41 More recently, a logistic regression model based on the ultrasound features of the original Simple Rules was developed—the Simple Rules risk model. This model is able to provide an individual estimated risk of malignancy for any type of lesion.35 A summary of the main models and scoring systems for the pre-operative diagnosis of ovarian tumors is shown in Table 2.

As many ovarian masses can be recognized relatively easily, the IOTA group also proposed four ‘Simple Descriptors’ of the features typical of common benign lesions and two suggestive of malignancy, which can give an ‘instant diagnosis’ and reflect the pattern recognition that is a key part of ultrasonography. These are applicable to about 43% of adnexal masses.42 A three-step strategy, consisting of the sequential use of Simple Descriptors, Simple Rules, and subjective assessment by an expert, had high accuracy for discriminating between benign and malignant adnexal lesions.43 A systematic review and meta-analysis reported better performance of the IOTA Simple Rules and the IOTA LR2 model compared with all other scoring systems, including the RMI.44 Besides confirming these findings, another meta-analysis highlighted that a two-step approach, with the IOTA Simple Rules as the first step and subjective assessment by an expert for inconclusive tumors as the second step, matched the test performance of expert ultrasound examiners.11 The IOTA Simple Rules have been integrated into several national clinical guidelines for the evaluation and management of adnexal masses,49,50 and they were considered the main diagnostic strategy51 as part of a first international consensus report for the assessment of adnexal masses.

A randomized controlled trial assessing surgical intervention rates and the oncologic safety of decision-making processes using an RMI-based protocol developed by the British Royal College of Obstetricians and Gynaecologists (RCOG) versus triage using the IOTA Simple Rules42 showed that the IOTA protocol resulted in lower surgical intervention rates compared with the RMI-based RCOG protocol. The IOTA Simple Rules did not result in more cases in which a diagnosis of cancer was delayed. It was found that the addition of biomarkers such as serum CA 125 and HE4 when using the IOTA Simple Rules, with or without subjective assessment by an expert sonographer, offered no additional diagnostic advantage for the characterization of ovarian masses, but was more costly than a three-step strategy based on the sequential use of the IOTA Simple Descriptors, Simple Rules, and expert evaluation.53,54 The IOTA group have also developed the Assessment of Different Neoplasias in the adnexa (ADNEX) model. This multiclass prediction model is the first risk model to differentiate between benign and malignant tumors, while also offering sub-classification of any malignancy into borderline tumors, Stage I, and Stage II–IV primary
### Table 1: Clinical and ultrasound features typical of different histological sub-types of adnexal tumor

<table>
<thead>
<tr>
<th>Category/type</th>
<th>Age (years)</th>
<th>Laterality</th>
<th>Appearance</th>
<th>Typical features</th>
<th>Color score</th>
<th>Picture</th>
<th>Ref</th>
</tr>
</thead>
<tbody>
<tr>
<td>Endometriosis-related tumors</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Endometrioma</td>
<td>Median, 34</td>
<td>Uni/bi</td>
<td>Uni- or multilocular (1–4 locules)</td>
<td>Groundglass content; papillations in 10%, but most often without internal blood flow; premenopausal patient; raised CA 125 (median 44 U/mL)</td>
<td>1/2/3</td>
<td><img src="image" alt="" /></td>
<td>171</td>
</tr>
<tr>
<td>Benign tumors</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sex cord-stromal tumor</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fibroma/ fibrothecoma (65%)</td>
<td>Median, 50; 65% postmenopausal</td>
<td>Uni</td>
<td>Regular round, oval or slightly lobulated solid tumors; sometimes multilocular-solid (15–20%)</td>
<td>Fan-shaped shadowing; often raised CA 125 (34%) and/or ascites</td>
<td>(1)/2/3</td>
<td><img src="image" alt="" /></td>
<td>172</td>
</tr>
<tr>
<td>Sertoli cell tumor (most benign)</td>
<td>≤30 (75%)</td>
<td>Uni</td>
<td>Solid; median diameter 90 mm</td>
<td>Hormonally inactive or estrogen-producing (abnormal bleeding)</td>
<td>3/4</td>
<td><img src="image" alt="" /></td>
<td>173</td>
</tr>
<tr>
<td>Leydig cell tumor (almost all benign)</td>
<td>Median, 58</td>
<td>Uni</td>
<td>Solid; median diameter 24 mm</td>
<td>Endocrine symptoms (75% virilization); testosterone/androstenedione</td>
<td>3/4</td>
<td><img src="image" alt="" /></td>
<td>173</td>
</tr>
<tr>
<td>Germ cell tumor</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Continued
### Table 1 Continued

<table>
<thead>
<tr>
<th>Category/type</th>
<th>Age (years)</th>
<th>Laterality</th>
<th>Appearance</th>
<th>Typical features</th>
<th>Color score</th>
<th>Picture</th>
<th>Ref</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mature cystic teratoma (dermoid)</td>
<td>Median, 33</td>
<td>Uni (88%)</td>
<td>Uni- (58%) or multilocular (or uni-/multilocular-solid)</td>
<td>Mixed echogenicity/white ball and stripes/shadowing; CA 19-9 elevated in 30%</td>
<td>1/2/3</td>
<td>†</td>
<td></td>
</tr>
<tr>
<td>Struma ovarii (entirely or predominantly thyroid tissue); 3% of all ovarian teratomas</td>
<td>Median, 40</td>
<td>Uni/bi</td>
<td>Multilocular/multilocular-solid; rarely, papillations; fluid anechoic or low-level</td>
<td>'Struma pearl': smooth; roundish solid area; thyrotoxicosis may occur</td>
<td>1/2/3</td>
<td>174</td>
<td></td>
</tr>
<tr>
<td>Epithelial</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Serous cystadenoma</td>
<td>40–60</td>
<td>Uni (80–90%)</td>
<td>Uni- or multilocular (2–10 locules)</td>
<td>Anechoic cystic fluid; often papillations without internal blood flow</td>
<td>1/2</td>
<td>‡</td>
<td></td>
</tr>
<tr>
<td>Serous cystadenofibroma</td>
<td>40–60</td>
<td>Uni (84%)</td>
<td>Multilocular (37%), unilocular-solid (30%), multilocular (19%) or unilocular (13%); median diameter 50–80 mm</td>
<td>1 (52%), 2 (17%) or 3 (13%) papillations; absent color Doppler signals (80%) and shadows behind papillations (40%)</td>
<td>1/2</td>
<td>175</td>
<td></td>
</tr>
<tr>
<td>Mucinous cystadenoma</td>
<td>Median, 50</td>
<td>Uni (95%)</td>
<td>Multilocular (65%) &gt;10 locules; sometimes unilocular (18%) or multilocular-solid (16%); median diameter 112 mm</td>
<td>Sometimes 'honeycomb nodule'</td>
<td>1/2/3</td>
<td>176</td>
<td></td>
</tr>
</tbody>
</table>

Continued
## Table 1  Continued

<table>
<thead>
<tr>
<th>Category/type</th>
<th>Age (years)</th>
<th>Laterality</th>
<th>Appearance</th>
<th>Typical features</th>
<th>Color score</th>
<th>Picture</th>
<th>Ref</th>
</tr>
</thead>
<tbody>
<tr>
<td>Brenner tumor (99% benign)</td>
<td>30–70</td>
<td>Uni</td>
<td>Small solid tumors, 20–60 mm; often extensive calcifications; sometimes multilocular-solid</td>
<td>Small cysts often seen in solid tumors; shadowing; CA 125 raised in 10%</td>
<td>1/2/(3)</td>
<td><img src="image1.png" alt="Image" /></td>
<td>177 178</td>
</tr>
<tr>
<td>Tumor-like lesions</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Abscess</td>
<td>16–50</td>
<td>Uni/bi</td>
<td>Uni-/multilocular</td>
<td>Cogwheel appearance; mixed echogenicity; acute pain; raised CA 125</td>
<td>3/4</td>
<td><img src="image2.png" alt="Image" /></td>
<td>178</td>
</tr>
<tr>
<td>Malignant tumors</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Epithelial</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Borderline serous</td>
<td>Median, 42; 30% &lt;40</td>
<td>Uni (73%)/bi (27%)</td>
<td>Unilocular-solid (55%) or multilocular-solid (30%); cystic fluid anechoic (47%) or low-level</td>
<td>&gt;3 irregular papillations (81%) with internal blood flow and anechoic spaces; no shadowing</td>
<td>2/3</td>
<td><img src="image3.png" alt="Image" /></td>
<td>179 180 181 182</td>
</tr>
<tr>
<td>Borderline mucinous (intestinal type) (30–50%)</td>
<td>Median, 50</td>
<td>Uni</td>
<td>Multilocular (80%) or unilocular (15%); very large tumor (median diameter 195 mm)</td>
<td>Multiple small loculi, often ‘honeycomb nodule”; no papillations; cystic fluid low-level</td>
<td>2/3</td>
<td><img src="image4.png" alt="Image" /></td>
<td>176 180</td>
</tr>
</tbody>
</table>

Continued
<table>
<thead>
<tr>
<th>Category/type</th>
<th>Age (years)</th>
<th>Laterality</th>
<th>Appearance</th>
<th>Typical features</th>
<th>Color score</th>
<th>Picture</th>
<th>Ref</th>
</tr>
</thead>
<tbody>
<tr>
<td>Borderline mucinous (endocervical type)</td>
<td>30–40</td>
<td>Uni</td>
<td>Unilocular-solid; sometimes multilocular-solid; median diameter 37 mm</td>
<td>Papillations (60%); cystic fluid low-level or ground-glass</td>
<td>2/3</td>
<td></td>
<td>176 180</td>
</tr>
<tr>
<td>Borderline seromucinous (new category)</td>
<td>Median, 42</td>
<td>Uni</td>
<td>Contain endometrioid-, indifferent- and squamous-type epithelium</td>
<td>Frequently associated with endometriosis</td>
<td>—</td>
<td></td>
<td>176 180</td>
</tr>
<tr>
<td>Low-grade serous carcinoma</td>
<td>Median, 53</td>
<td>Bi (60%)</td>
<td>Multilocular-solid (55%) or solid (32%)</td>
<td>Small calcifications in solid tissue; papillations (32%)</td>
<td>2/3/4</td>
<td></td>
<td>181</td>
</tr>
<tr>
<td>High-grade serous carcinoma</td>
<td>55–65</td>
<td>Bi (50%)</td>
<td>Solid (64%) or multilocular-solid (33%)</td>
<td>Areas of necrosis in solid tissue; rarely, papillations (7%)</td>
<td>2/3/4</td>
<td></td>
<td>181</td>
</tr>
<tr>
<td>Mucinous carcinoma (3%)</td>
<td>Median, 53</td>
<td>Uni (80%)</td>
<td>Multilocular-solid (55%), multilocular or solid</td>
<td>Very large tumor (median diameter 197 mm); cystic fluid low-level</td>
<td>2/3/4</td>
<td></td>
<td>176</td>
</tr>
</tbody>
</table>

Continued
<table>
<thead>
<tr>
<th>Category/type</th>
<th>Age (years)</th>
<th>Laterality</th>
<th>Appearance</th>
<th>Typical features</th>
<th>Color score</th>
<th>Picture</th>
<th>Ref</th>
</tr>
</thead>
<tbody>
<tr>
<td>Endometrioid carcinoma (10–15%)</td>
<td>Median, 55</td>
<td>Uni (79%); co-exist with endometrial carcinoma (20%)</td>
<td>Multilocular-solid (48%) with low-level (53%) or ground-glass (16%) cystic fluid, or solid (34%); median diameter 102 mm</td>
<td>Cockade-like appearance; papillations in 29%; 20% develop from endometriosis</td>
<td>(2)/3/4</td>
<td><img src="image1" alt="Image" /></td>
<td>183</td>
</tr>
<tr>
<td>Clear cell carcinoma (5–25%)</td>
<td>Median, 55</td>
<td>Uni (85%)</td>
<td>Multilocular-solid (41%) or unilocular-solid (35%) with low-level (44%) or ground-glass (22%) cystic fluid, or solid (24%); median diameter 117 mm</td>
<td>Solid nodules; papillations in 38%; 20–30% develop from endometriosis</td>
<td>(2)/3/4</td>
<td><img src="image2" alt="Image" /></td>
<td>184</td>
</tr>
<tr>
<td>Carcinosarcoma</td>
<td>Median, 66 (range 33–91)</td>
<td>Bi (50%)</td>
<td>Solid (72.5%); multilocular-solid (24.5%); median diameter 100 mm</td>
<td>Most tumors solid with irregular margins and cystic areas</td>
<td>3/4</td>
<td><img src="image3" alt="Image" /></td>
<td>§</td>
</tr>
<tr>
<td>Sex cord-stromal tumor</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Granulosa cell tumor (70%)</td>
<td>50% premenopausal; Uni 3–10% prepubertal (juvenile type)</td>
<td>Large multilocular-solid/solid; median diameter 100 mm; heterogeneous solid tissue with areas of necrosis and hemorrhage; echogenicity of fluid mixed or low-level; rarely, papillations</td>
<td>'Swiss cheese’ pattern; hyper-estrogenic (abnormal bleeding, thick endometrium); CA 125 normal; estradiol elevated in postmenopause</td>
<td>3/4</td>
<td><img src="image4" alt="Image" /></td>
<td>185</td>
<td></td>
</tr>
<tr>
<td>Sertoli-Leydig-cell tumor</td>
<td>≤30 (75%)</td>
<td>Uni (100%)</td>
<td>Large multilocular-solid or solid; median diameter 50–150 mm</td>
<td>Endocrine symptoms (one-third virilization); testosterone/androstenedione</td>
<td>3/4</td>
<td><img src="image5" alt="Image" /></td>
<td>173</td>
</tr>
</tbody>
</table>

*Table 1 Continued*
<table>
<thead>
<tr>
<th>Category/type</th>
<th>Age (years)</th>
<th>Laterality</th>
<th>Appearance</th>
<th>Typical features</th>
<th>Color score</th>
<th>Picture</th>
<th>Ref</th>
</tr>
</thead>
<tbody>
<tr>
<td>Germ cell tumor</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dysgerminoma</td>
<td>Median, 20 (range 16–31)</td>
<td>Uni</td>
<td>Highly vascularized, purely solid tumors with heterogeneous internal echogenicity divided into several lobules; smooth and sometimes lobulated contour; well-defined relative to surrounding organs</td>
<td>Internal lobular appearance; raised LDH, sometimes AFP</td>
<td>3/4</td>
<td><img src="image1.png" alt="Image" /></td>
<td>186</td>
</tr>
<tr>
<td>Yolk sac tumor*</td>
<td>20–30</td>
<td>Uni</td>
<td>Large and irregular multilocular-solid/solid (100–200 mm)</td>
<td>Fine-textured slightly hyperechoic solid tissue; raised AFP</td>
<td>3/4</td>
<td><img src="image2.png" alt="Image" /></td>
<td>187, 188</td>
</tr>
<tr>
<td>Immature teratoma</td>
<td>15–30</td>
<td>Uni</td>
<td>Large, predominantly solid</td>
<td>Very inhomogeneous solid tissue with hyper-reflective areas; raised AFP</td>
<td>2/3/4</td>
<td><img src="image3.png" alt="Image" /></td>
<td>¶</td>
</tr>
<tr>
<td>Choriocarcinoma</td>
<td>Median, 36</td>
<td>Uni</td>
<td>Large, solid (inhomogeneous echogenicity) with small and irregular cystic spaces</td>
<td>Raised hCG</td>
<td>(3/4)</td>
<td><img src="image4.png" alt="Image" /></td>
<td>189</td>
</tr>
</tbody>
</table>

*Yolk sac tumor*
<table>
<thead>
<tr>
<th>Category/type</th>
<th>Age (years)</th>
<th>Laterality</th>
<th>Appearance</th>
<th>Typical features</th>
<th>Color score</th>
<th>Picture Ref</th>
</tr>
</thead>
<tbody>
<tr>
<td>Embryonal carcinoma</td>
<td>14–20</td>
<td>Uni</td>
<td>Large, solid (inhomogeneous echogenicity) with small and irregular cystic spaces</td>
<td>Raised hCG and AFP</td>
<td>(3)/4</td>
<td>189</td>
</tr>
<tr>
<td>Malignant mixed germ cell tumor</td>
<td>Median, 18</td>
<td>Uni</td>
<td>Large, solid (inhomogeneous echogenicity) with small and irregular cystic spaces</td>
<td>Raised hCG/LDH/AFP</td>
<td>(3)/4</td>
<td>189</td>
</tr>
<tr>
<td>Secondary metastatic</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Breast, stomach, lymphoma or uterus</td>
<td>Median, 56</td>
<td>Bi (50–75%)/uni</td>
<td>Solid; median diameter 70 mm</td>
<td>'Lead-vessel' sign; CA 125 moderately raised in 75%; CA 15-3 raised (breast)</td>
<td>3/4</td>
<td>190</td>
</tr>
<tr>
<td>Colon, rectum, appendix or biliary tract</td>
<td>Median, 56; appendix younger</td>
<td>Bi (50–75%)/uni</td>
<td>Multilocular/multilocular-solid; median diameter 120 mm; many locules; irregular; papillations</td>
<td>CA 125 moderately raised in 75%; CEA raised (colon, rectum); CA19-9 raised (biliary tract)</td>
<td>(2)/3/(4)</td>
<td>190</td>
</tr>
<tr>
<td>Tumor of Fallopian tube: epithelial</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Continued
Joint statement

Cancers and secondary metastatic tumors. The IOTA ADNEX model was developed and validated using parameters collected by experienced ultrasound examiners.36 Several external validation studies have shown good to excellent performance of the ADNEX model in discriminating different types of ovarian tumor, with a higher clinical value than the RMI.55–61 A study aiming to validate the ADNEX model when applied by Level II examiners has confirmed that it can be used successfully by less-experienced examiners.62 A large multicenter cohort study of 4905 masses in 17 centers, comparing six different prediction models (RMI, LR2, Simple Rules, Simple Rules risk model, and ADNEX model with or without CA 125), demonstrated the IOTA ADNEX model and the IOTA Simple Rules risk model to be the best models for the characterization of ovarian masses in patients who present with an adnexal lesion.63

Gynecologic Imaging Reporting and Data System (GI-RADS)
The Gynecologic Imaging Reporting and Data System (GI-RADS) was first introduced by Amor et al64 in 2009 and was validated prospectively by the same team in a multicenter study 2 years later.65 This reporting system quantifies the risk of malignancy into five categories: GI-RADS 1, definitively benign (estimated probability of malignancy (EPM) 0%); GI-RADS 2, very probably benign (EPM <1%); GI-RADS 3, probably benign (EPM 1–4%); GI-RADS 4, probably malignant (EPM 5–20%); and GI-RADS 5, very probably malignant (EPM >20%). More recently, several studies have demonstrated the value of the GI-RADS system for the assessment of malignant adnexal masses in women who are candidates for surgical intervention. Furthermore, the addition of GI-RADS to CA 125 improves the identification of adnexal masses at high risk of malignancy compared with using CA 125 alone.66–71

Ovarian-Adnexal Reporting and Data System (O-RADS)
The Ovarian-Adnexal Reporting and Data System (O-RADS) lexicon for ultrasound was published in 2018, providing a standardized glossary that includes all appropriate descriptors and definitions of the characteristic ultrasound appearance of normal ovaries and various adnexal lesions.72 73 The O-RADS ultrasound working group developed an adnexal mass triage system based either on the O-RADS descriptors or on the risk of malignancy assigned to the mass using the IOTA ADNEX model to classify ovarian tumors into different risk categories.74 However, at present, neither the triage system nor the O-RADS descriptors have been externally validated. Basha et al75 determined the malignancy rates, validity, and reliability of the O-RADS approach when applied to a database of 647 adnexal masses collected before the development of the O-RADS system. In this retrospective study, the O-RADS system had significantly higher sensitivity than did the GI-RADS system and the IOTA Simple Rules, with a non-significant slightly lower specificity compared with both GI-RADS and IOTA Simple Rules, and with similar reliability.

Statements on ultrasonography (Statements 1–6)
1. Subjective assessment by expert (Level III) ultrasound examiners has the best performance to distinguish between benign and malignant ovarian tumors.
   - Level of evidence: 1a
   - Grade of statement: A

---

Table 1

<table>
<thead>
<tr>
<th>Category/Type</th>
<th>Age (years)</th>
<th>Laterality</th>
<th>Appearance</th>
<th>Typical features</th>
<th>Color score</th>
<th>Picture</th>
<th>Ref</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tubal cancer</td>
<td>55–60</td>
<td>Uni (90%)</td>
<td>Completely solid or with large solid component(s) and anechoic cystic fluid; average 50 mm</td>
<td>Well-vascularized ovoid or sausage-shaped structure; normal ovarian tissue adjacent in 50%</td>
<td>3/4</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

All example images in this table are reproduced from the cited references in Ultrasound in Obstetrics and Gynecology. Color score indicates amount of blood flow within lesion, classified as 1: no detectable flow (score, 1); 2: minimal blood flow (score, 2); 3: abundant blood flow (score, 3); score in parentheses under ‘Appearance’ is blood flow score, J: moderate flow; L: low flow; H: high flow.§Ciccarone et al (personal communication).¶Landolfo et al (personal communication).AFP, alpha-fetoprotein; Bi, bilateral; hCG, human chorionic gonadotropin; LDH, lactate dehydrogenase; postmeno, post-menopausal; premeno, pre-menopausal; Ref, reference; Uni, unilateral.
## Joint statement

<table>
<thead>
<tr>
<th>Model or system: type</th>
<th>Predictor variables</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Simple descriptors: classification as benign or malignant</td>
<td>Benign descriptor (BD) 1: unilocular tumor with ground-glass echogenicity in a pre-menopausal woman BD2: unilocular tumor with mixed echogenicity and acoustic shadows in a pre-menopausal woman BD3: unilocular anechoic tumor with regular walls and maximum diameter of lesion &lt;10 cm BD4: remaining unilocular tumor with regular walls Malignant descriptor (MD) 1: Tumor with ascites and at least moderate color Doppler blood flow in a post-menopausal woman MD2: age &gt;50 years and CA 125 &gt;100 U/mL</td>
<td>No risk estimates Based on clinical, ultrasound and CA 125 information Possible to calculate result without computer</td>
</tr>
<tr>
<td>RMI: score</td>
<td>CA 125, menopausal status, ultrasound score based on five binary ultrasound variables (multilocular cyst, solid areas, bilateral lesions, ascites, evidence of metastases on abdominal ultrasound)</td>
<td>No risk estimates Based on clinical, ultrasound and CA 125 information Possible to calculate result without computer Online calculators available</td>
</tr>
<tr>
<td>Simple Rules: classification as benign, inconclusive or malignant</td>
<td>Classification based on 10 binary features—five benign and five malignant features: Benign features: unilocular cyst, smooth multilocular cyst with largest diameter &lt;100 mm, presence of solid areas with largest diameter &lt;7 mm, acoustic shadows, no vascularization on color Doppler Malignant features: irregular solid tumor, irregular multilocular solid tumor with largest diameter ≥100 mm, presence of ascites, ≥4 papillary projections, very strong vascularization on color Doppler</td>
<td>No risk estimates Classification into only three groups Based on dichotomized ultrasound features Easy to use without computer Available as smartphone app</td>
</tr>
<tr>
<td>LR2: risk model based on logistic regression</td>
<td>Age (years), presence of acoustic shadows, presence of ascites, presence of papillary projections with blood flow, maximum diameter of largest solid component, irregular internal cyst walls</td>
<td>Risk estimates Based on clinical and ultrasound information Requires computer Available as smartphone app</td>
</tr>
<tr>
<td>Simple Rules risk: risk model based on logistic regression</td>
<td>The 10 binary features used in the Simple Rules, type of center (oncology center vs other)</td>
<td>Risk estimates Based on dichotomized ultrasound features Developed to add risk estimates for Simple Rules Available as online calculator; available in ultrasound machines from some manufacturers</td>
</tr>
<tr>
<td>ADNEX without CA 125: risk model based on multinomial logistic regression</td>
<td>Age (years), maximum diameter of lesion (mm), maximum diameter of largest solid component (mm), number of papillary projections (ordinal), presence of acoustic shadows, presence of ascites, presence of more than 10 cyst locules, type of center (oncology center vs other)</td>
<td>Risk estimates Also estimates risk of four subtypes of malignancy Based on clinical and ultrasound information Subjective predictors are avoided a priori (eg, color score or irregular cyst walls) Requires computer Available as smartphone app and as online calculator; available in ultrasound machines from some manufacturers</td>
</tr>
</tbody>
</table>

Continued
Table 2  Continued

<table>
<thead>
<tr>
<th>Model or system: type</th>
<th>Predictor variables</th>
<th>Remarks</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADNEX with CA 125: risk model based on multinomial logistic regression</td>
<td>Same variables as for ADNEX without CA 125, and additionally serum CA 125 (IU/L)</td>
<td>Risk estimates based on clinical, ultrasound, and CA 125 information. Also estimates risk of four subtypes of malignancy: Based on clinical, ultrasound, and CA 125 information. Subjective predictors are avoided a priori (e.g., color score or irregular cyst walls). Requires computer. Available as smartphone app and as online calculator; available in ultrasound machines from some manufacturers.</td>
</tr>
</tbody>
</table>

ADNEX, Assessment of Different NEOplasias in the adneXa; CA 125, cancer antigen 125; RMI, risk of malignancy index.

1. Consensus: yes, 95% (n=19); no, 0% (n=0); abstain, 5% (n=1)
2. If an expert ultrasound examiner is not available, the use of ultrasound-based diagnostic models can assist clinicians to distinguish between benign and malignant ovarian tumors.
   - Level of evidence: 2a
   - Grade of statement: B
   - Consensus: yes, 90% (n=18); no, 0% (n=0); abstain, 10% (n=2)
3. Ultrasound-based diagnostic models (IOTA Simple Rules risk model or IOTA ADNEX model) are preferable to CA 125 level, HE4 level, or ROMA as they are superior in distinguishing between benign and malignant ovarian tumors.
   - Level of evidence: 2b
   - Grade of statement: B
   - Consensus: yes, 95% (n=19); no, 0% (n=0); abstain, 5% (n=1)
4. The IOTA ADNEX model and the IOTA Simple Rules risk model are recommended as they outperform existing morphological scoring systems, including the RMI.
   - Level of evidence: 1b
   - Grade of statement: A
   - Consensus: yes, 95% (n=19); no, 0% (n=0); abstain, 5% (n=1)
5. The IOTA ADNEX model is a multiclass model and is helpful to differentiate between benign tumors, borderline tumors, early- or advanced-stage ovarian cancer, and secondary metastatic tumors.
   - Level of evidence: 3b
   - Grade of statement: C
   - Consensus: yes, 85% (n=17); no, 0% (n=0); abstain, 15% (n=3)
6. The threshold risk of there being a secondary metastatic tumor (as predicted by the IOTA ADNEX model), above which additional investigations to detect the primary organ of origin should be triggered, is 10%.
   - Level of evidence: 5
   - Grade of statement: D
   - Consensus: yes, 95% (n=19); no, 0% (n=0); abstain, 5% (n=1)
7. CA 125 is the best single-protein biomarker for the pre-operative characterization of ovarian tumors. However, it is not useful as a screening test for ovarian cancer.
   - Level of evidence: 2b
   - Grade of statement: B
   - Consensus: yes, 95% (n=19); no, 0% (n=0); abstain, 5% (n=1)
8. Neither HE4 nor ROMA improves the discrimination between benign and malignant masses compared with CA 125 alone.
   - Level of evidence: 2b
   - Grade of statement: B
   - Consensus: yes, 70% (n=14); no, 0% (n=0); abstain, 30% (n=6)

Levels of evidence and grades are described in Online supplemental appendix 2.

Tumor markers
According to a systematic quantitative review assessing the accuracy of CA 125 level in the diagnosis of benign, borderline, and malignant ovarian tumors, CA 125 is the best available single-protein biomarker identified to date.46 Although it lacks sensitivity and specificity for early stages of the disease and has a relatively low specificity overall, it can help direct treatment options in patients with suspicious ovarian masses. Pooled analyses have highlighted that a high body mass index and ethnicity might influence CA 125 levels, representing an additional diagnostic challenge.73 Other factors that influence CA 125 levels are the age of the patient, pregnancy, inflammatory processes, and the presence of fibroids or endometriosis.77–80

Multiple studies, including meta-analyses, have highlighted the role of HE4 as a potential complement to CA 125, especially in differentiating benign endometriotic and inflammatory lesions in younger women.25 81–103 Additional tumor markers (as in the ROMA test) have failed to improve significantly the discrimination between benign and malignant masses compared with CA 125 alone.53 81 84 91 96–109 The combination of a more extended tumor marker profile, including the addition of carcinoembryonic antigen (CEA) and/or carbohydrate antigen (CA 19-9) to CA 125, is useful mainly for differentiating between metastatic tumors from the gastrointestinal tract or pancreas and primary ovarian malignancy.110–113

Statements on tumor markers (Statements 7–12)

7. CA 125 is the best single-protein biomarker for the pre-operative characterization of ovarian tumors. However, it is not useful as a screening test for ovarian cancer.
   - Level of evidence: 2b
   - Grade of statement: B
   - Consensus: yes, 95% (n=19); no, 0% (n=0); abstain, 5% (n=1)
8. Neither HE4 nor ROMA improves the discrimination between benign and malignant masses compared with CA 125 alone.
   - Level of evidence: 2b
   - Grade of statement: B
   - Consensus: yes, 70% (n=14); no, 0% (n=0); abstain, 30% (n=6)
9. CA 125 does not increase the performance of ultrasound-based risk models to distinguish between benign and malignant tumors.
   - Level of evidence: 2b
   - Grade of statement: B
   - Consensus: yes, 60% (n=12); no, 10% (n=2); abstain, 30% (n=6)

10. CA 125 is helpful as a biomarker in cases of suspected malignancy and it helps to distinguish between sub-types of malignant tumors, such as borderline and early- and advanced-stage primary ovarian cancers and secondary metastatic tumors.
    - Level of evidence: 2b
    - Grade of statement: B
    - Consensus: yes, 90% (n=18); no, 5% (n=1); abstain, 5% (n=1)

11. CEA may be useful in specific cases to differentiate between primary ovarian cancer and secondary (ovarian) tumors.
    - Level of evidence: 3b
    - Grade of statement: C
    - Consensus: yes, 90% (n=18); no, 0% (n=0); abstain, 10% (n=2)

12. CA 19-9 can help to differentiate secondary metastatic tumors in the ovary.
    - Level of evidence: 3b
    - Grade of statement: C
    - Consensus: yes, 75% (n=15); no, 5% (n=1); abstain, 20% (n=4)

Levels of evidence and grades are described in Online supplemental appendix 2.

Magnetic resonance imaging/computed tomography/positron emission tomography-computed tomography

Magnetic resonance imaging

Several reports have found that magnetic resonance imaging (MRI), alone or in combination with computed tomography (CT), predicts accurately the presence of peritoneal carcinomatosis in patients undergoing pre-operative evaluation for cytoreductive surgery, particularly when the assessment is carried out by an experienced radiologist.114-117 Recently, a prospective study reported higher specificity of the IOTA LR2 model compared with subjective interpretation of MRI findings by an experienced radiologist, as well as similar sensitivities for both imaging modalities for discriminating between benign and malignant tumors.118 The addition of diffusion-weighted techniques to conventional imaging modalities has been shown in multiple pooled studies to increase diagnostic accuracy in discriminating between benign tumors and ovarian cancer, especially in the Caucasian population, with data even suggesting a value in predicting resectability.119-123 However, the true extent of such a benefit needs to be validated further in multicenter large-scale prospective randomized studies, which are currently being designed or underway.121 The addition of
Joint statement

Use of diffusion- and perfusion-weighted MRI

Management in specialized gynecological oncology

Expert subjective assessment or IOTA ADNEX with CA 125

CA 125, CEA

Other tumor markers in specific cases

(αFP, LDH, hCG, inhibin B, AMH, CA19.9, CA15.3)

Staging procedures: CT, MRI, PET-CT, diagnostic biopsies

Gynecological oncology specialist in charge or multidisciplinary team meeting

Borderline tumor

Early-stage primary ovarian cancer

Advanced stage primary ovarian cancer

Secondary metastasis

Options for fertility-sparing surgery if desired by patient

Therapy depending on staging and histopathological diagnosis

Investigation of other primary tumor origin

Figure 3 Flowchart of steps necessary to differentiate between subgroups of malignancy and extent of disease within gynecological oncology centers. *Early stage and advanced stage might differ according to different ADNEX models (stage I vs stages III–IV) and oncologically (stages I–II vs stages I–IV). αFP, alpha-fetoprotein; AMH, anti-Müllerian hormone; CA 125, cancer antigen 125; CA 15-3, cancer antigen 15-3; CA 19-9, carbohydrate antigen 19-9; CEA, carcinoembryonic antigen; CT, computed tomography; hCG, human chorionic gonadotropin; IOTA ADNEX, International Ovarian Tumor Analysis group Assessment of Different NEoplasias in the adnexa; LDH, lactate dehydrogenase; MRI, magnetic resonance imaging; PET-CT, positron emission tomography-computed tomography.

Quantitative dynamic contrast-enhanced MRI to diffusion-weighted imaging and anatomical MRI sequences and the development of a 5-point scoring system (O-RADS MRI score) is another modern diagnostic development with promising potential for the differentiation between benign and malignant adnexal masses in cases in which ultrasound is unable to arrive at a clear diagnosis (ie, indeterminate masses). When this technique is enhanced with volume quantification, it can help to discriminate between type I and type II epithelial ovarian cancers. However, there are only limited data available on the impact of these modern MRI techniques on clinical decision-making, and further studies are needed with larger sample populations.

Computed tomography

Dedicated multidetector CT protocols with standardized peritoneal carcinomatosis index forms are the most common diagnostic tool used in routine clinical practice to assess the extent of tumor dissemination and the presence of peritoneal carcinomatosis. A radiological peritoneal carcinomatosis index applied at pre-operative CT within an expert setting has been shown to have low performance scores as a triage test to identify patients who are likely to have complete cytoreduction to no macroscopic residual disease. On retrospective analysis, pre-operative CT imaging showed high specificity but rather low sensitivity in detecting tumor involvement at key sites in ovarian cancer surgery. Multiple studies that have attempted to cross-validate the accuracy of CT scans in predicting unresectable disease and incomplete cytoreduction have shown a substantial drop in accuracy rates when attempts have been made to validate them in other cohorts. Thus, CT should not be used as the sole tool to predict the resectability of peritoneal carcinomatosis and exclude patients from surgery; rather, the full clinical context should be taken into account. Its widespread availability makes CT useful as a first-line diagnostic tool to identify patients who should not be selected for cytoreductive surgery, such as those with large/multifocal intra-parenchymatous distant metastases, acute thromboembolic events, or secondary metastatic tumors that limit the prognosis. The role of radiomics as an additional quantitative mathematical segmentation of conventional pre-operative CT images has shown some promising results in preliminary studies; however, larger studies are necessary for validation before this technique is implemented in clinical practice.

Positron emission tomography-computed tomography

Positron emission tomography-computed tomography (PET-CT) may be useful in differentiating malignant from borderline or benign ovarian tumors, with the limitation that its diagnostic performance can be impacted negatively by certain tumor histological sub-types due to the lower fluorodeoxyglucose uptake in clear cell and mucinous invasive subtypes. PET-CT can also play a role as an additional technique in the diagnosis of
lymph node metastases, especially outside the abdominal cavity, or in characterizing unclear lesions in key areas that would alter clinical management (eg, chest lesions). However, PET-CT does not seem to be a relevant additional diagnostic modality for the true extent of peritoneal spread of ovarian cancer, specifically bowel and mesenteric serosa, and therefore fails to predict resectability in those key sites, especially in the presence of low-volume disease. Furthermore, PET-CT has been shown to have a low diagnostic value in differentiating borderline from benign tumors and should therefore not be used in clinical decision-making processes in that context, especially when considering fertility-sparing procedures.

Statements on MRI, CT, and PET/CT (Statements 13–17)

13. MRI with the inclusion of the functional sequences, dynamic contrast-enhanced and diffusion-weighted MRI, is not a first-line tool but may be used as a second-line tool after ultrasoundography to further differentiate between benign, malignant, and borderline masses.
   - Level of evidence: 2a
   - Grade of statement: B
   - Consensus: yes, 100% (n=20); no, 0% (n=0); abstain, 0% (n=0)

14. PET-CT and whole-body diffusion MRI as a second step can help to detect non-ovarian origin of secondary metastatic tumors if suspicions are raised by the initial ultrasound examination.
   - Level of evidence: 4
   - Grade of statement: C
   - Consensus: yes, 90% (n=18); no, 0% (n=0); abstain, 10% (n=2)

15. PET-CT cannot differentiate reliably between borderline and benign tumors.
   - Level of evidence: 4
   - Grade of statement: C
   - Consensus: yes, 95% (n=19); no, 0% (n=0); abstain, 5% (n=1)

16. Imaging alone cannot detect reliably the entire extent of either peritoneal carcinomatosis (especially in cases of small-volume carcinomatosis) or mesenteric and bowel serosal involvement.
   - Level of evidence: 3b
   - Grade of statement: B
   - Consensus: yes, 85% (n=17); no, 5% (n=1); abstain, 10% (n=2)

17. Imaging alone should not be used for surgical decision-making in terms of the prediction of peritoneal tumor resectability.
   - Level of evidence: 3b
   - Grade of statement: B
   - Consensus: yes, 80% (n=16); no, 15% (n=3); abstain, 5% (n=1)

Levels of evidence and grades are described in Online supplemental appendix 2.

Circulating cell-free DNA and circulating tumor cells

Circulating cell-free DNA and circulating tumor cells as non-invasive cancer biomarkers and in non-invasive biopsy (sometimes called ‘liquid biopsy’) have been investigated in multiple studies. DNA methylation patterns in cell-free DNA show potential to detect a proportion of ovarian cancers up to 2 years in advance of diagnosis. They may potentially guide personalized treatment, even though validation studies are lacking. The prospective use of novel collection vials, which stabilize blood cells and reduce background DNA contamination in serum/plasma samples, will facilitate the clinical implementation of liquid biopsy analyses. A prospective evaluation of the potential of cell-free DNA for the diagnosis of primary ovarian cancer using chromosomal instability as a read-out suggested that this might be a promising method to increase the specificity of the pre-surgical prediction of malignancy in patients with adnexal masses. However, even though these circulating biomarkers play a key role in understanding metastasis and tumorigenesis and provide comprehensive insight into tumor evolution and dynamics during treatment and disease progression, they still have not been established as part of routine clinical practice.

One meta-analysis suggested that quantitative analysis of cell-free DNA has unsatisfactory sensitivity but acceptable specificity for the diagnosis of ovarian cancer. In a more recent meta-analysis, cell-free DNA appeared to be slightly better than CA 125 and similar to HE4 with respect to its diagnostic ability to discriminate individuals with from those without ovarian cancer. Nevertheless, the diagnostic value of cell-free DNA in patients with ovarian cancer remains unclear and the data should be interpreted with caution. Further large-scale prospective studies are strongly recommended to validate the potential applicability of using circulating cell-free DNA, alone or in combination with conventional markers, as a diagnostic biomarker for ovarian cancer, and to explore potential factors that may influence the accuracy of ovarian cancer diagnosis.

Statement on circulating cell-free DNA and tumor cells (Statement 18)

18. Circulating cell-free DNA and circulating tumor cells should not yet be used in routine clinical practice to differentiate between benign and malignant ovarian masses.
   - Level of evidence: 4
   - Grade of statement: C
   - Consensus: yes, 85% (n=17); no, 5% (n=1); abstain, 10% (n=2)

Levels of evidence and grades are described in Online supplemental appendix 2.

OVERVIEW OF CONSENSUS

The experts also reached a consensus on a flowchart describing steps recommended to distinguish between benign and malignant tumors (Figure 2) and to direct patients towards appropriate treatment pathways. Ultrasonography is recommended as a first step to stratify patients with symptoms suggestive of an adnexal mass, and in those with an incidental finding of an adnexal mass on imaging. If the scan rules out normal ovaries and physiological changes (ie, rules out O-RADS 1), the IOTA ADNEX model could be applied as a next step in order to determine the risk of malignancy. Any ultrasonographic examination in the case of a suspected ovarian mass should be performed by an expert sonographer. The resulting classification of the lesion into one of the O-RADS categories can

further guide the management and selection of patients for referral to a dedicated gynecological oncology center.

A consensus was also reached on further steps necessary to differentiate between sub-groups of malignancy and extent of disease within gynecological oncology centers (Figure 3). Ultrasound assessment by an expert or application of the IOTA ADNEX model in combination with the tumor marker profile (CA 125 and CEA, complemented with other markers in specific cases) can often indicate the specific sub-type of malignancy. If available, diagnosis of the primary lesion can be confirmed with diffusion- and perfusion-weighted MRI, especially in cases in which fertility-sparing surgery is considered. A CT scan of chest, abdomen, and pelvis is mandatory before planned surgery for presumed malignancy, in order to exclude secondary cancers, thromboembolic events, and multifocal intraparenchymal distant metastases that would preclude resectability. The final management and treatment journey of the patient should be determined within an expert multidisciplinary setting, taking into account both the diagnostic findings and the overall patient profile, including symptoms, patient preferences and prior surgical, medical and reproductive history, with the ultimate aim of defining an individualized approach for every patient.

Author affiliations
1 Gynecology and Obstetrics, University Hospitals KU Leuven, Leuven, Belgium
2 Development and Regeneration, KU Leuven, Leuven, Belgium
3 Clinical Research Unit, Institut Bergonie, Bordeaux, France
4 Metabolism Digestion and Reproduction, Queen Charlotte’s & Chelsea Hospital, Imperial College, London, UK
5 Woman, Child and Public Health, Fondazione Policlinico Universitario A Gemelli IRCCS, Rome, Italy
6 Gynaecology and Gynaecological Oncology, Evangelische Kliniken Essen-Mitte, Essen, Germany
7 Gynaecology and Obstetrics, University Clinic of Navarra, Madrid, Spain
8 Obstetrics and Gynaecology, First Faculty of Medicine, Charles University and General University Hospital in Prague, Prague, Czech Republic
9 Obstetrics and Gynaecology, Medical University of Innsbruck, Innsbruck, Austria
10 Radiology, University Clinic of Navarra, Madrid, Spain
11 European Network of Gynaecological Cancers Advocacy Groups (ENGAGE) Executive Group, Prague, Czech Republic
12 KU - Patient Organisation for Women with Gynaecological Cancer, Copenhagen, Denmark
13 Clinical Physiology, Nuclear Medicine & PET, Rigshospitalet, Copenhagen University Hospital, Copenhagen, Denmark
14 Gynecology and Obstetrics, Gynecologic Oncology Unit, Santa Chiara Hospital, Trento, Italy
15 Gynaecological Surgery, Institut Gustave Roussey, Villejuif, France
16 Gynecologic Oncology, Fondazione Policlinico Universitario A Gemelli IRCCS, Rome, Italy
17 Obstetrics and Gynecologic Oncology, University Hospital, Strasbourg, France
18 Obstetrics and Gynecology, Universita Cattolica del Sacro Cuore, Rome, Italy
19 Obstetric and Gynaecology, and Gynaecologic Oncology, University Hospital Leuven, Leuven Cancer Institute, Leuven, Belgium
20 Radiology, University Hospitals Leuven, Leuven, Belgium
21 Division of Translational MRI, Department of Imaging and Pathology, KU Leuven, Leuven, Belgium
22 Gynaecologic Oncology, Hammersmith Hospital, Imperial College, London, UK

Presented at
This paper is being published simultaneously and jointly, in International Journal of Gynecological Cancer, Ultrasound in Obstetrics & Gynecology and Facts, Views and Vision in ObGyn, by the European Society of Gynecological Oncology (ESGO), the International Society of Ultrasound in Obstetrics and Gynecology (ISUOG), the International Ovarian Tumour Analysis (IOTA) group and the European Society for Gynaecological Endoscopy (ESGE).

Twitter Guillermo Gallardo Madueño @gggallar

Acknowledgements The authors thank ESGO, ISUOG, IOTA, and ESAGE for their support. We wish also to express sincere gratitude to Maciej Malecki (University Hospital Leuven, Leuven, Belgium) for providing technical support during the conference call.

Contributors The development group (including all authors) is collectively responsible for the decision to submit for publication. CF (chair), DT (chair), and FP (methodologist) wrote the first draft of the manuscript. All other contributors have actively given personal input, reviewed the manuscript, and have given final approval before submission.

Funding All costs relating to the development process were covered from ESGO, ISUOG, IOTA and ESAGE funds. There was no external funding of the development process or manuscript production.

Competing interests DT: senior investigator FWO (Fund for Scientific Research Flanders), and research sponsored by Roche Diagnostics. KU Leuven has consultancy agreements with GE Healthcare, Samsung Healthcare, GSK, and Canon. TB: research sponsored by Roche Diagnostics, Samsung Medison and Illumina, and grants for traveling from Samsung Medison. LC: advisory boards for AstraZeneca, GSK, Takeda, and Roche. DC: advisory boards for Gennab, AstraZeneca, Roche, and Sotio. NC: advisory boards for AstraZeneca, Seattle Genetics, Mersana and eTheRNA Immunotherapies NV; grants for traveling from Roche, Gennab and Agenin, and educational fees from MSD and Medscape Oncology. AD: advisory boards for Roche, AstraZeneca, GSK/Tesaro, BICDAD, Clovis, Gennab/Seattle Genetics, Pfizer, and Agenin, and grants for traveling from Roche and AstraZeneca. IV: consulting activities for Agenin, AstraZeneca, Clovis Oncology, Carrick Therapeutics, Diebiopharm International SA, Deciphera Pharmaceuticals, Elevare Therapeutics, F Hoffmann-La Roche Ltd, Gennab, GSK, Immunogen Inc, Medical University of Vienna, Mersana, Millennium Pharmaceuticals, MSD, Novocure, Octimet Oncology NV, Oncoinvent AS, Pharmamar, Sotio,s, Tesaro Inc, Varstem Oncology, and Zenitalis, contracted research (via KU Leuven) from Oncoinvent AS and Gennab, grants (corporate sponsored research) from Agenin and Roche, and accommodations/travel expenses from Agenin, AMG, Tesaro, AZ/MSD, Clovis, Sequana and Ethicon, and grants for traveling from GSK and Roche. FP, CL, DF, WF, GG, BL, AL, LM, PM, DG, ACT, W and GS: no conflicts of interest.

Patient consent for publication Not required.

Provenance and peer review Not commissioned; internally peer reviewed.

Data availability statement All data relevant to the study are included in the article or uploaded as supplementary information.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, an indication of whether changes were made, and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/.

ORCID iDs
Dirk Timmerman http://orcid.org/0000-0002-3707-6645
Tom Bourne http://orcid.org/0000-0003-1421-6059
Chiara Landolfo http://orcid.org/0000-0001-9808-7957
Luis Chiva http://orcid.org/0000-0002-1908-3251
David Cibuła http://orcid.org/0000-0001-6387-9356
Nicole Concini http://orcid.org/0000-0002-9795-2643
Daniela Fischerova http://orcid.org/0000-0002-7224-3218
Weider Fremyman http://orcid.org/0000-0002-1398-9124
Guillermo Gallardo Madueño http://orcid.org/0000-0002-7502-3544
Denis Querleu http://orcid.org/0000-0002-3984-4812
Antonia Carla Testa http://orcid.org/0000-0003-2217-8726
Vincent Vandecaveye http://orcid.org/0000-0002-0890-3279
Christina Fotopoulou http://orcid.org/0000-0001-6375-9645

REFERENCES
Joint statement

Cancer 2006;106:589–98.


Gynecol Oncol 2009;112:422–36.


7 Vernooi F, Heintz P, Witteveen E, et al. The outcomes of ovarian cancer treatment are better when provided by gynecologic oncologists and in specialized hospitals: a systematic review. 
Gynecol Oncol 2007;105:801–12.


10 Timmerman D. The use of mathematical models to evaluate pelvic masses; can they beat an expert operator? 


12 Valent L. Prospective cross-validation of Doppler ultrasound examination and gray-scale ultrasound imaging for discrimination of benign and malignant pelvic masses. 


15 Education and Practical Standards Committee, European Federation of Societies for Ultrasound in Medicine and Biology, Minimum training recommendations for the practice of medical ultrasound. 


20 Dochez V, Ramanon M, Renaudeau C. Efficacy of HE4, CA125, risk of malignancy index and risk of ovarian malignancy index to detect ovarian cancer in women with presumed benign ovarian tumours: a prospective, multicentre trial. 

21 Al-Musalihi K, Al-Kindi M, Al Aisayer F, et al. Evaluation of HE4, CA-125, risk of ovarian malignancy index (ROMA) and risk of malignancy index (RMI) in the preoperative assessment of patients with adnexal mass. 


J Gynecol Oncol 2011;22:177–82.


25 Anton C, Carvalho FM, Oliveira EL, et al. A comparison of CA125, HE4, risk of malignancy index (ROMA), and risk malignancy index (RMI) for the classification of ovarian masses. 


Gynecol Oncol 2013;130:398–9.

28 Timmerman D, Van Calster B, Jurkovic D, et al. Inclusion of CA-125 does not improve mathematical models developed to distinguish between benign and malignant adnexal tumors. 

29 Timmerman D, Van Calster B, Bourne TH, et al. Terms, definitions and measurement to describe the sensitivity and specificity of aadnexal tumours: a consensus opinion from the International Ovarian Tumor Analysis (IOTA) group. 

30 Timmerman D, Testa AC, Bourne T, et al. Logistic regression model to distinguish between benign and malignant adnexal masses before surgery: a multicenter study by the International Ovarian Tumor Analysis group. 

31 Timmerman D, Van Calster B, Testa A, et al. Predicting the risk of malignancy in adnexal masses based on the simple rules from the International Ovarian Tumor Analysis group. 

BMJ 2014;349:g5920.


36 Timmerman D, Ameye L, Fischervoda D, et al. Simple ultrasound rules to distinguish between benign and malignant adnexal masses before surgery; prospective validation by IOTA group. 
BMJ 2013;341:e6398.

37 Hartman CA, Juliato CRT, Sarian LO, et al. Ultrasound criteria and CA 125 as predictive variables of ovarian cancer in women with adnexal tumours. 


Kelly PJ, Archbold P, Price JH, et al. Serum CA19.9 levels are commonly elevated in primary ovarian mucinous tumours but cannot be used to predict the histological subtype. *J Clin Pathol* 2010;63:169–73.


Nam EJ, Yun MJ, Oh YT, et al. Diagnosis and staging of primary ovarian cancer: correlation between PET/CT, Doppler US, and CT or MRI. Gynecol Oncol 2010;116:389–94.


Joint statement


188 Corrigendum. Ultrasound Obstet Gynecol 2020;56.


Appendix 1 Identification of scientific evidence: literature search in MEDLINE

<table>
<thead>
<tr>
<th>Research period</th>
<th>1 May 2015–1 May 2020</th>
</tr>
</thead>
</table>

### Indexing terms
- adnexal masses, alpha fetoprotein, assessment of different neoplasias in the adnexa, assessment of different neoplasias in the adnexa masses, assessment of different neoplasias in the adnexa model, benign ovarian masses, benign ovarian tumours, beta-human chorionic gonadotropin, biomarker, borderline tumours, carbohydrate antigen 19.9, carbohydrate antigen 125, carcinoembryonic antigen, cell-free deoxyribonucleic acid, circulating cancer cells, circulating cell-free deoxyribonucleic acid, circulating free deoxyribonucleic acid, circulating tumour cells, circulating tumour deoxyribonucleic acid, clinical routine, computed tomography, consensus statement, daily practice, diagnosis, diagnostic performance, diagnostic models, diffusion-weighted imaging, diffusion-weighted magnetic resonance imaging, dynamic contrast-enhanced magnetic resonance imaging, expert ultrasound examiners, first line test, functional sequences, gynecology imaging reporting and data system, human epididymis protein, imaging, imaging methods, immunohistochemical diagnosis, inhibin, international ovarian tumor analysis, international ovarian tumor analysis methods, international ovarian tumor analysis rules, intraoperative ultrasound, investigations, logistic regression 1 test, logistic regression 2 test, magnetic resonance imaging, malignant ovarian masses, malignant ovarian tumours, marker, maximum standardized uptake value, molecular biology, molecular marker, morphological scoring system, multivariate analysis, ovarian cancer, ovarian masses, ovarian tumours, ovary, positron emission tomography, positron emission tomography–computed tomography, pre-operative characterization, pre-operative diagnosis, prognostic factor, prognostic value, protein biomarker, risk factors, risk of malignancy score, risk of malignancy index, risk of ovarian malignancy algorithm, scoring system, screening test, secondary metastatic tumours, second line test, simple rules, simple rules risk, simple rules risk model, single protein biomarker, standardized uptake value, suspected malignancy, suspected metastatic tumour, test performances, threshold risk, transabdominal ultrasound, transvaginal ultrasound, tumour markers, ultrasonography, ultrasound, ultrasound (3D), ultrasound-based diagnostic models, ultrasound-based risk models, ultrasound examiners, vascular endothelial growth factor, vascular endothelial growth factor, whole body diffusion magnetic resonance imaging.

### Language
- English

### Study design
- Priority was given to high-quality systematic reviews, meta-analyses and validating cohort studies, but lower levels of evidence were also evaluated. Search strategy excluded editorials, letters and case reports. Reference list of each identified article was reviewed for other potentially relevant papers.
Table 3 Levels of evidence and grades of statement used in this Consensus Statement

<table>
<thead>
<tr>
<th>Levels of evidence</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>1a Systematic review (with homogeneity) of Level-1 diagnostic studies; or clinical decision rule with Level-1b studies from different clinical centers</td>
<td></td>
</tr>
<tr>
<td>1b Validating cohort study with good reference standards; or clinical decision rule tested within one clinical center</td>
<td></td>
</tr>
<tr>
<td>1c Absolute SpPins and SnNouts*</td>
<td></td>
</tr>
<tr>
<td>2a Systematic review (with homogeneity) of Level &gt; 2 diagnostic studies</td>
<td></td>
</tr>
<tr>
<td>2b Exploratory cohort study with good reference standards; or clinical decision rule after derivation, or validated only on split-sample or databases</td>
<td></td>
</tr>
<tr>
<td>3a Systematic review (with homogeneity) of studies Level ≥ 3b</td>
<td></td>
</tr>
<tr>
<td>3b Non-consecutive study; or without consistently applied reference standards</td>
<td></td>
</tr>
<tr>
<td>4 Case–control study, poor or non-independent reference standard</td>
<td></td>
</tr>
<tr>
<td>5 Expert opinion without explicit critical appraisal, or based on physiology, bench research or ‘first principles’</td>
<td></td>
</tr>
</tbody>
</table>

Grades of statement

<table>
<thead>
<tr>
<th>Code</th>
<th>Quality of evidence</th>
<th>Definition</th>
</tr>
</thead>
</table>
| A    | High                | Further research is very unlikely to change our confidence in the estimate of effect.  
|      |                     | ● Several high-quality studies with consistent results  
|      |                     | ● In special cases: one large, high-quality multicenter trial |
| B    | Moderate            | Further research is likely to have an important impact on our confidence in the estimate of effect and may change the estimate.  
|      |                     | ● One high-quality study  
|      |                     | ● Several studies with some limitations |
| C    | Low                 | Further research is very likely to have an important impact on our confidence in the estimate of effect and is likely to change the estimate.  
|      |                     | ● One or more studies with severe limitations |
| D    | Very low            | Any estimate of effect is very uncertain.  
|      |                     | ● Expert opinion  
|      |                     | ● No direct research evidence  
|      |                     | ● One or more studies with very severe limitations |

Note: A minus sign ‘−’ may be added to denote evidence that fails to provide a conclusive answer because it is either (a) a single result with a wide confidence interval; or (b) a systematic review with considerable heterogeneity. Such evidence is inconclusive, and therefore can only generate Grade D recommendations. *‘Absolute SpPins’ is a diagnostic finding whose specificity is so high that a positive result rules in the diagnosis; ‘Absolute SnNout’ is a diagnostic finding whose sensitivity is so high that a negative result rules out the diagnosis.*