### Evidence tables:

<table>
<thead>
<tr>
<th><strong>Author(s):</strong></th>
<th>Geomini <em>et al.</em>, 2009</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Settings:</strong></td>
<td>Women with adnexal mass before surgery.</td>
</tr>
<tr>
<td><strong>Participants:</strong></td>
<td>109 studies were included in the review: reporting on 21750 adnexal masses: 15490 benign, 5826 malignant (27%) and 434 (2%) of borderline malignancy.</td>
</tr>
<tr>
<td><strong>Study Design:</strong></td>
<td>Systematic review and meta-analysis. The included studies were observational, at least 56% were prospective, in 77% blinding of the pathologist was not mentioned and in 14% verification bias could not be excluded. Literature search included papers published up to 2008</td>
</tr>
<tr>
<td><strong>Target Condition:</strong></td>
<td>The target condition was ovarian malignancy; the reference standard test was the histopathological diagnosis following surgery.</td>
</tr>
<tr>
<td><strong>Tests:</strong></td>
<td>Index and comparator tests were diagnostic models predicting malignancy in ovarian masses. Models had to contain at least two parameters. 83 models were reported in the included studies: incorporating ultrasound parameters, age, menopausal status and CA 125 level. Some models relied on ultrasound parameters only (Sassone, Alcazar, Lerner, Ferrazzi, DePriest) others included additional parameters such as age, CA-125 level, and menopausal status (RMI I to IV, Tailor) The model with the optimal combination of sensitivity and specificity was the RMI I: sensitivity 78% (95% CI 71 to 85%), specificity 87% (95% CI 83 to 91%) to with a cut off value of 200). See evidence summary for the estimated accuracy of models for prediction of malignancy on ultrasound parameters.</td>
</tr>
<tr>
<td><strong>Follow Up:</strong></td>
<td>Not applicable.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th><strong>Author(s):</strong></th>
<th>Im <em>et al.</em>, 2005</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Settings:</strong></td>
<td>Women undergoing surgical exploration for a pelvic mass at one of 6 university hospitals or a large tertiary community hospital. Exclusion criteria were age &lt; 18 years or prior invasive gynaecologic malignancy.</td>
</tr>
</tbody>
</table>
Participants:
1035 women: 318 (30.7%) with primary malignancy, 50 (4.8%) with metastases to the ovaries and 667 benign masses. Women of 50 years or older were assumed to be postmenopausal. The prevalence of ovarian cancer was 77/454 (17%) in premenopausal women and 240/530 (45%) in postmenopausal women.

Study Design:
Retrospective case series.

Target Condition:
Identification of ovarian malignancy. The reference standard was histopathology of the surgical specimen.

Tests:
Pelvic examination (in post-menopausal women)
Likelihood ratio for a positive test result 2.30
Likelihood ratio for a negative test result 0.50

Follow Up:
Not applicable.

Notes:
Study was designed to validate the SCO and ACOG referral guidelines for women with pelvic masses suspicious of ovarian cancer, and reports the predictive value of the various referral criteria.

Author(s): Kinkel et al., 2000

Settings:
Systematic review of studies of US characterisation of adnexal masses, not discovered during screening for ovarian cancer.

Participants:
The review included a total of 46 studies with 5159 patients: The studies used the following ultrasound techniques: morphologic assessment (34 studies, N=3377), Doppler arterial resistance measurement (24 studies, N=2712), colour Doppler flow imaging (10 studies, N=1408) and combined techniques (7 studies, N=832)

Study Design:
Systematic review and meta analysis

**Target Condition:**

The target condition was the identification of malignancy in adnexal masses. The reference standard was histopathology.

**Tests:**

Ultrasound, four techniques were considered: morphologic assessment, Doppler arterial resistance measurement, colour Doppler flow imaging and combined techniques

**Results - diagnostic accuracy for benign versus malignant masses**

The meta-analysis calculated summary ROC curves for each of the US techniques, and obtained the Q* statistic in each case. The Q* values correspond to the point on the summary ROC curve where sensitivity and specificity are equal.

- combined techniques: Q* 0.92 (95% C.I. 0.87 to 0.96)
- morphologic assessment: Q* 0.85 (95% C.I. 0.83 to 0.88)
- Doppler arterial resistance measurement: Q* 0.82 (95% C.I. 0.78 to 0.86)
- colour Doppler flow imaging: Q* 0.73 (95% C.I. 0.58 to 0.87)

**Subgroup analyses**

The authors considered the following covariates: year of publication; proportion of pre-menstrual women in each study; proportion of women with mucinous tumours, endometriomas, and non-neoplastic cysts; prevalence of malignancy; stage distribution of ovarian cancer; study design (prospective versus retrospective); diagnostic criteria; US technical factors (like transabdominal versus endovaginal probes); country of publication; specialty of the person reading the US images (radiology versus gynaecology).

These factors did not have a statistically significant effect on diagnostic accuracy except in the following cases:

- Prevalence of malignancy (34 studies, P=0.05) but only in US morphologic assessment studies. The accuracy of ultrasound was higher in studies with lower prevalence of malignancy.
- Diagnostic criteria (89 studies, P=0.2). Accuracy of US was higher in studies which used validated diagnostic criteria.
- Percentage of mucinous tumours (P=0.03, 28 studies). Accuracy of US was higher in studies with fewer cases of mucinous tumours.

**Author(s):** Liu et al., 2007

**Settings:**

Women with adnexal mass (not discovered during screening for ovarian cancer), who had
Participants:

69 studies with 6364 patients. Ultrasound was evaluated in 65 studies with 126 data sets, of these 54 articles with 58 data sets (5524 patients) used morphologic information alone. Colour/power Doppler were used in 42 studies. Combined morphologic and colour/power Doppler were used in 7 studies. Literature search included papers published between 1990 and 2006.

Menopausal status was mentioned in 34/69 studies. There were 2016/3125 (64.5%) premenopausal women in these 34 studies.

At least 49% of studies were prospective, at least 53% of studies used blinded interpretation of test results but reporting of the study population was inadequate in 36% of the studies.

Prevalence of malignant tumours was 24%.

Study Design:

Systematic review and meta-analysis.

Target Condition:

Target condition was identification of malignancy in adnexal mass, the reference standard was histopathology of the adnexal mass.

Results:

Any ultrasound: sensitivity 89% (95% CI 88 to 90%), specificity 84% (82% to 86%)

Morphologic assessment ultrasound: sensitivity 85% (95% CI 83 to 87%), specificity 83% (81% to 85%)

Colour Doppler flow imaging: sensitivity 75% (95% CI 72 to 77%), specificity 73% (71% to 75%)

Combined Doppler and morphologic US: sensitivity 87% (95% CI 85 to 90%), specificity 88% (85% to 91%)

Contrast enhanced US: sensitivity 90% (95% CI 87 to 93%), specificity 89% (87% to 91%)

Follow Up:

not applicable

Notes:

The review does not report the setting of each study (primary, secondary or tertiary care), unclear what diagnostic tests women had already had before the ultrasound.

Author(s): Medeiros et al., 2009a
### Settings:

Women with clinically suspected adnexal mass, evaluated using 5 MHz transvaginal probe ultrasonography with colour Doppler, who went on to have histopathological analysis of the adnexal mass.

### Participants:

12 studies included (2398 women): 7 were prospective studies, all were non-blinded. Prevalence of malignant tumours was 20% and borderline tumours 3%. Literature search included studies published between 1990 and 2007.

### Study Design:

Systematic review and meta-analysis.

### Target Condition:

The target condition was the identification of malignancy in adnexal masses. The reference standard was histopathology in all cases.

### Tests:

**Transvaginal colour Doppler ultrasound** (resistance index of 0.5 or less) for malignant/borderline tumours versus benign tumours.

Pooled sensitivity was 84% (95% CI 84% to 90%)

Pooled specificity was 89% (95% CI 84% to 90%)

### Follow Up:

Not applicable.

### Notes:

Uncertain US results were excluded from the analysis (would inflate the estimates of diagnostic accuracy). The setting of each study is not reported (primary, secondary or tertiary care).

### Author(s): Medeiros et al., 2009b

### Settings:

Women with clinically suspected adnexal mass, whose CA 125 levels were measured and who went on to have histopathological analysis of the adnexal mass.

### Participants:

17 primary studies were included, with a total of 2374 women. The prevalence of ovarian cancer was 25.5%, and prevalence of borderline tumours was 3%. Literature search included studies.
### Published between 1985 and 2007.

#### Study Design:

Systematic review.

#### Target Condition:

The target condition was ovarian cancer. Reference standard diagnosis was histopathology.

#### Tests:

**Serum CA 125 level**, >35 U/ml cut-off value for malignancy

Pooled sensitivity was 80% (95%CI 76% to 82%) for the detection of malignant/borderline tumours versus benign tumours

Pooled sensitivity was 75% (95%CI 73% to 77%) for the detection of malignant/borderline tumours versus benign tumours

#### Follow Up:

Not applicable.

#### Notes:

Results were not analysed according to menopausal status

---

### Author(s): Menon et al., 2009

#### Settings:

Post menopausal women aged 50-74 years, who were not at high risk of ovarian cancer.

#### Participants:

202638 women.

#### Study Design:

Randomised trial of screening strategies. This paper reports the results of the prevalence (initial) screen.

#### Target Condition:

The target condition was ovarian cancer. The reference standard was histopathology in women who had surgery or clinical/radiological follow up in others.

#### Tests:
Women were randomised to no screening, annual CA-125 screening with transvaginal ultrasound scan as a second test (multimodal screening) or annual screening with transvaginal ultrasound. If initial tests (called level 1 screens) suggested intermediate or elevated risk of ovarian cancer women went for a level 2 screening test - an ultrasound scan done by an experienced gynaecologist, radiologist or senior sonographer with particular expertise in gynaecological scanning. Women with abnormal level scans were referred for clinical assessment.

**Diagnostic accuracy of multimodal screening for detection of primary epithelial and tubal cancers**

Sensitivity 89%, specificity not reported; for invasive cancers (within 1 year of screen) sensitivity 89.5%, specificity 99.9%

**Diagnostic accuracy of ultrasound screening for detection of primary epithelial and tubal cancers**

Sensitivity 85%, specificity not reported; for invasive cancers (within 1 year of screen) sensitivity 75.0%, specificity 98.3%

---

**Author(s):** Myers et al., 2006

**Settings:**

Four clinical settings: patients with suspected adnexal masses, patients with adnexal masses, patients with suspected benign adnexal masses and patients with suspected malignant adnexal masses.

**Participants:**

14 studies examined pelvic examination, 153 studies ultrasound, Almost all studies were case series, although 13 population based screening studies were also included.

**Study Design:**

Systematic review

**Target Condition:**

Target condition was detection of adnexal mass, discrimination of malignant from benign adnexal masses,

**Tests:**

Bimanual pelvic examination, ultrasound morphology (Sassone.DePriest, Ferrazzi, Finkler or other scoring systems), ultrasound Doppler (resistance index, pulsatility index and maximum systolic velocity), combined morphology and Doppler, MRI, CT, FDG-PET, serum tumour markers (CA-125)
Settings:
Women presenting with a pelvic mass to gynaecology department. Inclusion criteria: age 45 or older, amenorrhoeic for at least 1 year, scheduled for surgical exploration with biopsy and/or excision of pelvic mass.

Participants:
228 women. 95 malignant tumours were found (41%) and 6 borderline tumours (2.6%).
199 of the pelvic masses were initially identified by pelvic examination and 28 by ultrasound.

Study Design:
Prospective multi centre case series.

Target Condition:
Target condition was the prediction of malignancy in pelvic masses. The reference standard was histopathology.

Tests:

Pelvic examination (PE) done by gynaecologist (clinical impression of malignant disease or not)
Sensitivity 93% (85 to 97%), specificity 63% (55 to 71%)

Transvaginal ultrasound (US) (Finkler score of 7-10 was the criteria for malignancy)
Sensitivity 88% (80 to 95%), specificity 64% (56 to 72%)

CA-125 level (>35 U/ml was the threshold for malignancy)
Sensitivity 72% (61 to 81%), specificity 80% (73 to 87%)

CA 125, US and PE all positive
Sensitivity 62% (51% to 92 72%), specificity 92% (87% to 97%)

PE and US positive
Sensitivity 83% (74% to 91%), specificity 79% (72% to 86%)

CA 125 and PE positive
Sensitivity 67% (56% to 77%), specificity 90% (85% to 95%)

US and CA 125 positive
Sensitivity 64% (53% to 74%), specificity 89% (84% to 94%)

The diagnostic accuracy of other combinations of the test results were reported

Follow Up:
<table>
<thead>
<tr>
<th>Notes:</th>
</tr>
</thead>
<tbody>
<tr>
<td>High prevalence of malignancy, patients had already had pelvic exam / ultrasound before entry into the study.</td>
</tr>
</tbody>
</table>