Progress and Challenges in Screening for Early Detection of Ovarian Cancer*

Ian J. Jacobs‡ and Usha Menon

Ovarian cancer is characterized by few early symptoms, presentation at an advanced stage, and poor survival. As a result, it is the most frequent cause of death from gynecological cancer. During the last decade, a research effort has been directed toward improving outcomes for ovarian cancer by screening for preclinical, early stage disease using both imaging techniques and serum markers. Numerous biomarkers have shown potential in samples from clinically diagnosed ovarian cancer patients, but few have been thoroughly assessed in preclinical disease and screening. The most thoroughly investigated biomarker in ovarian cancer screening is CA125. Prospective studies have demonstrated that both CA125 and transvaginal ultrasound can detect a significant proportion of preclinical ovarian cancers, and refinements in interpretation of results have improved sensitivity and reduced the false-positive rate of screening. There is preliminary evidence that screening can improve survival, but the impact of screening on mortality from ovarian cancer is still unclear. Prospective studies of screening are in progress in both the general population and high-risk population, including the United Kingdom Collaborative Trial of Ovarian Cancer Screening (UKCTOCS), a randomized trial involving 200,000 postmenopausal women designed to document the impact of screening on mortality. Recent advances in technology for the study of the serum proteome offer exciting opportunities for the identification of novel biomarkers or patterns of markers that will have greater sensitivity and lead time for preclinical disease than CA125. Considerable interest and controversy has been generated by initial results utilizing surface-enhanced laser desorption/ionization (SELDI) in ovarian cancer. There are challenging issues related to the design of studies to evaluate SELDI and other proteomic technology, as well as the reproducibility, sensitivity, and specificity of this new technology. Large serum banks such as that assembled in UKCTOCS, which contain preclinical samples from patients who later developed ovarian cancer and other disorders, provide a unique resource for carefully designed studies of proteomic technology. There is a sound basis for optimism that further developments in serum proteomic analysis will provide powerful methods for screening in ovarian cancer and many other diseases. Molecular & Cellular Proteomics 3:355–366, 2004.

THE RATIONALE OF OVARIAN CANCER SCREENING

Ovarian cancer is the most frequent cause of death from gynecological cancer and the fourth most frequent cause of death from cancer in women in Europe and the United States. Most ovarian cancers occur after menopause when the ovaries have no physiological role and consequently abnormal ovarian function causes no symptoms. As a result of this factor, combined with the anatomical location of the ovaries deep in the pelvis, ovarian cancers typically cause few symptoms until they reach a large size or have disseminated. As a result, ovarian cancer is usually diagnosed at an advanced stage when despite advances in surgical and chemotherapeutic management during the last decade survival rates are poor. Almost 90% of patients are diagnosed with metastatic disease in the pelvis or abdomen and for these patients 5-year survival rates are less than 30%. In contrast, the small proportion of patients diagnosed with stage I ovarian cancer confined to the ovaries have a 5-year survival rate in excess of 90%.

A premalignant precursor lesion for ovarian cancer has not been identified, limiting the focus of screening at present to detection of asymptomatic, early stage disease (1). The relationship between stage at presentation and survival in ovarian cancer has long provided a rationale for efforts to improve outcome by detection of early stage disease. Ovarian cancer satisfies many of the World Health Organization criteria (2) for population screening. However, it remains uncertain whether the currently available screening tests can detect ovarian cancer sufficiently early to allow intervention to alter the natural history of the disease. A major effort has been made during the last 20 years to evaluate the tumor marker CA125 and ultrasound scanning in screening for ovarian cancer. Considerable success has been achieved in refining these tests, and large prospective trials are currently in progress to assess the impact of general population screening. Recent progress in serum proteomic analysis has generated much interest in the prospect of novel and sensitive combinations of serum protein markers.

THE CHALLENGE OF OVARIAN CANCER SCREENING

The consequence of a positive screening test for ovarian cancer is surgical intervention of some kind (either laparoscopy or laparotomy). Although ovarian cancer is an important cause of mortality, it is still a relatively uncommon disease, with an incidence no greater than 40 per 100,000 per year even in the postmenopausal population. There is therefore a
considerable danger that the morbidity (and potentially mortality) associated with complications of surgery for false-positive screening results will outweigh the benefits of early detection in women with true-positive results. Most clinical researchers in this field agree that in order to be acceptable in this context an ovarian cancer screening strategy must achieve a minimum positive predictive value (PPV)\(^1\) of 10% (i.e., no more than nine false positives for each true positive). In order to achieve this 10% PPV target on screening the general population of postmenopausal women with an incidence of 40/100,000/year, a screening test for ovarian cancer will need to achieve a minimum of 99.6% specificity. It follows that extremely high specificity is a prerequisite of a screening strategy for ovarian cancer. This is a challenging target for any individual marker and explains the need to combine tests (such as CA125 and ultrasound or multiple tumor markers) in a serial fashion.

A second challenge is the identification of a test with adequate sensitivity for ovarian cancer. This requirement causes considerable confusion. Most markers are developed and initially evaluated using samples from patients with clinically diagnosed and often advanced stage cancer. The challenge is to find a test that can detect not only clinically apparent ovarian cancer but also early disease long before it causes symptoms. There are many markers with high sensitivity for clinically diagnosed ovarian cancer, but few of these markers will have high sensitivity for preclinical disease. Reports based upon high sensitivity in samples from patients with clinically diagnosed disease can be highly misleading. The key to altering the outcome for ovarian cancer is intervention early in the natural history of the disease. Consequently, an important parameter is the duration of the marker-positive preclinical phase. This is difficult to study given the lack of suitable samples and the lack of information about the lead time required to successfully alter the natural history of ovarian cancer. Ultimately, the impact on ovarian cancer mortality can only be confirmed in a prospective randomized controlled trial, but before embarking on such large expensive trials evidence is required that potential tests have encouraging sensitivity for preclinical disease.

A third challenge is to define the most appropriate target population for screening. Risk groups for sporadic ovarian cancer are defined by postmenopausal status and age (\(\geq 50\)) and for hereditary ovarian malignancy by family history criteria and presence of BRCA1 and BRCA2 mutations. The majority of ovarian cancers are sporadic and occur in the general population. Over 90% of sporadic cancers occur in women aged over 50, and screening studies in the general population usually target this group. A number of other factors affect risk in the general population including menopause, years of oral contraceptive use, and parity. Hereditary syndromes account for \(~5-10\%) of ovarian cancers. First-degree female relatives of affected members from ovarian or breast and ovarian or hereditary nonpolyposis colon and ovarian cancer families have a lifetime risk of developing ovarian cancer of greater than 10%. Much of this risk is due to mutations arising in the BRCA1 and BRCA2 genes. The average cumulative risks by age 70 years for ovarian cancer is 39% (18–54%) in BRCA1 mutation carriers and 11% (2.4–19%) in BRCA2 mutation carriers (3). In women with strong evidence of a hereditary predisposition, screening from the age of 35 is frequently advocated although the efficacy of such surveillance is not yet established (4).

SCREENING STRATEGIES FOR OVARIAN CANCER SCREENING

Biochemical Markers—In the past 5 years, significant progress has been made in identifying novel tumor markers for early detection of ovarian cancer (Table I). It is important to note that most of these studies use samples from women with clinically diagnosed ovarian cancer as opposed to asymptomatic women with preclinical disease. The most thoroughly assessed ovarian cancer marker is still CA125. CA125 is an antigen expressed by fetal amniotic and coelomic epithelium.

In the adult, it is found in tissue derived from coelomic epithelium (mesothelial cells of the pleura, pericardium, and peritoneum) and Mullerian epithelium (tubal, endometrial, and endocervical). The surface epithelium of normal fetal and adult ovaries do not express CA125, except in inclusion cysts, areas of metaplasia, and papillary excrences (19). Expression has also been identified outside the female genital tract in epithelial cells of the lung, breast, conjunctiva, and glandular epithelium of the prostate gland (20). CA125 was initially detected using a murine monoclonal antibody OC125 raised in response to immunologic challenge with an ovarian cancer cell line (21). Subsequently, at least 26 other antibodies have been described (22). It is now known that the CA125 antigen carries two major antigenic domains classified as A, the domain binding monoclonal antibody OC125, and B, a domain binding monoclonal antibody M11 (22). Current immunoassays for the quantitation of serum CA125 levels are based on a heterologous assay (CA125 II) using both monoclonal antibodies (M11, OC125) in place of the original homologous assay with monoclonal antibody OC125 alone. Molecular analysis of the CA125 antigen has identified a mucin-type glycoprotein that is highly glycosylated with the protein moiety rich in serine, threonine, and proline (23). A serum CA125 of 35 U/ml, initially measured using the homologous assay and representing 1% of healthy female blood donors, is usually accepted as the upper limit of normal (24). This cut-off value is retained by the CA125 II assay (25), which is now preferred because of reduced interassay variation (26). Interest in CA125 as a screening test was initiated by the fact that \(~83\%) of patients with epithelial ovarian cancer had CA125

---

\(^1\) The abbreviations used are: PPV, positive predictive value; PI, pulsatility index; ROC, risk of ovarian cancer; UKCTOCS, United Kingdom Collaborative Trial of Ovarian Cancer Screening; SELDI, surface-enhanced laser desorption/ionization; TOF, time-of-flight.
Progress and Challenges in Screening for Ovarian Cancer

### Tumor markers that may have a role in ovarian carcinoma

<table>
<thead>
<tr>
<th>Tumor marker</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>CA72-4 or TAG 72</td>
<td>Cancer antigen 72 (CA72-4) or tumor-associated glycoprotein 72 (TAG 72) is a glycoprotein surface antigen found in colon, gastric, and ovarian cancer. It is more frequently elevated in mucinous tumors. There are conflicting reports regarding additional sensitivity for detection of ovarian cancer when combined with CA125 compared to CA125 alone (5–7).</td>
</tr>
<tr>
<td>M-CSF</td>
<td>Serum macrophage colony-stimulating factor (M-CSF) is a cytokine produced constitutively by normal as well as neoplastic ovarian epithelium. Levels are elevated in 68% of patients with ovarian cancer compared to 2.5% of apparently, healthy controls (8). Elevated levels have been found in ovarian cancer patients with normal levels of CA125 (8). While CA125 alone was elevated in 67% of 46 patients with stage I ovarian cancer, CA125 or M-CSF was elevated in 91% (9).</td>
</tr>
<tr>
<td>OVX1</td>
<td>Monoclonal antibody OVX1 recognizes an antigenic determinant present in ovarian and breast cancer cells (9). A combination of OVX1, M-CSF, and CA125 can detect a greater fraction of patients with stage I ovarian cancer than CA125 alone, but this is accompanied by an additive effect on false positives (9). However, the OVX1 radioimmunoassay is highly dependent on sample handling (11).</td>
</tr>
<tr>
<td>LPA</td>
<td>Lysophosphatidic acid (LPA) is a bioactive phospholipid with mitogenic and growth factor-like activities that stimulates the proliferation of cancer cells. Plasma LPA may represent a potential biomarker for ovarian and other gynecologic cancers. Elevated plasma LPA levels were detected in 9 of 10 patients with stage I ovarian cancer and all 24 patients with stage II–IV ovarian cancer. In comparison, only 28 of 47 had elevated CA125 levels, including 2 of 9 patients with stage I disease. LPA levels were also elevated in patients with other gynecologic cancers (12).</td>
</tr>
<tr>
<td>Prostasin</td>
<td>Prostasin is a serine protease normally secreted by the prostate gland. It was identified as a biomarker following identification of overexpression of the gene using microarray technology on RNA pooled from ovarian cancer and normal human ovarian surface epithelial cell lines. The combination of CA125 and prostasin in 37 patients with nonmucinous ovarian cancer and 100 control subjects resulted in a sensitivity of 92% (95% CI = 78.1–98.3%) and a specificity of 94% (95% CI = 87.4–97.7%) for detection of ovarian cancer (13).</td>
</tr>
<tr>
<td>Osteopontin</td>
<td>Osteopontin is another biomarker that has been identified by exploiting gene expression profiling techniques. Plasma levels of osteopontin were significantly higher in 51 patients with epithelial ovarian cancer compared with 107 healthy controls, 46 patients with benign ovarian disease, and 47 patients with other gynecologic cancers (14).</td>
</tr>
<tr>
<td>Inhibit</td>
<td>Serum inhibit is an ovarian product that decreases to nondetectable levels after menopause. However, certain ovarian cancers (mucinous carcinomas and sex cord stromal tumors such as granulosa cell tumors) continue to produce inhibit, which provides a basis for a serum diagnostic test. Available data show that inhibit assays that detect all inhibit forms, i.e. assays that detect the α subunit both as the free form and as an αβ subunit dimer provide the highest sensitivity/specificity characteristics as an ovarian cancer diagnostic test (15). Preliminary reports indicate that two kallikreins (hK6 and hK10) may be useful serum biomarkers for diagnosis of ovarian (16–18).</td>
</tr>
</tbody>
</table>

levels $\geq 35$ U/ml (24, 27). Elevated levels were found in 50% of patients with stage I disease and $>90\%$ of women with more advanced stages, though certain tumors (e.g., mucinous and borderline carcinomas) are likely to be associated with lower levels of CA125 than invasive serous cancers (28). In addition, it became apparent that CA125 could be elevated in the preclinical asymptomatic phase of the disease as raised levels were found in 25% of 59 stored serum samples collected 5 years prior to the diagnosis of ovarian cancer (29). In a prospective ovarian cancer screening study of Swedish women, a specificity of 97% and PPV of 4.6% was achieved using CA125 ($\geq 30$ U/ml) in 4,290 volunteers aged 50 years and older (30). The positive-rate ratio of CA125 for ovarian cancer is in part due to the marker being elevated in other cancers (pancreatic, breast, bladder, liver, lung) as well as benign disease (diverticulitis, uterine fibroids, endometriosis benign ovarian cysts, tubo-ovarian abscess, hyperstimulation syndrome, ectopic pregnancy, and physiological conditions (pregnancy and menstruation)) (31).

### Morphological Markers

Real-time ultrasound screening is aimed at detecting the earliest possible architectural changes in the ovary that accompany carcinogenesis. Criteria for assessment include both morphology and ovarian volume with cut-offs for volume ranging from 10 to 20 ml depending on menopausal status (32). The transvaginal route is preferred because of the more detailed images obtained. Persistence of abnormalities on repeat scanning 4–6 weeks following initial detection helps reduce false-positive rates (33). The lack of physiological changes in ovarian volume in postmenopausal women further decreases the number of false positives in this group compared with premenopausal women. However, even in older women there is a high prevalence of benign ovarian lesions. In an ultrasound and histopathological study at autopsy of 52 consecutive postmenopausal women (mean age 79, range 64–96 years), who died from causes other than gynaecological or intraperitoneal cancer, 56% were found to have a $\leq 50$-mm benign adnexal lesion (34). Ultrasonography used in this manner can therefore lead to the detection of...
many “false-positive” benign ovarian tumors, which results in unnecessary surgery in healthy, asymptomatic women. As data regarding outcome accumulates with long-term follow-up of the participants of the early screening trials, it has been possible to further define the risk of ovarian cancer associated with various ultrasound findings. The use of complex ovarian morphology to interpret pelvic ultrasound increases the sensitivity and PPV in multimodal screening (35). Similarly, follow-up of participants in an ultrasound-based screening trial has established that unilocular ovarian cysts <10 cm in diameter are found in 18% of asymptomatic postmenopausal women over 50 years and are associated with an extremely low risk of malignancy. In contrast, complex ovarian cysts with wall abnormalities or solid areas are associated with a significant risk for malignancy (33, 36). To decrease the number of false positives, many screening protocols use a weighted scoring system or morphological index based on ovarian volume, outline, presence of papillary projections, and cyst complexity (i.e. number of locules, wall structure, thickness of septae, and echogenicity of fluid). There is no standardized index as yet with systems varying on the number and type of variables evaluated (37–42). Others use subjective assessment of the grayscale images. Based on gross anatomic changes at the time of surgery, papillary projections have the highest and simple cysts and septal thickness the lowest correlation with a diagnosis of ovarian malignancy (43). Various second-line tests have been explored to reduce the false-positive rate and facilitate discrimination between benign and malignant ovarian lesions. These include CA125 (44), multiple serum tumor markers, computerized tomography and magnetic resonance imaging (45), and newer modalities like three-dimensional ultrasound and three-dimensional power Doppler (46, 47). Self-teaching computer models such as neural networks that may increase the reproducibility of results are being investigated in order to address the problem of the subjectivity of ultrasound (48–50).

Vascular Markers—Neovascularization is an obligate early event in tumor growth and neoplasia (51). Fast growing tumors contain many new vessels that have less smooth muscle in their walls and therefore provide less resistance to blood flow when compared with vessels within benign ovarian tumors. Flow Doppler imaging uses these altered blood flow patterns as markers to differentiate malignant from physiologic and benign lesions. It has been used as a first-line screening test in combination with transvaginal ultrasound (52, 53) as well as a second-line test following an abnormal ultrasound (54, 55) in both general and high-risk population screening. The initial promise of Doppler to differentiate between malignant and benign ovarian masses and therefore improve the specificity of ultrasound (52, 53) has not been sustained (37, 56, 57). While it has been demonstrated that the mean pulsatility index (PI) of vessels supplying ovarian cancers is lower than that of vessels supplying benign ovarian tumors, the overlap in vascular resistance between these two groups prevents reliable separation of malignant from benign ovarian tumors. The optimal parameters and cut-off levels (PI <1.0, resistance index < 0.4 or 0.6, or peak flow velocity) with the highest predictive value for malignancy have been difficult to define. It was reported that lack of blood flow in an ovarian tumor as detected by color Doppler may preclude cancer (58). This was not substantiated in data from the Kentucky screening trial where 6% of ovarian tumors without blood flow were malignant (37). Even when Doppler examinations were simplified and limited to the expression of internal color flow, grayscale sonography was a more sensitive indicator of malignancy than Doppler sonography (57). Recently, some studies have shown that three-dimensional power Doppler examinations may be more accurate than two-dimensional Doppler examinations (46). However, other experts dispute this (59).

Combining or Repeating Tests—It has been suggested that the use of multiple markers may increase sensitivity for early detection of ovarian cancer. However, increased sensitivity is usually associated with decreased specificity. A panel of eight different markers (CA125, M-CSF, OVX1, LASA, CA15-3, CA72-4, CA19-9, CA54/61) improved the sensitivity for discriminating malignant from benign pelvic masses (60). Using the same dataset, a subset of four markers analyzed using an artificial neural network demonstrated improved sensitivity over CA125 alone (87.5% versus 68.4%) while maintaining comparable specificity (61). In addition, greater specificity using multiple markers might be attained if serial values were employed as in the case of CA125. Preliminary data on a panel of five serum tumor markers (CA125, HER-2/neu, urinary gonadotropin peptide, lipid-associated sialic acid, and Dianon marker 70/K) obtained during 6 years of follow-up of 1,257 healthy women at high risk of ovarian cancer suggests that individual-specific screening rules may be developed with the potential to improve early detection of ovarian cancer (62).

Important progress has been made by combining CA125 with ultrasound and by repeating the CA125 test over time to establish a pattern. The specificity of screening with CA125 was initially improved by the use of pelvic ultrasound as a second-line test to assess ovarian volume and morphology. Using multimodal screening incorporating sequential CA125 and pelvic ultrasound, a specificity of 99.9% and PPV of 26.8% (~4 operations for each cancer) for detection of ovarian and fallopian tube cancer was achieved in 22,000 postmenopausal women (63, 64). With the accumulation of data in this study, ovarian morphology has been used to refine algorithms for the interpretation of ultrasound in postmenopausal women with elevated CA125 levels (65, 66). In a randomized control trial of ovarian cancer screening using a screening strategy incorporating sequential CA125 and transvaginal ultrasound, median survival was significantly increased in women with ovarian cancer in the screened group (72.9 months) when compared with the control group (41.8 months) (66). The same study dataset was used to develop a more so-
Cancer Genetics Network in the United States.

screening trial in being evaluated prospectively in a pilot ovarian cancer (UKCTOCS; www.ukctocs.org.uk). The ROC algorithm is also part of the multimodal screening strategy in the recently forms part of the multimodal screening strategy in the recently completed pilot randomized control trial of ovarian cancer screening at Bart’s London and is part of the ongoing United Kingdom Collaborative Trial of Ovarian Cancer Screening (UKCTOCS; www.ukctocs.org.uk). The ROC algorithm is also being evaluated prospectively in a pilot ovarian cancer screening trial in “high-risk” women under the auspices of the Cancer Genetics Network in the United States.

OVARIAN CANCER SCREENING TRIALS IN PROGRESS

Two distinct screening strategies have emerged, one ultrasound based and the other based on measurement of the serum tumor marker CA125 with ultrasound as the secondary test (multimodal screening). Overall, the data from large prospective studies of screening for ovarian cancer in the general population (Table III) suggests that sequential multimodal screening has superior specificity and PPV compared with strategies based on transvaginal ultrasound alone. However, ultrasound as a first-line test may offer greater sensitivity for early stage disease.

Trials in the General Population—Randomized controlled trials are now underway in the general population to assess the impact of screening on ovarian cancer mortality. The UKCTOCS has recruited over 120,000 postmenopausal women from 13 centers in the United Kingdom. A total of 200,000 women in all will be randomized to either control, screening with ultrasound, or multimodal screening. The primary endpoint is impact of screening on ovarian cancer mortality. The study also addresses the issues of target population, compliance, health economics, and physical and psychological morbidity of screening. Results are expected in 10 years (www.ukctocs.org.uk). The Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial has completed enrolling 74,000 women aged 55–74 at 10 screening centers in the United States with balanced randomization to intervention and control arms. For ovarian cancer, women are screened using both CA125 and transvaginal ultrasound for 3 years and CA125 alone for a further 2 years. Follow-up will continue for at least 13 years from randomization to assess health status and cause of death (78).

Trials in the High-Risk Population—Screening this population can be problematic as they are mainly premenopausal women who have a variety of both physiological (e.g., menstrual cycle variations) and benign conditions (e.g., endometriosis, ovarian cysts) that can give rise to false-positive abnormalities on ultrasound and CA125. Hence criteria for interpretation of the screening tests need to be different from that developed for postmenopausal women in the general population. To date, nine prospective studies have reported on screening for familial ovarian cancer (Table IV). Over 5,000 women have been screened and 33 primary invasive epithelial ovarian and peritoneal cancers detected using mainly ultrasound and CA125 as first-line tests. Criteria for interpreting the test results vary, and screening protocols are not always clearly reported. Only three of the studies have reported interval cancers, which presented between 2 and 24 months following the last screen (79, 86, 88). Multifocal peritoneal serous papillary carcinoma may be a phenotypic variant of familial ovarian cancer and screening strategies using ultrasonography and CA125 testing are not reliable in detecting this disease (86, 90). The other option for these women at high risk is risk-reducing salpingo-oophorectomy after completion of their families (91, 92). In order to develop an optimal screening strategy in the high-risk population, a multicenter National Familial Ovarian Cancer Screening Study (UKFOCSS) has started recruiting “high-risk” women in the United Kingdom. This is a prospective study based on annual screening with CA125 measurement and transvaginal ultrasound. The trial design includes collecting and storing serial serum samples every 4 months for retrospective analysis of CA125 and other markers (93). A similar trial is underway in the United States under the auspices of the Cancer Genetics Network of the National Cancer Institute with the scope for meta-analysis in the future. In the U.S. trial, screening is
### Table III
Prospective ovarian cancer screening studies in the general population

<table>
<thead>
<tr>
<th>Study</th>
<th>Main features</th>
<th>Screening strategy</th>
<th>No. screened</th>
<th>No. of invasive epithelial ovarian cancers detected</th>
<th>No. of positive operations/cancer detected</th>
</tr>
</thead>
<tbody>
<tr>
<td>CA125 alone</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Einhorn et al. 1992</td>
<td>Age ≥40 years</td>
<td>Serum CA125</td>
<td>5,550</td>
<td>62 stage I</td>
<td>175&lt;sup&gt;c&lt;/sup&gt;</td>
</tr>
<tr>
<td>Jacobs et al. 1993, 1996</td>
<td>Multimodal approach: CA125 (level I screen), then USS (level II screen)</td>
<td>Serum CA125 TAS, if CA125 ↑</td>
<td>22,000</td>
<td>114 stage I</td>
<td>41</td>
</tr>
<tr>
<td>Jacobs et al. 1999</td>
<td>Age ≥45 years (median 56) postmenopausal</td>
<td>RCT Serum CA125 TAS/TVS, if CA125 ↑</td>
<td>10,958</td>
<td>63 stage I</td>
<td>29</td>
</tr>
<tr>
<td>Grover et al. 1995</td>
<td>Age ≥40 years (median 51) or with family history (3%)</td>
<td>Serum CA125 TAS/TVS, if CA125 ↑</td>
<td>2,550</td>
<td>10 stage I</td>
<td>16</td>
</tr>
<tr>
<td>Adonakis et al. 1996</td>
<td>Age ≥45 years (mean 58)</td>
<td>Serum CA125 TVS, if CA125 ↑</td>
<td>2,000</td>
<td>1 (1) stage I</td>
<td>15</td>
</tr>
<tr>
<td>USS only approach: USS (level I screen), then repeat USS (level II screen)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>De Priest et al. 1997</td>
<td>Age ≥50 years and postmenopausal or ≥30 with family history</td>
<td>TVS annual screens; mean 4 screens/woman</td>
<td>14,469</td>
<td>11 (6) stage I</td>
<td>180</td>
</tr>
<tr>
<td>van Nagell et al. 2000</td>
<td></td>
<td>TVS TVS + markers at level II</td>
<td>51,550</td>
<td>2217 stage I</td>
<td>324</td>
</tr>
<tr>
<td>Sato et al. 2000</td>
<td></td>
<td>TVS TVS markers at level II</td>
<td>51,550</td>
<td>2217 stage I</td>
<td>324</td>
</tr>
<tr>
<td>Hayashi et al. 1999</td>
<td>Age ≥50 years</td>
<td>TVS</td>
<td>23,451</td>
<td>3 (3)</td>
<td>258&lt;sup&gt;ad&lt;/sup&gt;</td>
</tr>
<tr>
<td>Tabor et al. 1994</td>
<td>Aged 46–65 years</td>
<td>TVS</td>
<td>435</td>
<td>0</td>
<td>9</td>
</tr>
<tr>
<td>Campbell et al. 1989</td>
<td>Age ≥45 years (mean 53) or with family history (4%)</td>
<td>TVS 3 screens at 18 monthly intervals</td>
<td>5,479</td>
<td>2 (3) stage I</td>
<td>326</td>
</tr>
<tr>
<td>Minto et al. 1989</td>
<td>Age ≥45 years or postmenopausal (mean 54)</td>
<td>US (not specified)</td>
<td>500</td>
<td>0</td>
<td>11</td>
</tr>
<tr>
<td>Goswamy et al. 1983</td>
<td>Age 39–78 postmenopausal</td>
<td>TVS</td>
<td>1,084</td>
<td>11 stage I</td>
<td></td>
</tr>
<tr>
<td>US and CDI (level I screen)</td>
<td></td>
<td>TVS and CDI</td>
<td>5,013</td>
<td>44 stage I</td>
<td>38</td>
</tr>
<tr>
<td>Kurjak et al. 1995</td>
<td>Aged 40–71 years (mean 45)</td>
<td>TVS and CDI</td>
<td>1,364</td>
<td>(1)</td>
<td>5</td>
</tr>
<tr>
<td>Vuento et al. 1995</td>
<td>Aged 56–61 years (mean 59)</td>
<td>TVS and CDI</td>
<td>1,364</td>
<td>(1)</td>
<td>5</td>
</tr>
<tr>
<td>USS (level I) and other test (level II screen)</td>
<td></td>
<td>TVS then CDI if TVS positive</td>
<td>2,953</td>
<td>11 stage I</td>
<td>15&lt;sup&gt;e&lt;/sup&gt;</td>
</tr>
<tr>
<td>Parkes et al. 1994</td>
<td>Aged 50–64</td>
<td>TVS then CDI if TVS positive</td>
<td>478</td>
<td>11 stage I</td>
<td>33&lt;sup&gt;f&lt;/sup&gt;</td>
</tr>
<tr>
<td>Holbert 1994</td>
<td>Aged 30–89 years postmenopausal</td>
<td>TVS then CA125 if TVS positive</td>
<td>478</td>
<td>11 stage I</td>
<td>33&lt;sup&gt;f&lt;/sup&gt;</td>
</tr>
</tbody>
</table>

<sup>a</sup> RCT, randomized controlled trial; TAS, transabdominal ultrasound; TVS, transvaginal ultrasound.

<sup>b</sup> Primary invasive epithelial ovarian cancers. The borderline/granulosa tumors detected are shown in parentheses.

<sup>c</sup> Not all of these women underwent surgical investigation as the study design involved intensive surveillance rather than surgical intervention.

<sup>d</sup> Only 95 women consented to surgery, and there are no follow-up details on the remaining.

<sup>e</sup> 86 women had abnormal USS prior to CDI.

<sup>f</sup> Only 11 of these women underwent surgery.
<table>
<thead>
<tr>
<th>Study</th>
<th>Population*</th>
<th>Screening protocol</th>
<th>No. screened (premenopausal %)</th>
<th>No. referred for diagnostic tests (%)b</th>
<th>No. of invasive EOC detected (borderline tumors)c</th>
<th>Cancers in screen-negative women</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bourne et al. 1994 (79)</td>
<td>Aged &gt;17 (mean 47); F/H of Ov cancer</td>
<td>TVS then CDI</td>
<td>1,602 (60%)</td>
<td>62 (3.8)</td>
<td>4 (3) stage I</td>
<td>2 PP (2–8 months) 4 EOC (24–44 months)</td>
</tr>
<tr>
<td>Weiner et al. 1993 (80)</td>
<td>P/H Br cancer</td>
<td>TVS and CDI</td>
<td>600</td>
<td>12 (3)</td>
<td>31 stage I</td>
<td>Not stated</td>
</tr>
<tr>
<td>Muto et al. 1993 (81)</td>
<td>Aged &gt;25; F/H of Ov cancer</td>
<td>TVS and CA125</td>
<td>384 (85.4)</td>
<td>15 (3.9)</td>
<td>0</td>
<td>Not stated</td>
</tr>
<tr>
<td>Schwartz et al. 1995 (82)</td>
<td>Aged &gt;30; F/H of Ov cancer</td>
<td>TVS and CDI and CA125</td>
<td>247</td>
<td>1 (0.4)</td>
<td>0</td>
<td>Not stated</td>
</tr>
<tr>
<td>Belinson et al. 1995 (83)</td>
<td>Aged &gt;23 (mean 43); F/H of Ov cancer</td>
<td>TVS and CDI and CA125</td>
<td>137</td>
<td>2 (1.5)</td>
<td>1</td>
<td>Not stated</td>
</tr>
<tr>
<td>Menkiszak et al. 1998 (84)</td>
<td>Aged &gt;20; F/H of Br/Ov cancer</td>
<td>TVS and CA125 (6 monthly)</td>
<td>124</td>
<td>Not available</td>
<td>1 (3)</td>
<td>Not available</td>
</tr>
<tr>
<td>Karlan et al. 1993 (85)</td>
<td>Aged &gt;35; F/H of Ov, Br, Endo, colon cancer; P/H Br cancer</td>
<td>TVS and CDI and CA125 (6 monthly till 1995 then annually)</td>
<td>597d (75)</td>
<td>10 (1.7)</td>
<td>0 (1)</td>
<td>Not stated</td>
</tr>
<tr>
<td>Karlan et al. 1999 (86)</td>
<td>Aged &gt;35; F/H of Ov, Br, Endo, colon cancer; P/H Br cancer</td>
<td>TVS and CDI and CA125 (6 monthly till 1995 then annually)</td>
<td>1,261</td>
<td>Not stated</td>
<td>1 EOC, 3 PP (2) 1 stage I</td>
<td>4 PP (5, 6, 15, 16 months)</td>
</tr>
<tr>
<td>Dorum et al. 1996 (87)</td>
<td>Aged &gt;25 (mean 43); strict criteria for F/H of Br/Ov cancer</td>
<td>TVS and CA125</td>
<td>180d</td>
<td>16 (8.9)</td>
<td>4 (3)d</td>
<td>2e</td>
</tr>
<tr>
<td>Dorum et al. 1999 (88)</td>
<td>Aged &gt;25 (mean 43); strict criteria for F/H of Br/Ov cancer</td>
<td>TVS and CA125</td>
<td>803</td>
<td>Not stated</td>
<td>16 (4)</td>
<td>Not stated</td>
</tr>
<tr>
<td>Scheuer et al. 2002 (89)</td>
<td>Aged &gt;35; BRCA1 /2 mutation carriers</td>
<td>TVS and CA125 (6 monthly)</td>
<td>62</td>
<td>22 (35.5)</td>
<td>10 had surgery 54 EOC, 1 PP3 stage I</td>
<td>0f</td>
</tr>
</tbody>
</table>

* P/H, personal history; F/H, family history; Ov, ovarian; Br, breast; Endo, endometrial.

b Following positive secondary screens.
c EOC, epithelial ovarian cancer; PP, primary peritoneal cancer.
d Not included in total as there are more recent updates on the trial.
e A further 13 women underwent oophorectomy for breast cancer; two had ovarian cancer not detected by TVS.
f Two women who opted for oophorectomy with normal scans and CA125 had stage 1 ovarian cancer.
based on 3-monthly serum CA125 measurement interpreted using the ROC algorithm. Women with elevated ROC values are triaged to ultrasound. Over 2,000 high-risk women have been recruited onto this study to date with the scope for meta-analysis in the future.

PROTEOMIC TECHNOLOGY AS A POTENTIAL WAY FORWARD IN OVARIAN CANCER SCREENING

There is broad agreement that analysis of the human serum proteome has great potential for diagnosis and early detection of human disease. The challenges are immense given the complexity of the human proteome and the broad dynamic range in abundance of individual proteins. The key to unlocking this potential is the development of reproducible, sensitive, and specific technology for proteomic analysis. Recent advances in technology suggest that this may now be feasible. A wide range of techniques are now available for protein identification and characterization with high throughput. Surface-enhanced laser desorption/ionization time-of-flight (SELDI-TOF) analysis and matrix-associated laser desorption/ionization time-of-flight (MALDI-TOF) technology have the potential to identify patterns or changes in thousands of small proteins (<20 kDa). When combined with matrices that selectively absorb certain serum proteins, these approaches can globally analyze almost all small proteins in complex solutions, such as serum or plasma. Several groups have demonstrated the potential power of proteomic techniques for the discovery of novel disease biomarkers from complex body fluids (94–98). In the last 2 years, studies have investigated the use of mass spectrometry-based methods for early cancer detection from human serum through so-called "proteomic pattern diagnostics" combining the mass spectra generated by these new technologies with artificial-intelligence-based informatic algorithms (98–102). An important though controversial preliminary study reported that using SELDI-TOF to analyze the proteomic spectra patterns generated from 50 women with and 50 women without ovarian cancer, the algorithm identified a cluster pattern that, in the training set, completely segregated cancer from noncancer. The discriminatory pattern correctly identified all 50 ovarian cancer cases in the masked set including all 18 stage I cases. Of the 66 cases of nonmalignant disease, 63 were recognized as not cancer. This result yielded a sensitivity of 100%, specificity of 95% (87–99), and PPV of 94% (101). Although the limitations of this study design and its analysis have been discussed in some detail in the literature (103–106), the potential implications of such proteomic spectrum analysis for the identification of novel tumor markers is huge. It is possible that in the future, the early detection of ovarian cancer (and other cancers) will involve high-throughput proteomic profiling either alone or in combination with markers already in use.

The challenges in translating the potential of serum proteomic technology into a robust approach to clinical ovarian cancer screening are numerous. First, there are uncertainties about whether or not it will be possible to identify subtle changes caused by an early stage ovarian cancer in a low-abundance protein or protein fragment. Diamandis (107) argues that SELDI surfaces preferentially bind high-abundance proteins and that the technology will not be able to identify small alterations in the proteome caused by an early cancer. If so, it may prove difficult to achieve adequate sensitivity for early-stage preclinical ovarian cancer using current technology. To date, only samples from patients with clinically diagnosed disease have been analyzed, and the need for high sensitivity will be even greater in preclinical disease. Second, achieving a high degree of analytical reproducibility is also an area of concern. This relates to issues such as the reproducibility of the presence or absence of peaks and the quantification of the height of peaks. The currently available technology appears to be difficult to reproduce both within and between laboratories. A particular concern has been highlighted in relation to differences between cancer and control samples at very low m/z ratios <500, which suggests non-biological experimental bias (108). Third, the studies reported to date have reported uncharacterized m/z peaks and have not identified specific proteins. It is assumed that many low molecular mass peaks identified are protein fragments. Although knowledge of the identity of a marker is not a prerequisite for clinical utility, in due course information about the identity of a biomarker will increase confidence in the biological basis for a discriminative pattern. Fourth, there are many potential variables and consequently possible sources of bias related to differences between cases and controls, as well as variations in sample collection, processing, and storage. This is particularly important when utilizing a technique that produces such a large dataset. Statistically significant associations may relate to non-disease-related factors. For example in the Petricoin report (106) there was a 10-year difference in the mean age of the control group and the ovarian cancer group, so the difference observed could have been related to factors such as age or menopausal status. In addition, the control samples but not the cases were from individuals with a family history of cancer. Given the large number of data points yielded by the proteomic spectrum, it will be difficult to guard against all possible biases and hence the importance of validation in entirely independent sample sets.

Several of these points highlight the need for careful selection of samples in future proteomic studies. It is essential that studies utilize samples from carefully matched populations collected and stored under the same conditions. In studies of single markers, it is relatively straightforward to document changes due to factors such as collection and storage variations. This is extremely difficult in proteomic studies that generate large-spectra different components of which will be affected in different ways by these variables. Furthermore, in studies of single markers it is reasonable to start discovery with samples from patients with advanced stage disease on the general assumption that a marker that is unchanged in
advanced disease is unlikely to be altered earlier in the disease. There is a flaw in using this approach to evaluate proteomic technology in which numerous differences are likely to be identified between clinical cancer and controls. For example, Sorace and Zhan (108) reported 3,591 m/z values whose intensities differed by a p value of 10–6 or less in an analysis of an ovarian cancer dataset downloaded from the Clinical Proteomics Data bank website. Even if these differences are reproducible in advanced disease, most are unlikely to be present in early preclinical disease. Selection of a pattern of differences that distinguish clinically diagnosed cancer and controls may therefore ignore the relatively few but key differences in the early preclinical cancer and normal serum proteome. There is also of course a possibility that key differences will be present in preclinical disease that could be used in screening but are not present in clinical disease. Large serum banks from prospective screening samples are now assembled and will provide an ideal resource for future proteomic studies. For example, the UKCTOCS serum bank will contain 500,000 serum samples from 200,000 postmenopausal women age 50–74 years followed up carefully for an average of 6 years. It will be possible to produce carefully matched case control sets for a variety of disorders that develop in this cohort of women including the common cancers.

In summary, both the technology required for application of serum proteomics to cancer screening and the sample banks for optimal study design now exist. There is a sound basis for optimism that novel and robust approaches to cancer detection and screening will emerge in the near future, but further progress in refining the reproducibility and sensitivity of the technology will be required.

* The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

† To whom correspondence should be addressed: Department of Gynaecological Oncology, Cancer Institute, Bart’s and The London, Queen Mary’s School of Medicine & Dentistry, London EC1M 6GR, United Kingdom. Tel.: 20-7882-5959; Fax: 20-7882-5958; E-mail: i.j.jacobs@qmul.ac.uk.

REFERENCES

Progress and Challenges in Screening for Ovarian Cancer


