

Monitoring ovarian cancer patients during chemotherapy and follow-up with the serum tumor marker CA125

Suher Othman Abu Hassaan

This review has been accepted as a thesis together with three previously published papers by University of Copenhagen 29th of November 2016 and defended on 26th of January 2017.

Toturs: Dorte L. Nielsen, Malgorzata K. Tuxen and György Sölétormos

Official opponents: Claus K. Høgdall, Susanne Malander and Karina Dahl Steffensen

Correspondence: Department of Clinical Biochemistry and Department of Research Nordsjællands Hospital, University of Copenhagen, Dyrehavevej 29, 3400 Hillerød, Denmark.

E-mail: suherothman@hotmail.com

Dan Med J 2018;65(4):B5463

1.0 Introduction

Worldwide, ovarian cancer is the leading cause of death from gynecological cancers [1, 2]. In Denmark, epithelial ovarian/fallopian tube or primary serous peritoneal cancer is the fourth most frequent cause of death in women with approximately 450 new cases diagnosed each year. The incidence of ovarian cancer is 15 pr. 100.000 women, ranking Denmark second among countries with the highest rate of ovarian cancer in the world [3]. Approximately 75% of ovarian cancer patients are detected at advanced stages (Stage IIb, III or IV), when the 5-year survival rate is less than 20%. However, the 5-year survival rate may be as high as 90%, when ovarian cancer is detected early (stage I). Lack of specific symptoms in the early stages, resulting in late diagnosis, is the main reason for the high mortality [4].

Patients with advanced epithelial ovarian cancer (EOC) often have no macroscopic detectable tumor after initial surgery or they present with widespread diffuse peritoneal disease that may be difficult to detect or quantify with traditional methods such as gynecological examinations, transvaginal ultrasound (TVS), computed tomography (CT), and magnetic resonance (MR) scans. There is a need for reliable and easily performed quantitative biochemical tests that reflect tumor burden correctly and provide an early signal of tumor growth given the cost, inconvenience and limited sensitivity of imaging investigations. The serum cancer biomarker Cancer Antigen 125 (CA125) has been proposed as a supplement to non-invasive procedures among patients with advanced disease [5, 6]. However, challenges remain on how to

define increments in CA125 concentrations that allow an optimal interpretation being vital for early diagnosis of tumor growth. This PhD thesis focused on the utility of criteria that have been proposed to detect and interpret increments in serial CA125 concentrations during patients monitoring.

2.0 Background

2.1 Etiology

The majority (90%) of ovarian cancer incidents are sporadic. Epidemiologic and molecular genetic studies have identified several protective and risk factors associated with the development of ovarian cancer (*Table 1*) [7].

Table 1. Protective and risk factors associated with ovarian cancer [7].

Protective factors	Risk factors
Multiple pregnancies	Nulliparity
Lactation	Early menarche
Oral contraceptives	Late menopause
Tubal ligation procedure	Fertility drugs

Approximately 10% of ovarian cancers are hereditary. The most significant risk factors are alterations in *BRCA1* and *BRCA2* tumor suppressor genes accounting for around 40% of hereditary ovarian cancer cases, especially those of the high grade serous type [8]. Mutations in mismatch repair genes associated with Lynch syndrome and mutations in DNA-repair genes and the *TP53* gene account for a few percent. However, the underlying genetic aberrations in about half of the families with a hereditary predisposition remain unknown [9, 10].

2.2 Classification and staging

Ovarian cancer consists of three types; EOC, germ cell, and sex cord stromal tumors. EOC accounts for 90% of the ovarian cancers. The remaining types each account for approximately 5% [7].

Recent evidence based on histopathological research and high throughput genomic and molecular techniques have identified five distinct subtypes of EOC; high-grade serous, low-grade serous, clear cell, endometrioid, and mucinous. Each of these subtypes is associated with distinct underlying genetic aberrations and molecular pathways, expression of biomarkers as well as pattern of metastases, response to chemotherapy, and prognosis [11, 12]. In high-grade serous EOC, *TP53* and *BRCA1/2* mutations are frequently identified and high-grade serous EOC seems to develop de novo within a few months. Whereas, low-grade serous EOC is promoted by *KRAS* and *BRAF* mutations and develops in a stepwise fashion from serous cystadenoma to serous borderline cancer [13]. There is now persuasive evidence to classify these five types of EOC as distinct and different diseases [14, 15]. Substantial evidence also supports the view that EOC/fallopian tube or primary serous peritoneal cancer are closely related clinical entities, sharing histologic, molecular biologic, and genetic features as well as clinical behavior. Accordingly, the International Federation of Gynecology and Obstetrics (FIGO) staging system effective from January 2014, is the same for ovarian/fallopian tube and primary peritoneal carcinoma. FIGO stages are based on findings made mainly through surgical exploration and pathology results [6, 7].

- **2.3 Symptoms and treatment**

Symptoms associated with EOC can be non-specific such as fatigue, weight loss, and vaginal and rectal bleeding, and are often attributable to multiple and common disorders such as irritable bowel syndrome, gastritis and urinary tract infections [1, 16-19]. Overall, approximately 70% of patients with EOC present with Stage III and IV disease. The vast majority of those patients have high-grade serous tumors which are often chemo-sensitive and respond well to initial chemotherapy, but tumor recurrence is frequent and resistance to further therapy develops in nearly all patients over time [1]. In contrast, patients with low-grade serous tumors or tumors of the endometrioid, clear cell or mucinous types often have indolent tumors with localized disease at diagnosis. However, these tumors as a rule show no or only minor responses to chemotherapy and may thereafter progress slowly despite therapy [11, 14, 20]. Current standard treatment for all stages is aggressive cytoreductive surgery. In the case of a significant risk of recurrence (Stage IB-IV), surgery is followed by first-line taxane- and platinum-based chemotherapy [7]. In recent years, platinum-based neoadjuvant therapy followed by interval debulking has increasingly been adopted for disease stages III and IV [1].

- **2.4 Biomarkers in epithelial ovarian cancer**

By definition a biomarker is a gene, protein or process which deviates from the normal either qualitatively or quantitatively and can be identified in body tissues and/or fluids [21]. In order to assess the value of a biomarker in clinical practice principles of an ideal marker has been proposed [22]. The biomarker has to be absent in healthy persons as well as in benign conditions and is exclusively expressed by specific tumor cells in the target malignancy [23]. Among asymptomatic individuals, the biomarker should allow screening for early cancer or premalignant disease and among symptomatic patients the biomarker should help in

the differential diagnosis of benign and malignant disease. Following diagnosis, an ideal biomarker should also be used to assess the prognosis and to predict the most appropriate treatment. For patients receiving systemic therapy the level of expression should correlate with the therapeutic response and tumor burden [24]. A biomarker should contribute to enhance beneficial clinical outcomes such as increasing overall survival (OS), progression free survival (PFS) or reductions in the costs of care [21, 25]. CA125 is currently the most used serological biomarker for the management of patients with EOC/fallopian tube or primary serous peritoneal cancer. Other serological biomarkers to be used among different types of ovarian cancers are listed in *Table 2* [26-35].

Table 2. Histological types of ovarian tumors and associated serological cancer biomarkers [26, 27, 31-34, 36].

Tumor type	Tumor marker
Epithelial ovarian tumor	
High-grade serous	CA125, HE4
Low-grade serous	CA125, HE4
Mucinous	CEA, CA 19-9
Endometrioid	CA125, HE4
Clear cell	CA125
Germ cell tumor	
Dysgerminoma	βHCG, LDH
Endodermal sinustumor	AFP
Immature teratoma	AFP, LDH
Embryonal carcinoma	AFP, LDH
Choriocarcinoma	βHCG, LDH
Mixed type tumor	AFP, βHCG, LDH
Sex stromal cancer	
Granulosa theca cell tumor (adult)	Estrogen, Inhibin, AMH
Granulosa theca cell tumor (juvenile)	Estrogen, Inhibin

CA125: Cancer Antigen 125. CEA: Carcinoembryonic Antigen. CA 19-9: Cancer Antigen 19-9. LDH: Lactate Dehydrogenase. βHCG: Beta Human Chorion Gonadotropin. AFP: Alpha-fetoprotein. AMH: Antimüllerian Hormone.

Cancer Antigen 125

Originally, CA125 was defined by a monoclonal antibody (OC125) generated by immunizing laboratory mice with the OVCA 433 cell line derived from a patient with ovarian serous carcinoma [37]. The most frequently used reference interval quoted for CA125 is ≤ 35 kU/L [38]. By using an immunoradiometric CA125 assay (Centocor, Malvern, PA) Bonfrer et al. reported that the 95th percentile values among healthy women were 36 kU/L (age 40–44 years), 30 kU/L (age 45–55 years), and 25 kU/L (>55 years). With an automated luminescence assay (LIA, Byk-Sangtec, Dietzenbach, Germany) on the same samples, the 95th percentiles were 31, 29, and 21 kU/L, respectively. None of the healthy women >55 years, had a CA125 level >35 kU/L [39]. CA125 concentrations >35 kU/L can occur in healthy premenopausal women, in women with benign gynecologic diseases, women with several benign non-gynecological diseases, and in women with other malignancies (Table 3) [40–42]. Approximately 80% of women with EOC have CA125 concentrations >35 kU/L. The frequency of elevated CA125 concentrations is highest among patients with serous EOC followed by endometrioid and clear cell types [40, 41]. CA125 is poorly expressed in pure mucinous tumors, CEA or CA 19-9 may be more useful markers among these patients (Table 2) [4, 32, 40, 43, 44].

The value of CA125 in clinical practice

In screening, the main challenge with CA125 measurements in asymptomatic women is the lack of sensitivity for early stage serous EOC. Another challenge is lack of specificity, with frequent false positive (FP) signals of disease in terms of elevated CA125 concentrations, especially among premenopausal women (Table 3) [40, 45]. To detect primary EOC at an early stage, the Risk of Ovarian Cancer Algorithm (ROCA) based on serial CA125 measurements was developed by Steven Skates et al. [46, 47].

Table 3. Different conditions causing elevated CA125 concentrations [40–42].

Healthy premenopausal women	Benign gynecologic diseases	Benign non-gynecological diseases	Malignancies
During menses	Ovarian cysts	Peritoneal, inflammatory disorders	Breast cancer
Pregnancy	Endometriosis	Pelvic inflammatory disease	Colorectal cancer
	Adenomyosis	Liver disease	Pancreas cancer
	Uterine leiomyomas	Renal disease	Lung cancer
	Benign ovarian tumors	Musculoskeletal disease	Endometrial cancer
		Cardiac disease	Cervical cancer

The ROCA was investigated in several screening trials i.e. The UK Collaborative Trial of Ovarian Cancer Screening (UKCTOCS) and the Normal Risk Ovarian Screening Study (NROSS). The results from UKCTOCS showed that ROCA may have a role in detecting early ovarian cancer; however, further follow-up is needed before the magnitude of mortality reduction can be determined [48]. CA125 measurements differentiated between benign and malignant ovarian lesions [40, 41], and postmenopausal women with CA125 concentrations >35 kU/L should be considered for referral to a gynecologist [40]. CA125 concentrations >95 kU/L were stated to discriminate malignant from benign pelvic masses with a positive predictive value (PPV) of 95% [4]. The American College of Obstetrics and Gynecologists suggested, that premenopausal women with a pelvic mass and CA125 concentrations >200 kU/L, should be referred to a gynecologist for consultation [49]. For diagnostic purposes an algorithm to calculate the Risk of Malignancy Index (RMI) were developed by Jacobs [50] and by Tingulstad, respectively [51]. Both RMI scoring systems were the product of ultrasound score x menopausal score x CA125 (kU/L) and recommended by the Danish Gynecologic Cancer Group (DGCG) (Table 4) [3].

CA125 predicted the prognosis in women with EOC along with tumor stage, grade, histological type, and size of residual tumor after primary cytoreductive surgery [1, 52]. Among patients with a preoperative CA125 concentration >65 kU/L, the 5-year survival rates were reported to be significantly lower as compared to patients with concentrations <65 kU/L [45, 53–55]. According to guidelines from the European Group on Tumor Markers (EGTM) from 2015, a change in sequential CA125 measurements during primary therapy should be considered as a prognostic indicator for response to treatment [55]. Currently, CA125 is widely used for monitoring of EOC patients [4, 52].

Table 4. The Risk of Malignancy Index.

Feature	RMI 1 scoring system [56]	RMI 2 scoring system [57]
<i>Ultrasound features</i>		
Multilocular cyst	0= no abnormality	0= no abnormality
Solid areas	1= one abnormality	1= one abnormality
Bilateral lesions	1= one abnormality	1= one abnormality
Ascites	3= two or more abnormalities	4= two or more abnormalities
Intra-abdominal metastases		
Premenopausal	1	1
Postmenopausal	3	4
CA125	kU/L	kU/L

RMI score = ultrasound score x menopausal score x CA125 concentration in kU/L.

RMI: Risk of malignancy scoring system. RMI > 200 indicates risk of ovarian malignancy [55].

Standard CA125 monitoring schemes among epithelial ovarian cancer patients

A monitoring scheme starts with a preoperative baseline concentration followed by serial measurements postoperatively, during chemotherapy, and subsequent follow-up periods [4, 5, 52, 55]. A change of tumor size is usually assessed clinically by radiological imaging and measurement of evaluable tumor lesions according to the Response Evaluation Criteria in Solid Tumors (RECIST) [58-60]. Most of EOC patients maintain small amounts of residual tumors at multiple sites throughout the peritoneal cavity (dissemination) after surgery or chemotherapy, thus staging with CT and TVS remains a challenge [5, 61]. There is a need for reliable and easily performed quantitative biochemical tests that may provide an early prediction of tumor growth among patients with diffuse peritoneal carcinomatosis. CA125 has been suggested as a biochemical supplement to other non-invasive procedures because CA125 concentrations may increase in parallel with higher stage of disease as defined by FIGO [5, 6]. Still, a major challenge in monitoring patients with CA125 is to define an increment in concentrations that reliably correlates with recurrence, and progression.

Human Epididymis Protein 4

The molecule is a precursor of the protein human epididymis protein, encoded by the gene located on chromosome 20q12-13.1 [62]. HE4 is frequently over-expressed in ovarian cancer, especially among those with serous and endometrioid histology [63]. Expression of HE4 has also been identified in pulmonary, endometrial and breast carcinomas as well as mesotheliomas, but less frequently in gastrointestinal, renal and transitional cell carcinomas [29]. HE4 was the first biomarker after CA125 to be approved by the Food and Drug Administration for EOC [29, 64-66].

2.5 Criteria to interpret CA125 increments during treatment and follow-up

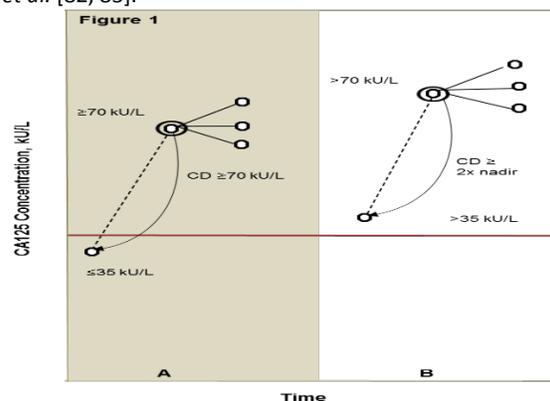
In the last three decades several criteria were proposed to interpret increments in serial concentrations, but none of them have been implemented into routine clinical practice [37, 67-86]. Already the earliest reports on the monitoring performance of CA125 recognized the challenge to develop assessment criteria that separated CA125 increments associated with clinical recurrence and progression from random fluctuations unrelated to tumor growth [37, 67, 68, 70, 73, 75, 77-79, 87]. The most extensive and recent studies evaluating the reliability of assessment criteria were reported by Rustin et al. 1992-2011 (Figure 1) [60, 80-83, 88-92], Tuxen et al. 2001-2002 (Figure 2) [84, 85, 93-95], and by Liu et al. 2007 (Figure 3) [86]. Rustin et al. and Tuxen et al. tested the ability of their criteria to signal tumor growth during first-line chemotherapy and during the subsequent follow-up. Most of the patients investigated by Rustin et al. and all patients investigated by Tuxen et al. were allocated to The North Thames Ovary Trial of 5 versus 8 courses of carboplatin or cisplatin [80-82, 84, 85]. Rustin et al. developed new criteria in each of their studies where some criteria required a defined percent of change from below to above an arbitrarily set cut-off value [80-82, 88]. Other criteria simply required an increment starting above cut-off to higher levels [83]. Tuxen et al. stated that increasing CA125 concentrations were not only due to tumor growth but concen-

trations were also influenced by analytical variation (CV_A) and within-subject background biological variation (CV_I) where CV_I was the random fluctuation around the individual's own hemostatic set-point [95]. Alternative to Rustin et al. and Tuxen et al. the criteria suggested by Liu et al. were generated to interpret CA125 increments starting from baseline concentrations well below cut-off [86].

Design of CA125 assessment criteria

According to Rustin et al. [80-83] an increment between two concentrations should exceed a defined percentage of change before considered indicative of progression. Tuxen et al. [84, 85] used two approaches to generate their criteria. Their first approach was similar to the criteria reported by Rustin et al. Both Rustin et al. [80-83] and Tuxen et al. [84, 85] calculated the percentage change as $((\text{concentration a} - \text{concentration b}) / (\text{concentration a})) \times 100$. Criteria following the second approach suggested by Tuxen et al. involved a more detailed statistical estimate of the significance of an increment. The percentage increase between two concentrations was calculated as $((\text{concentration a} - \text{concentration b}) / ((\text{concentration a} + \text{concentration b}) / 2)) \times 100$. The increase was statistically significant at $p < 0.05$, if the increment exceeded the random variation inherent in the two test results, the Reference Change Value (RCV). The RCV was calculated as $\sqrt{2} \times Z \times \sqrt{(CV_A^2 + CV_I^2)}$ [96], where $\sqrt{2}$ is a constant (two measurements). Z is the Z-statistic; the value is 1.65 if the expected change is unidirectional or 1.96 if the expected change is bidirectional (either an increment or decrement). The CV_A corresponding to the baseline CA125 concentration was read from the precision profile as 8.6%. The average CV_I of CA125 obtained from monitoring studies of healthy individuals and from ovarian cancer patients monitored during steady state of disease was 19.8% and 24%, respectively [93, 94]. Finally, an alternative set of criteria were proposed by Liu et al. in 2007 [86]. The criteria were named Early Signal of Progressive Disease (EPD), and were basically based on a progression criterion suggested by Rustin et al. in 1996 [82], where the upper limit of normal was lowered from 35 kU/L to 10 kU/L.

Figure 1. Criteria to detect CA125 increments generated by Rustin et al. [82, 83].

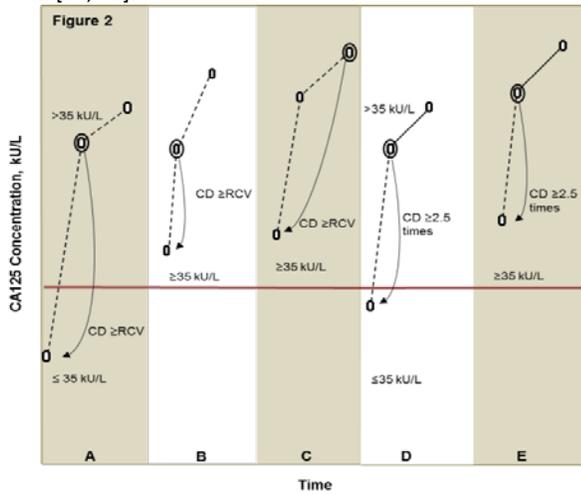


CD: The required critical difference

1A: The criterion corresponds to criterion in figure 6A in study 1 and criterion 2A in study 2 and 3

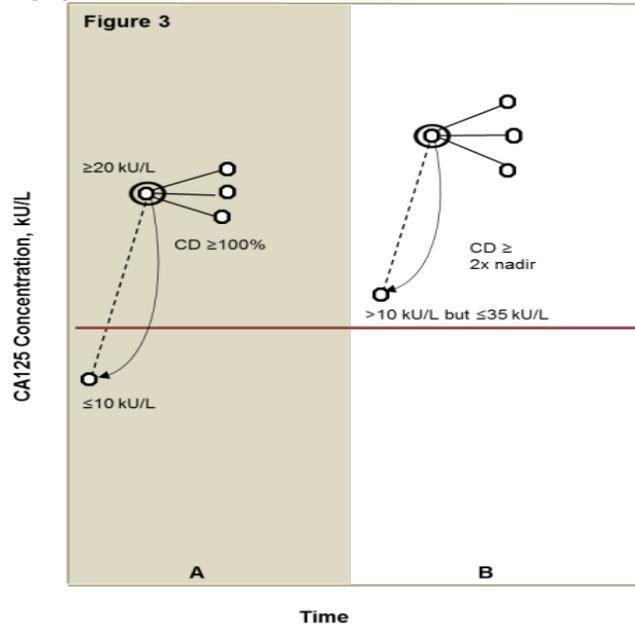
1B: The criterion corresponds to criterion in figure 6B in study 1 and criterion 1A in study 2 and 3

Figure 2. Criteria to detect CA125 increments generated by Tuxen et al. [84, 85].



- CD: The required critical difference
 2A: The criterion corresponds to criterion in figure 7A in study 1 and criterion 2C in study 2 and 3
 2B: The criterion corresponds to criterion in figure 7B in study 1 and criterion 1D in study 2 and 3
 2B: The criterion corresponds to criterion in figure 7C in study 1 and criterion 1C in study 2 and 3
 2D: The criterion corresponds to criterion in figure 7D in study 1 and criterion 2B in study 2 and 3
 2E: The criterion corresponds to criterion in figure 7E in study 1 and criterion 1B in study 2 and 3

Figure 3. Criteria to detect CA125 increments generated by Liu et al. [86].



- CD: The required critical difference
 3A: The criterion corresponds to the criterion in figure 8A in study 1

3B: The criterion corresponds to the criterion in figure 8A in study 1

2.6 Computer simulation model

A computer model that generated simulated biomarker concentrations was previously suggested for comparison and validation of criteria to interpret serial measurements of a given biomarker [97-100]. The model system was developed for monitoring patients with metastatic breast cancer where criteria to interpret serial measurements of Cancer Antigen 15.3 (CA15.3), CEA, and Tissue Polypeptide Antigen (TPA) underwent preclinical evaluation [97-99]. The model system proved useful because assessment criteria could be compared in a variety of simulated conditions i.e. steady-state and progressive disease. It was also used to fine-tune existing assessment criteria and to develop new types of criteria [101, 102].

3.0 Hypothesis and aims

3.1 Hypothesis

An increment in CA125 concentrations is only significant if the change exceeds what can be explained by random variation. Assessment criteria considering the random variation may have a better accuracy and lead-time ability than criteria based on an arbitrary percentage of change.

3.2 Aims

The aim of this PhD project was to investigate the hypothesis through different approaches.

1. Review of the literature to identify criteria to assess CA125 increments during monitoring of patients with EOC (study 1, systematic review).
2. Comparison of the accuracy of the identified CA125 criteria under standardized conditions in a prospective preclinical computer simulation model (study 2, phase I trial).
3. Comparison of the accuracy of the identified CA125 criteria applied to serial CA125 concentrations obtained prospectively from EOC patients during first-line chemotherapy and the subsequent follow-up period (study 3, phase II trial).
4. Comparison of CA125 accuracy obtained in the three different types of studies (studies 1-3).

4.0 Methodological considerations

Details regarding the design are available in the individual articles study 1, study 2, and study 3. Additionally, some issues are addressed further below.

4.1 Study 1

PRISMA

The Medline and Ovid versions of EMBASE were used to search the peer-reviewed literature in English. The clinical trials that tested the ability of CA125 to monitor the growth of EOC during first-line chemotherapy and follow-up were reviewed. The review

was conducted and reported according to Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA guidelines) providing a 27 item checklist and a flow diagram [103]. The review fulfilled all items in the PRISMA statement except items 14, 16 and 21 which refer to meta-analyses [104].

Data source and data collection

The search terms were: “increments” OR “rising CA125 concentrations” OR “monitoring” OR “progression criteria”. These were all combined with “cancer antigen 125” OR “ovarian neoplasms”. Limits were set to human subjects, English language, and articles published between January 1982 and August 2014. Results were downloaded into the reference management software, Reference Manager 12. Excluded were reviews, observational studies and case reports. Included were original articles that met the requirements: women with ovarian cancer and criteria to assess CA125 increments during therapy and/or the follow-up period.

Quality Assessment of Diagnostic Accuracy Studies

Two reviewers (SH and GS) evaluated the identified articles independently and determined whether the report was excluded or included for the systematic review. If disagreement concerning inclusion or exclusion occurred, this was resolved by consensus. The risk of bias regarding the included studies was critically evaluated according to Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) [105]. All of the four key domains of QUADAS-2 regarding patient selection, index test, reference standard, and flow of patients through the study were discussed during the evaluation process.

Statistics

In order to compare the accuracy of each criterion; the number of true positive (TP) signals, false negative (FN) signals, FP signals, and the true negative (TN) signals were extracted from the retrieved papers. Lead-times were also extracted if available. Using a 2 x 2 contingency table the sensitivities, FP-rates, and FN-rates based on the extracted data were calculated/recalculated (Table 5) [106].

Table 5. 2x2 contingency table used to calculate the accuracy of the different criteria.

	Disease positive	Disease negative	
CA125 positive	TP	FP	PPV TP/(TP+FP)
CA125 negative	FN	TN	NPV TN/(TN+FN)
	Sensitivity TP/(TP+FN)	Specificity TN/(TN+FP)	

TP: True positive, FN: False negative, FP: False positive, TN: True negative, PPV: Positive predictive value, NPV: Negative predictive value.

• 4.2 Study 2

The computer simulation model

In study 2 the criteria reported by Rustin et al. and Tuxen et al. in study 1 were applied to simulated data sets corresponding to 1000 surrogate healthy individuals and 4000 surrogate patients each with 50 serial CA125 measurements. The robustness of the criteria against FP signals and their potential to provide early signals of tumor growth were investigated under standardized conditions. By applying the same data sets to all the tested criteria the obtained results were directly comparable and could be used for ranking the criteria in terms of accuracy. The simulation model applied the cut-off level for CA125 of 35 kU/L [67].

The within subject biological variation

We hypothesized that the CV_i among healthy women or the CV_i among ovarian cancer patients were relevant to be considered in the simulations. The CV_i was assumed to be randomly distributed around a homeostatic set point [93]. Because of the assumption that the CV_i could either be Gaussian or In-Gaussian distributed, the simulation study provided the baseline concentrations following both Gaussian and In-Gaussian distributions [102]. Earlier simulation studies in breast cancer showed that increasing values of CV_i were associated with higher numbers of FP marker signals [101]. The applied biological variation in the current study was derived from 20 healthy postmenopausal women and was a combination of the short- and long-term CV_i (19.8%) [93]. Another publication among women with ovarian cancer observed a CV_i of 24% [94]. The CV_i for healthy postmenopausal women were chosen, but it could be argued that the biological variation from ovarian cancer patients might have been a better choice to use in the computer simulation model. However, using 24% instead of 19.8% as CV_i would have little impact on the RCV calculations. *Steady-state*

We assumed that all the simulated patients had baseline concentrations in a steady-state, irrespective of whether their values started above or below the applied cut-off [107]. The applied nadir values may not represent the real steady-state but may represent the extreme high or low steady-state value. This procedure could result in a higher frequency of FP signals for the criteria 2B and 2C, (Figure 2) and moderate increase of FP signals for criterion 2A (Figure 2), because of an additional requirement from below to above the cut-off.

Tumor growth

The rate of CA125 increase among women with progressing EOC was assumed to follow an exponential model. Thus, the CA125 doubling times were transformed into the exponential function lambda (λ) [101]. Han et al. [108] found that the median days required for CA125 to double were 40 days with a range from 2 days to 1675 days. The simulation model applied doubling times of 20 days, 40 days, 80 days, and 160 days in order to illustrate the performance of the criteria under different conditions. The model generated serial CA125 baseline concentrations as well as

increasing concentrations related to tumor growth separately and enabled a combination of the two types of CA125 kinetics.

Statistics

The simulated data were generated with Microsoft Excel version 2003. It was assumed that the CV_i in general followed a Gaussian distribution. However, this assumption may not be correct because some studies suggested that original serial biological data are not always normally distributed, and \ln -transformation was a more relevant approach [102, 109, 110]. Therefore, the simulations generated CA125 concentrations for both situations to investigate whether there were differences in the performance of the investigated criteria among Gaussian and \ln -transformed data, respectively. The numbers of TP increments and the CA125 detection time were obtained among 4 x 1000 surrogate patients (1000 simulations for each doubling time). The number of FP increments was obtained among 1000 simulated healthy individuals with CA125 concentrations in a steady state.

• 4.3 Study 3

Study design

The patients were recruited at two departments of oncology in Denmark during 1995-2001. The design of the study was prospective even though the serum samples and clinical data were collected approximately two decades ago. The design was in accordance with the items outlined by "Standards for Reporting Diagnostic Accuracy Studies" STARD [111] published in 2003, recently updated in 2015 [112]. The study was approved by the relevant committees before the samples were collected, consecutively analyzed for CA125 and frozen, thus, the CA125 data were generated from fresh material. One requirement was not fulfilled according to STARD, because CA125 data, clinical information, and results of the imaging analysis were available to the participating departments, thus the study was not blinded.

Patients

A power calculation of sample size was performed prior to estimation the CA125 results obtained in the phase II study. For a criterion to be valid for detecting CA125 increments it was assumed that the criterion should provide 70 TP signals in terms of progressive disease. By fixating the type 1 error on 0.05 and the type 2 error on 0.10, calculation of power showed that a total of 156 patients should be included for each criterion in order to detect a difference in their performance. A total of 231 patients with newly diagnosed and histologically verified EOC were enrolled to first-line chemotherapy. Patients with other primary malignancies or age <18 years were excluded. During treatment 189 patients were eligible for monitoring and 143 patients were eligible during follow-up (Figure 4). Date of primary surgery, the duration of treatment and follow-up periods were registered for all patients. Evaluation of disease was performed at every third treatment cycle and every three months in the first three years of follow-up, and every 6 months in the last two years. Additional evaluations were performed when needed. The clinical status was recorded in terms of progressive disease, complete response, partial response, and no change according to World Health Organization (WHO) criteria

because these criteria were used at the time the study was conducted [113].

Serum samples for CA125 measurements

Samples were collected prospectively over a six year period from 1995-2001 during first-line chemotherapy and the subsequent follow-up period. The specimens were collected on the days of treatment or on the days of the clinical evaluation, and if possible when routine analytes were requested outside the scheduled time points.

CA125 measurements

Initially, the CA125 concentrations in serum were measured with the ELISA-CA125 II assay from CIS Bio International (Gif-sur-Yvette, France), a solid-phase two-site immune radiometric assay. Following August 1996, the CA125 concentrations were measured by the Immuno 1, a one-step solid phase enzyme immunoassay based on the sandwich principle. The applied cut-off value (the 95th percentile value) for both assays were 35 kU/L as recommended by Bast et al. [37] and the manufacturer.

Quality assurance

To ensure a stable analytical quality throughout the study, three control samples were included in each assay run with different concentrations of CA125. The analytical imprecision comprised both the intra- and inter-assay variation because each sample from an individual subject was analyzed consecutively in different assay runs. The study was conducted according to recently published guidelines by EGTM on how to evaluate biomarkers in phase II trials [24, 114].

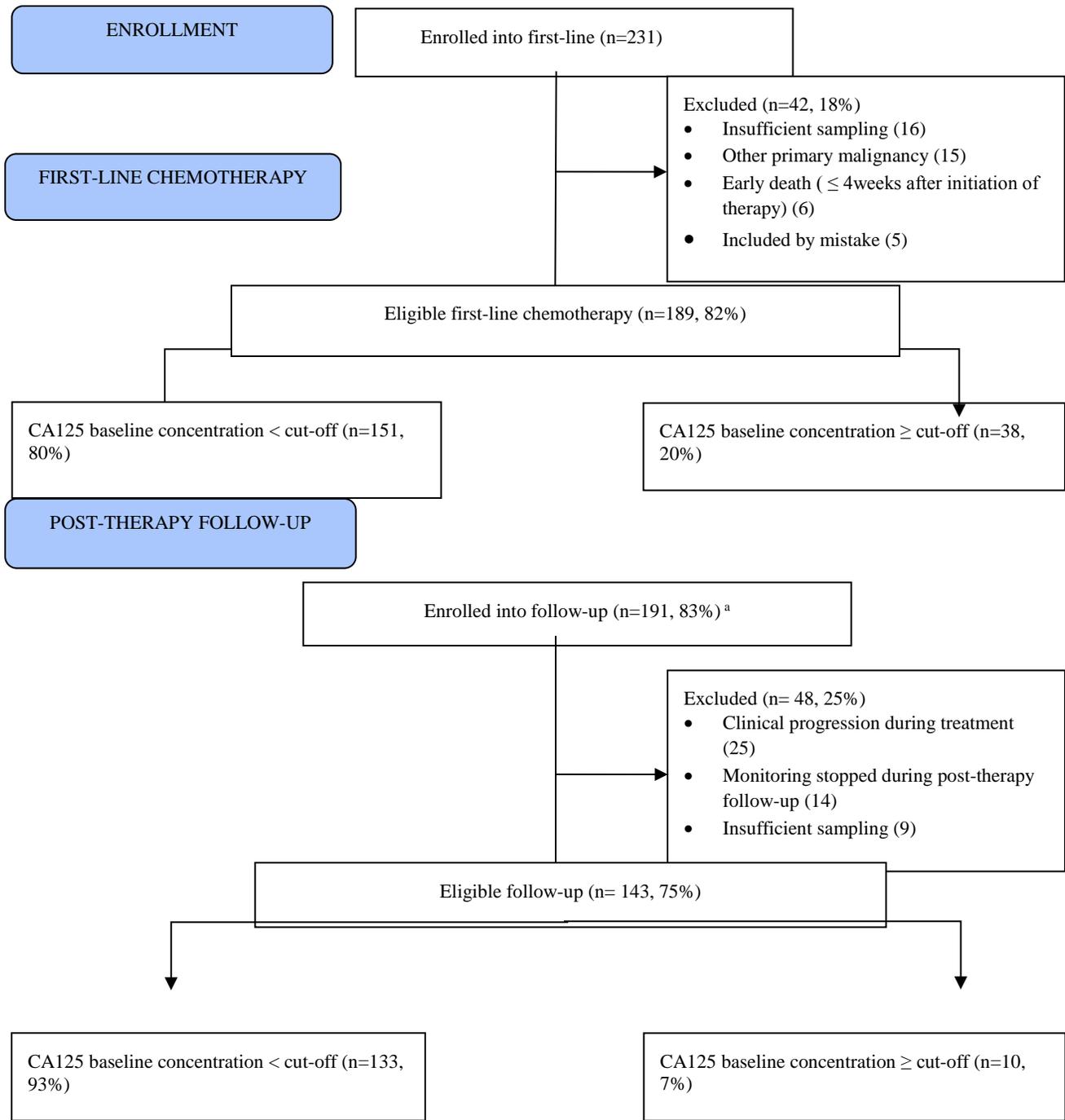
Statistics

According to Linnet et al. a test of biomarker accuracy should discriminate between the absence and presence of a particular disease. In order to compare the accuracy of each criterion the numbers of TP, FN, FP, and TN results were counted [106]. The sensitivities (percentage of patients with tumor growth detected by CA125 increments), the specificities (percentage of patients without new tumor growth confirmed by unchanged CA125 concentrations), the PPV (probability of clinical progression following CA125 progression), the negative predictive values (NPV) (probability of clinical non-progression given CA125 non-progression), the FP-rates (percentage CA125 increments among patients without new tumor growth), and the FN-rates (percentage without CA125 increments among patients with new tumor growth) were calculated using a 2x2 contingency table (Table 5). The 95% confidence intervals (95% CI) were estimated according to the Geigy formulae 771 and 772 [115]. Furthermore, the time interval between a CA125 increment and new tumor growth (lead-time) for each criterion was calculated as described in previous investigations [85, 116].

Ethics

Study III was approved by the Ethical Committee of Region Hovedstaden (KA 94162m) (H-3-2013-FSP43) and the Danish Data Protection Agency (1995-1200-655) (2013-41-2366). An informed consent was obtained from all participants.

Figure 4. Overview of patients with ovarian cancer recruited for CA125 assessment during first-line chemotherapy and follow-up-all histological types.



^a 2 patients excluded from CA125 assessment during first-line chemotherapy due to insufficient sampling (<3 samples), had sufficient sampling during post-therapy period.

5.0 Main results

This section presents a summary of studies 1-3.

• 5.1 Study 1

“Systematic review of monitoring criteria to interpret CA125 increments during first-line chemotherapy and the subsequent follow-up period among patients with advanced epithelial ovarian cancer”

In total 21 original articles fulfilled the inclusion criteria for the systematic review. CA125 assessment criteria and their accuracies were evaluated in 13 reports during primary therapy and in 8 reports during the subsequent follow-up. The sensitivities for detecting CA125 increments were not reported consistently, but could be calculated from data provided in the articles. The median sensitivity of all the investigated criteria for recurrence was 57% (range 33%-95%) during primary therapy and 85% (range 62%-93%) during follow-up. During primary therapy the calculated FP- and FN-rates were in median 1% (range 0%-13%) and 44% (range 5%-67%), respectively. During follow-up the FP- and FN-rates were in median 9% (range 0%-33%) and 15% (range 7%-38%), respectively. Most reports were inflicted by heterogenic study design and format of presentation. The criteria that were further evaluated in the phase I trial were proposed by Rustin et al. [82, 83, 116] and Tuxen et al. [84, 85]. The criteria are illustrated in *Figures 1-2*. The EPD criteria proposed by Liu et al. [86] operated in another range than the criteria proposed by Rustin et al. and Tuxen et al. and were not evaluated further in study 2.

• 5.2 Study 2

“Monitoring performance of progression assessment criteria for cancer antigen 125 among patients with ovarian cancer compared by computer simulation”

For increments starting from baseline concentrations \geq cut-off, the best performing criterion in terms of low number of FP signals was based on a confirmed increment of ≥ 2.5 times the nadir concentration (*Figure 2E*). For increments starting from baseline concentrations \leq cut-off, the best performing criterion in terms of low number of FP signals was based on a confirmed increment from \leq cut-off to >2 times the cut-off (*Figure 1A*). The lead-time potential of the criteria assessing increments from baseline concentrations \geq cut-off were similar as were the lead-time potential of increment starting \leq cut-off. Ln-transformation reduced the number of FP increment but did not influence on the lead-time potential. The best performing criteria were based on an arbitrary percentage of change without considering the random variation.

• 5.3 Study 3

“Performance of seven criteria to assess CA125 increments among ovarian cancer patients monitored during first-line chemotherapy and the subsequent follow-up period”

The accuracy of the seven investigated criteria (*Figures 1-2*) during first-line chemotherapy and follow-up among all histological tumor types and serous tumors only were similar with overlapping 95% CI. The sensitivities for CA125 increments during first-line chemotherapy and follow-up ranged from 30% to 55%. The

FP-rates ranged from 0% to 17%; however, the FN-rates ranged from 45% to 70%. For increments starting above cut-off including all histological tumor types the lead-times ranged from 27 days to 87 days depending on the applied criterion. For increments starting below cut-off including all histological types the median lead-times ranged from 41 days to 46 days depending on the applied criterion. Inclusion of serous tumor types only did not improve the accuracy or the median lead-times for any of the assessment criteria.

6.0 Discussion

The results from the three studies are discussed at a general level in this section. A detailed discussion is provided in the individual studies.

• 6.1 The monitoring performance of CA125 among patients with EOC

The purpose of CA125 monitoring is to detect treatment failure and recurrence, which may lead to supplementary imaging, abandoning an ineffective treatment or initiating an early effective intervention. The parameters to consider in terms of accuracy are sensitivity, specificity, PPV, FP-rate, NPV, and FN-rate. Relevant biomarkers and reliable criteria to interpret increments in serial concentrations are needed.

Study 1

The systematic review estimated the accuracy of several different types of assessment criteria applied for monitoring of patients with EOC. Most studies were based on small and inhomogeneous patient populations without clear information on stage of disease and histological type of ovarian cancer [117]. Some studies reported performance data based on a single criterion whereas some of the data reported by Rustin et al. and all data reported by Tuxen et al. were based on a combination of more criteria. Even though the studies reporting on monitoring performance during primary therapy were heterogeneous in several aspects, the 95% CI for the calculated values for sensitivity, FP-rates, and FN-rates were mostly overlapping as were the 95% CI for studies reporting on monitoring performance during follow-up after primary therapy. Apparently, complex criteria based on consecutive measurements, defined sampling intervals, and/or analytical and biological variation did not outperform simple criteria based on an arbitrary percentage of increase i. e. 50% or 100%. It was surprising that different approaches among different patient populations provided similar results both during therapy and follow-up, respectively. It may be concluded that studies conducted almost through three decades basically provided similar results. Comparison of performances between therapy and follow-up also revealed mostly overlapping 95% CIs suggesting that the respective sets of criteria applied during therapy and follow-up performed similarly. However, it may be speculated that the requirements were more demanding for criteria intended for monitoring therapy as compared to criteria intended for monitoring follow-up because the sensitivities and FP-rates tended to be lower and the FN-rates higher during therapy. Thus, the medians and ranges for sensitivities, FP-rates, and FN-rates during therapy and follow-up, respectively, were: 57% (range 33%-95%), 1% (range 0%-13%), and 44% (range 5%-67%) vs. 85% (range 62%-

93%), 9% (range 0%-33%), and 15% (range 7%-38%). Overall, the results suggested that regardless of the approach, fine-tuning of the assessment criteria did not seem to improve their monitoring performances indicating that CA125 used as a tumor marker for monitoring has inherent limitations in terms of accuracy.

Study 2

This is the first investigation designed as a preclinical phase I trial based on simulated CA125 data. The monitoring performance of each individual criterion proposed by Rustin et al. (*Figures 1A-B*) and Tuxen et al. (*Figures 2A-E*) was estimated under standardized conditions in terms of frequency of FP CA125 increments among the surrogate healthy individuals and the time needed to detect 100% of the TP CA125 increments among the surrogate patients. Assuming a Gaussian distribution of CA125 baseline concentrations, the criterion suggested by Tuxen et al. (*Figure 2E*) performed best for CA125 increments starting above cut-off. The criterion provided 5.4% FP increments cumulated during 12 months and 22.5% FP increments cumulated during 36 months. It could be argued that the criterion (*Figure 2E*) provided too many FP events to be useful for excluding tumor growth in clinical routine. However, following ln-transformation of baseline concentrations, the number of FP signals was reduced considerably to 2% and 3% cumulated from 12 months and 36 months of monitoring suggesting that the criterion may be considered for use in clinical practice if CA125 data are ln-transformed before interpretation [107]. Our observation is supported by Fokkema et al. who also reported a reduction of FP signals following ln-transformation of data for B-type natriuretic peptide, a biomarker used to monitor patients with heart failure [109]. Assuming a Gaussian distribution of CA125 baseline concentrations, the criterion suggested by Rustin et al. (*Figure 1A*) performed best for CA125 increments starting below cut-off because there were no FP increments during 36 months of monitoring. Following ln-transformation of baseline concentrations the number of FP increments remained at 0% suggesting that the criterion may be considered for use in clinical practice irrespective of Gaussian distribution or ln-transformation of CA125 data [107]. The reason why the criteria suggested by Tuxen et al. (*Figure 2E*) and by Rustin et al. (*Figure 1A*) tended to provide fewer FP signals as compared to the other criteria may be due to their design requiring larger increments and consecutive measurements [107].

Study 3

In the phase II trial the sensitivities of the individual CA125 assessment criteria suggested by Tuxen et al. (*Figures 2A-E*) ranged from 30% to 51%. The data were cumulated from EOC patients during first-line chemotherapy and the subsequent follow-up period including all histological tumor types [118]. In a previous study Tuxen et al. reported sensitivities of 33.3% (*Figures 2D-E*) and 45.8% (*Figures 2A-C*) by combining more criteria during first-line chemotherapy by including patients with all histological types of EOC [84]. Even though study 3 investigated the performance of individual criteria during therapy and follow-up and Tuxen et al. reported on a combination of criteria during chemotherapy the results were similar [84]. The required large increments and consecutive measurements required by the criteria may explain the low sensitivities in both studies. For increments starting below cut-off, the criterion suggested by Rustin et al. (*Figure 1A*) provided a high FN-rate in study 3 and a low FP-rate in study 2. It may

be argued that the finding is explained by the inverse relationship between the number of FN and FP events owing to the nature of a quantitative biochemical test - the higher the number of FN events the lower the number of FP events [106]. Thus, study 3 is in accordance with study 2 supporting the view that simulation models may be helpful for a preclinical evaluation of the monitoring potential of different assessment criteria to interpret increments in serial cancer biomarker concentrations. In a study focusing on CA125 for monitoring ovarian cancer during follow-up, Tuxen et al. reported sensitivities of the criteria (*Figures 2D-E*) to be higher than the sensitivities calculated in study 3 [85]. However, Tuxen et al. reported the combined sensitivity of both criteria which may have provided more TP events than a single criterion. A multicenter study, conducted by the UK-based Medical Research (MRC) and the EORTC [91], reported a median CA125 lead-time of 4.8 months for recurrence among EOC patients in complete remission after first-line chemotherapy and with CA125 concentrations <35 kU/L. The study applied the criterion suggested by Rustin et al. (*Figure 1A*) to confirm progression [91]. In study 3 the same criterion provided a median lead-time of 41 days which was shorter than the median lead-time reported by the multicenter study. Even though the criterion requires that the latest concentration is used for lead-time calculations, the multicenter study based the calculation on the second latest concentration which may have overestimated the lead-time of the criterion [91]. The clinical utility of the length of positive lead-times should be evaluated separately for treatment and follow-up periods. During treatment the positive lead-time, even though a new treatment could be initiated, may not alter the prognosis in EOC patients. However, several patients may avoid unnecessary toxicity associated with one or more cycles of ineffective treatment. One could believe that the length of a positive lead-time at progression during follow-up would influence the overall survival by starting treatment before the clinical progression occurred, remarkably, the multicenter study conducted by MRC and the EORTC showed no benefit from early detection of recurrence by CA125 measurements [91]. New treatments options for EOC may change this.

• 6.2 Image techniques versus CA125

Owing to limited accuracy of CA125, supplementary investigative procedures for surveillance of EOC patients are needed [119]. Among patients with primary ovarian cancer the conventional imaging techniques TVS and CT have a sensitivity of 71-96% and 72-82%, respectively, and specificities of 23-83% and 53-81%, respectively [120]. Chandrashekhara et al. [121] reported that CT had a sensitivity of 17% to 88% with a specificity of >90% for detection of peritoneal metastases among patients with EOC. Accordingly, the conventional imaging techniques underestimated the stage of disease in ovarian cancer patients due to difficulties in identifying peritoneal dissemination. Magnetic resonance (MR) proved more precise than CT in detecting lymph node metastasis but had similar low accuracy for peritoneal and liver surface evaluation. Combining CT or MR with single positron emission tomography with fluorodeoxyglucose (FDG-PET) improved detection of distant metastases from ovarian cancer in the pelvis [61, 122]. However, a major limitation of PET imaging is visualization of widespread peritoneal carcinomatosis. The problem derives from the reduced spatial resolution of PET imaging (5-

6 mm) that makes it unable to detect small tumor lesions <5 mm resulting in high FN-rates. Other well-known limitations of PET imaging are due to bowel and urinary artefacts during excretion of the radiotracer [122]. Additionally, repetitive evaluation by PET/CT can give a significant radiation burden [61]. Thus, a study among pediatric patients showed a triple risk of leukemia after a cumulative CT dose of 50 milligram [123]. The radiation exposure from PET is minor in comparison to the radiation dose from CT [61]. The potential adverse effects of CT and PET imaging need to be considered when monitoring EOC patients because of repetitive evaluations during the course of their disease. Overall, the combination of PET and MR may be a better choice than PET and CT when monitoring ovarian cancer patients because of a high resolution in soft tissue, however, the implementation in clinical practice may be challenging due to the high costs [61].

- **6.3 Gynecological Cancer Intergroup criteria**

In 2010 the progression criteria proposed by Rustin et al. in 1996 and in 2001 (*Figures 1A-B*) were officially recommended by the Gynecological Cancer Intergroup (GCIg) to be incorporated into clinical trials as the criteria were considered to be sufficiently validated [60, 124]. Accordingly, patients should be declared to have progressive disease during follow-up either on the basis of the RECIST 1.1 progression criteria or on the basis of CA125 progression criteria where the date of progression should be ascribed to the earliest signal of progression [60].

Several randomized clinical trials incorporated the two GCIg CA125 progression criteria to evaluate new treatment modalities [125-130]. In 2010 the CALYPSO group incorporated both the GCIg progression criteria along with the RECIST criteria to assess progression among women with platinum-sensitive recurrent ovarian carcinoma Stage I-IV [128]. The trial concluded that the GCIg criteria and the RECIST criteria performed similarly in terms of ability to detect tumor growth, and that CA125 could be used as an alternative to imaging techniques in clinical trials [131]. In 2013, Levy et al. reported a study using the GCIg CA125 progression criterion (*Figure 1A*) to assess whether CA125 increments from below to above the normal range at recurrence was associated with outcome among patients with a complete clinical response after initial treatment. The results suggested that the rate of CA125 increase among patients with recurrence was of prognostic significance [132]. Even though the GCIg CA125 progression criteria are used in clinical trials, studies 2 and 3 indicated that the criterion generated by Rustin et al. (*Figure 1B*) may signal high rates of FP and FN information, respectively, questioning the utility of the criterion in daily clinical practice. A recent study from 2016 by Lindemann et al. also questioned the clinical utility of CA125 [133]. The authors investigated the concordance between CA125 defined and RECIST defined progression using data from the GCIg randomized phase III trial AURELIA trial in platinum-resistant ovarian cancer. CA125 progression was defined according to GCIg (except that confirmatory CA125 measurement was not applied). The authors concluded that progression was typically detected earlier by imaging than by CA125 and disease status at the time of progression was not prognostic for overall survival.

- **6.4 Early Signals of Progressive Disease criteria**

In 2007 Liu et al. investigated the monitoring performance of their two EPD criteria among 178 patients with stage III-IV ovarian

cancer, with complete response to primary therapy and with baseline CA125 concentrations below 30 kU/L [86]. They compared the performances of the EPD criteria (*Figures 3A-B*) with the performance of the GCIg criterion (*Figure 1A*) and reported that the proposed EPD criteria provided a low FP-rate and early prediction of disease progression in more than 50 % of the patients. Accordingly, Liu et al. suggested that their criteria were better to signal tumor growth as compared to the GCIg criterion. However, there were major drawbacks of the study because it was impossible to calculate the sensitivity and specificity of CA125 monitoring, additionally the histology of ovarian tumors among the investigated patients were not specified. In 2009, Prat et al. reported on the EPD criteria in a retrospective study based on 96 patients with advanced EOC (Stage III-IV) [134]. They found that an increase in serum CA125 concentration within the normal range was an independent predictive factor for disease recurrence. In 2012 Levy et al. performed a similar analysis using the EPD criteria and reported that the criteria had predictive value in terms of early signals of clinical recurrence [135]. The EPD criteria were not investigated in study 2 and 3 because the criteria operate within the normal range and therefore need to be addressed in more detail with a different approach.

- **6.5 Clinical value of early detection of recurrence by CA125**

Even though rising CA125 concentrations may signal recurrent disease the clinical value of early detection remains unclear, the issue has only been addressed by one randomized controlled clinical trial in 2010 [91]. The results showed that neither survival nor quality of life improved by early CA125 guided therapy [91]. The authors advised to abandon CA125 monitoring in the routine follow-up of patients with ovarian cancer; instead they should be informed about the most common symptoms prompting an appointment with a specialist and rapid access to CA125 testing. However, the study design has been debated because it was conducted over a decade where the therapies varied and the patients did not receive the most effective therapies, the study may have underestimated the benefit of early recurrence detection [136, 137]. Additionally, CA125 analysis was decentralized to several laboratories without information on the analytical quality [136]. The European Society of Gynecologic Oncologists (ESGO) recently advised against universally abandoning CA125 in the routine follow-up of all patients with ovarian cancer based on this single randomized trial [138]. Accordingly, CA125 monitoring should be considered in patients who i) after complete response on primary treatment have been or are being treated as part of a clinical trial, ii) are eligible for (future) clinical trials on second-line treatment, iii) refuse routine (3 monthly) follow-up including regular imaging, and iv) are eligible for secondary surgery at recurrence. Also the European Society of Medical Oncology (ESMO) advise against abandoning CA125 monitoring during follow-up because regular CA125 measurements may signal tumor growth in some patients before symptoms appear [136, 137]. The current position of the EGTM is that CA125 is recommended for monitoring subgroups of patients if monitoring is likely to have clinical consequences [55]. The Danish Gynecologic Cancer Group (DGCG) suggest that CA125 should only be measured among patients if the need has been expressed during the follow-up period [3].

Our findings in study 3 supported the study conducted by the UK-based MRC and the EORTC in questioning the clinical utility of serial CA125 measurements. However, several aspects need to be investigated before drawing firm conclusions regarding the clinical value of CA125 in detecting recurrence. Since the conduction of study 3 and the UK-based MRC and the EORTC trial, the conventional histological classification system has changed. Currently, there is substantial evidence that identifies EOC as a heterogeneous disease with five distinct subtypes and different origins [11]. The new classification system may have an impact on the monitoring approaches in EOC due to differences in clinical behavior, response to chemotherapy, pattern of metastases, and survival [12]. There is a need for specific and individualized treatment for each identified subtype of EOC that may change and perhaps optimize the clinical utility of CA125 [11, 12].

At this stage, challenges remain on how tumor marker monitoring studies should be designed. The EGTM has recognized the challenges associated with planning, conducting, and reporting clinical tumor marker surveillance programs and now offers advice on how to conduct longitudinal monitoring studies. The EGTM has proposed a four-phase model to design biomarker monitoring trials analogous to the stepwise approach used to investigate new drugs [114].

- **6.6 Other biomarkers**

Human Epididymis Protein 4

HE4 was recently proposed for the diagnosis of EOC because of increased specificity of HE4 as compared to CA125, mainly in premenopausal women [139]. The sensitivity of HE4 and CA125 was suggested to be similar [29, 64-66]. As regards measurements during the menstrual cycle Hallamaa et al. found that HE4 among healthy premenopausal women and women with endometriosis contrary to CA125 was not influenced by the menstrual cycle or hormonal medication [140]. Moore et al. developed the Risk of Ovarian Malignancy Algorithm (ROMA), combining HE4 and CA125 to predict the risk of serous EOC in women with a pelvic mass [141]. The algorithm was solely based on the two biomarkers HE4 and CA125, as compared to RMI, which required information on CA125, menopause status, and an ultrasound examination [56, 142, 143]. Van Gorp et al. reported that neither HE4 nor the ROMA performed better in the differentiation of ovarian cancer from other pelvic masses as compared to CA125 [66]. Molina et al. suggested that HE4 and CA125 in combination predicted PFS and OS at least as well as HE4 and significantly better than CA125 [29]. Karlsen et al. compared a new algorithm, the Copenhagen Index (CPH-I) based on CA125, HE4, and age with the ROMA and the RMI [144]. The CPH-I, RMI, and ROMA performed similarly in distinguishing benign ovarian tumors from malignant ovarian cancer. Due to the conflicting results further prospective studies are needed to elucidate whether HE4 measurements should be implemented into clinical routine [55].

Cancer Antigen 19-9 and Carcinoembryonic Antigen

CA 19-9 and CEA have been suggested as useful biomarkers in patients with the mucinous type of ovarian cancer [31, 32], however, their relevance for monitoring remains unclear and they are not widely used [31, 145].

7.0 Strengths and limitations

- **7.1 Study 1**

Strengths

The systematic review approach was chosen to address part of the aims due to the robustness of the methodology, appraisal of the literature, and transparency. The procedure for the systematic review followed the PRISMA guidelines because of the standardized format for presentation [103]. QUADAS-2, a well-known tool for studies reporting on diagnostic accuracies was used to evaluate the quality of the included studies and the risk of bias [105].

Limitations

The original articles identified in study 1 were not subjected to a meta-analysis due to inconsistent reporting and lack of information. The systematic review method was used instead [104]. The clarity of some of the original reports regarding procedures to calculate the accuracy of CA125 and the lead-time was not always quite clear which rendered interpretation difficult of some of the identified results. It may be argued that study 1 may be inflicted by: 1) Reporting bias: owing to incomplete identification of all relevant publications, 2) Publication bias: owing to selective reporting of complete studies and the tendency for positive results to achieve publication. 3) Selection bias: owing to the possibility that the study participants were unrepresentative for the target disease. 4) Confirmation bias: owing to the possibility that the individual authors had a tendency to seek information that confirmed their hypothesis rather than data that facilitated efficient testing of competing hypothesis. 5) Incorporation bias: when information related to the test result is incorporated into the diagnostic criteria.

Based on the thorough search strategy and knowledge of the literature it is likely that all major studies addressing CA125 assessment criteria were identified. It is difficult to argue on the impact of the different types of bias on the individual original studies. However, it may be argued that the search strategy and inclusion as well as the exclusion criteria were too narrow which may have led to omission of some studies. Publications in other languages than English were also excluded and thereby some minor trials could be missing. Additionally, most of the included original studies were heterogenic due to study design and selection of the patients. Overall, even though study 1 has shortcomings it is the first study summarizing the accuracy of the large number of CA125 assessment criteria developed during the last three decades [117].

- **7.2 Study 2**

Strengths

The model system is effective because the robustness against FP signals and the ability to early detection of tumor growth by the individual criteria could be estimated among a large number of surrogate healthy individuals and patients. The system has the ability to test the performance of several assessment criteria on the same sets of simulated data which allows a solid comparison. Simulation models have the potential to be used as supplementary tools during phase 1 trials to investigate the performances of assessment criteria for new cancer biomarkers and as tool to optimize already existing criteria thereby saving time and money

compared to large clinical trials. It is important to recognize that computer simulations should not be used as a substitute for clinical investigations.

Limitations

A drawback of computer simulations is that the basic parameters used in the model have to be available from clinical publications; otherwise the obtained results are misleading. So far the model system has only been used in a few studies validating the performance of assessment criteria intended to interpret increments in serial concentrations of other biomarkers than CA125. Consequently, it is difficult to compare the current simulated CA125 data with those of others [97, 98, 100-102]. The applied biological variation (19.8%) was derived from healthy women [93] which may have influenced the results obtained for FP-rate and lead-time potential. It may be argued that the biological variation obtained among ovarian cancer patients with steady state CA125 concentrations (24%) should have been applied instead [94]. On the other hand, because of the resemblances of the two biologic variations it is less likely that the presented approach had major influence on the results. Another drawback of the simulation model is that the number of FN signals provided by the investigated criteria could not be estimated.

- **7.3 Study 3**

Strengths

The study was a longitudinal phase II monitoring trial based on a homogeneous cohort of EOC patients where the samples and the clinical information were collected prospectively according to the STARD guidelines which provides guidance on how to conduct diagnostic accuracy studies of biomarkers [111]. The EGTM has recommended that tumor biomarker monitoring trials should involve four phases, where phase II trials should validate the performance of relevant assessment criteria that were identified in previous phase I trials [114]. Study 3 followed the recommended design. It strengthens study 3 that the assessment criteria were validated among a cohort of patients where increasing tumor burden was very likely. As compared to previously reported monitoring trials [72-76, 79] study 3 provided a long surveillance period of six years. Previous monitoring studies mostly reported shorter monitoring periods or did not report the duration of their studies [37, 67-70, 74]. Following collection, the CA125 samples were analyzed shortly after in separate assay runs according to the local quality assurance procedures at one routine laboratory which strengthens the validity of the presented CA125 results [114]. Study 3 investigated the performance of the assessment criteria both among all eligible tumor types and among the serous EOC type only [55].

Limitations

One of the limitations of study 3 is that the investigation was not blinded, the CA125 data was available throughout the study period together with the results of clinical examinations and imaging [111]. This may have influenced the length of the positive lead-time because the clinicians had the opportunity to request earlier imaging based on CA125 increments and thereby shorten the potential lead-time. Another concern is that a power calculation was not performed before planning the study and collection of the blood samples during 1995-2001. However, a power-

calculation was performed in 2013 before the registers containing the collected CA125 and clinical data were opened, scrutinized, systematized, and evaluated. According to the power calculation, 156 eligible patients were required in order to obtain a solid estimate of the monitoring performances of the investigated criteria. By including ovarian cancer patients with all histological tumor types the number of eligible patients was 189, thus meeting the requirement of the power calculation. However, when including ovarian cancer patients with serous tumors only the number of eligible patients was 112 being well below the required number. It may be argued that the results obtained among ovarian cancer patients with serous tumors only did not meet the requirement and therefore should be evaluated cautiously. On the other hand, inclusion of serous tumor types only did not improve the accuracy of the investigated criteria as compared to situations when all tumor types were included. Therefore a potential difference between the monitoring performances of the criteria among the two groups of patients may be minor or not existing. Another weakness is the changed histological classification of EOC now considered a heterogeneous disease with five different subtypes [12]. Most likely, the new classification system will provide a different distribution of patients among the two groups, all tumors and serous tumors only, respectively, inflicting the presented results with some uncertainty.

A further weakness adheres to the clinical response evaluation which was based on criteria of the WHO in use at the time of study 3 [113]. In 2000, the WHO standards were replaced by a set of new guidelines to evaluate the response to treatment in solid tumors (RECIST) [58]. The RECIST guidelines were further updated in 2009 (RECIST version 1.1) [59]. Re-evaluation of clinical response among the investigated patients according to the new standards may have some impact on the obtained results but would hardly influence the overall impression of the validity of CA125 as a monitor of patients with EOC.

8.0 Future perspective

- **8.1 Alternative assessment criteria**

Study 1-3 indicated a need for reliable alternative assessment criteria to detect recurrence among ovarian cancer patients. ROCA identified significant rises in CA125 above baseline and could be an interesting alternative for interpretation of serially measured CA125 concentrations over time [47]. It would be relevant to investigate the performance of ROCA in phase I and II monitoring trials according to EGTM guidelines [114]. Another challenge is situations with slow rates of CA125 increase from low baseline concentrations providing FN events in terms of progression. In an effort to reduce the rate of FN events without numerous FP signals it seems relevant to explore criteria that assess increments within normal range. The EPD criteria suggested by Liu et al. in 2007 (*Figures 3A-B*) should be explored further in phase I computer simulation studies [86].

- **8.2 Alternative biomarkers**

The monitoring performance of HE4 in combination with CA125 has as yet not been investigated among EOC patients undergoing first-line chemotherapy and during the subsequent follow-up period. Evidently, identification of new biomarkers is important and the area is developing fast i.e. circulating tumor cells, Deoxy-

ribonucleic acid (DNA) and Ribonucleic acid (RNA) fragments as well as epigenetic alterations [146, 147].

9.0 Conclusions

The presented PhD study highlighted that the seven investigated CA125 assessment criteria showed low ability to detect and to exclude tumor growth. Criteria based on random variation (analytical and biological) did not perform better than simple criteria based on an increment that had to exceed an arbitrary defined percentage. Simulation models could be valuable for a preclinical evaluation of the monitoring potential of different assessment criteria to interpret increments in serial cancer biomarker concentrations. Our results are in accordance with a much debated previous report as well as a recent report questioning the clinical utility of CA125 monitoring. Alternative assessment criteria along with discovery of new biomarkers based on the new classification of EOC are needed. As yet, CA125 remains the most used serological biomarker for monitoring of patients with EOC.

Abbreviations

AFP: Alpha-fetoprotein

AMH: Antimüllerian Hormone

β HCG: Beta Human Chorion Gonadotropin

CA125: Cancer Antigen 125

CA 19-9: Carbohydrate Antigen 19-9

CA 15.3: Cancer Antigen 15.3

CD: Critical difference

CEA: Carcinoembryonic Antigen

CI (95%): 95% confidence intervals

CPH-I: Copenhagen Index

CT: Computed tomography

CVA: The total analytical imprecision

CVI: The within-subject-biological variation

DGCG: Danish Gynecologic Cancer Group

DNA: Deoxyribonucleic acid

EGTM: European Group on Tumor Markers

EOC: Epithelial Ovarian Cancer

EPD: Early Signal of Progressive Disease

ESGO: European Society of Gynecologic Oncologists

ESMO: European Society of Medical Oncology

FDG: Fluorodeoxyglucose

FIGO: The International Federation of Gynecology and Obstetrics

FN: False negative

FP: False positive

GCIG: Gynecological Cancer Intergroup

HE4: Human Epididymis Protein 4

λ : Exponential function lambda

LDH: Lactate Dehydrogenase

MR: Magnetic resonance

MRC: Medical Research

NPV: Negative predictive value

NROSS: the Normal Risk Ovarian Screening Study

OS: Overall survival

PET: Positron emission tomography

PFS: Progression free survival

PPV: Positive predictive value

PRISMA: Preferred Reporting Items for Systematic Reviews and Meta-Analyses

QUADAS: Quality Assessment of Diagnostic Accuracy Studies

RCV: Reference change value

RECIST: Response Evaluation Criteria in Solid Tumor

RMI: Risk Malignancy Index

RNA: Ribonucleic acid

ROCA: The Risk of Ovarian Cancer Algorithm

ROMA: The Risk of Ovarian Malignancy Algorithm

STARD: Standards for Reporting Diagnostic Accuracy Studies

TN: True negative

TP: True positive

TPA: Tissue Polypeptide Antigen

TVS: Transvaginal ultrasound

UKCTOCS: UK Collaborative Trial of Ovarian Cancer Screening

WHO: World Health Organization

Funding

The following foundations and institutions funded this PhD-thesis:

- Klinisk Biokemisk Afdeling, Nordsjællands Hospital
- Helen Rudes Fond
- Fru Olga Bryde Nielsens Fond
- Ingeniør August Frederik Wedell Erichsens Legat
- Fonden til fremme af Klinisk Cancerforskning

Summary

Cancer Antigen 125 (CA125) is frequently used in the routine monitoring of patients with epithelial ovarian cancer (EOC). The potential benefit is based on the assumption that changes in serial concentrations may provide early and reliable information on tumor growth expediting an early and potentially effective treatment. However, it has remained a challenge to interpret increments in concentrations that correlate with increasing tumor burden in the individual patient. It has been hypothesized that CA125 assessment criteria taking the random variation (analytical and biological) into account may have better accuracy and lead-time potential than criteria based only on an arbitrary percentage of increase.

The aims of the current PhD project was to i) identify different types of assessment criteria intended to interpret CA125 increments, ii) compare their ability to signal tumor growth, and iii) estimate the time interval between marker progression and clinical progression (lead-time).

Study 1 was a systematic review of the literature identifying 21 relevant original articles reporting on 37 different assessment criteria to interpret serial CA125 concentrations. Study 2 was a preclinical phase I trial investigating the monitoring potential of seven selected criteria in a computer-based simulation model under standardized conditions. Study 3 was a clinical phase II trial comparing the performances of the seven criteria among 189 patients with EOC stage IC-IV during first-line chemotherapy and the subsequent follow-up period.

Study 1 reported that the median sensitivity of the investigated criteria for recurrence was 57% (range 33%-95%) during primary therapy and 85% (range 62%-93%) during follow-up. The calculated false positive (FP) and false negative (FN) rates, respectively, were in median 1% (range 0%-13%) and 44% (range 5%-67%) during primary therapy and 9% (range 0%-33%) and 15% (range 7%-38%) during follow-up. Most of the reports were heterogeneous in terms of study design and format of presentation. Study 2 reported that for increments starting from baseline concentrations \geq cut-off, the best performing criterion in terms of low number of FP signals was based on a confirmed increment ≥ 2.5 times the nadir concentration. For increments starting from baseline concentrations \leq cut-off, the best performing criterion, also in terms of low number of FP signals, was based on a confirmed increment from \leq cut-off to >2 times the cut-off. Accordingly, the best performing criteria in terms of low number of FP events were based on an arbitrary required percentage of change without defining the random variation. Study 3 reported that the accuracy of the seven criteria observed during first-line chemotherapy and follow-up among all histological tumor types and serous tumors only was similar with overlapping 95% confidence intervals. The sensitivities for detecting CA125 increments ranged from 30% to 55%. The FP rates ranged from 0% to 17%; however, the FN rates ranged from 45% to 70%. The median lead-times

ranged from 26 days to 87 days. The performances of the CA125 assessment criteria showed low sensitivities and low ability to exclude tumor growth. The chance of developing clinical progression following CA125 progression was high (range of positive predictive value 90%-100%); however, the lead-times were short among several patients. Thus, study 3 questioned the clinical utility of CA125 monitoring.

Overall, the PhD study showed, that the different CA125 assessment criteria basically provided similar results thus rejecting the hypothesis that criteria based on calculating the random variation would outperform criteria based on a simple percentage of change. The simulated data proved useful for a preclinical evaluation of CA125 assessment criteria. The results suggested that regardless of the approach, fine-tuning of the assessment criteria did not seem to improve the monitoring performance of CA125 probably indicating that CA125 used as a tumor marker for monitoring has inherent limitations in terms of accuracy. Supplementary markers and alternative assessment criteria are needed.

References

- [1] Lutz AM, Willmann JK, Drescher CW, et al. Early diagnosis of ovarian carcinoma: is a solution in sight? *Radiology* 2011;259:329-45.
- [2] Siegel RL, Miller KD, Jemal A. Cancer statistics, 2016. *CA Cancer J Clin* 2016;66:7-30.
- [3] Dansk Selskab for Obstetrik og Gynækologi, Dansk Selskab for Onkologi, Dansk Patologiselskab. Retningslinier for visitation, diagnostik, behandling og opfølgning af epitelial ovarie-, tuba- og primær peritonealcancer samt borderline tumore. 2016. <http://www.dgco.dk/index.php/guidelines/ovariecancer-guidelines> (Assessed September 2016).
- [4] Sturgeon CM, Duffy MJ, Stenman UH, et al. National Academy of Clinical Biochemistry laboratory medicine practice guidelines for use of tumor markers in testicular, prostate, colorectal, breast, and ovarian cancers. *Clin Chem* 2008;54:e11-e79.
- [5] Guppy AE, Rustin GJ. CA125 response: can it replace the traditional response criteria in ovarian cancer? *Oncologist* 2002;7:437-43.
- [6] New FIGO ovarian cancer staging guidelines. 2014. <https://www.sgo.org/clinical-practice/guidelines/new-figo-ovarian-cancer-staging-guidelines/> (Assessed September 2016).
- [7] Aletti GD, Gallenberg MM, Cliby WA, Jatoi A, Hartmann LC. Current management strategies for ovarian cancer. *Mayo Clin Proc* 2007;82:751-70.
- [8] Ramus SJ, Gayther SA. The contribution of BRCA1 and BRCA2 to ovarian cancer. *Mol Oncol* 2009;3:138-50.

- [9] Kauff ND, Mitra N, Robson ME, et al. Risk of ovarian cancer in BRCA1 and BRCA2 mutation-negative hereditary breast cancer families. *J Natl Cancer Inst* 2005;97:1382-4.
- [10] NCCN Clinical Practice guidelines in Oncology: Genetic/familial high-risk assesment: Breast and Ovarian Cancer. 2010. https://www.nccn.org/professionals/physician_gls/f_guidelines.asp (Assessed September 2016).
- [11] Integrated genomic analyses of ovarian carcinoma. The Cancer Genome Atlas Research Network. *Nature* 2011;474:609-15.
- [12] Banerjee S, Kaye SB. New strategies in the treatment of ovarian cancer: current clinical perspectives and future potential. *Clin Cancer Res* 2013;19:961-8.
- [13] Forstner R, Meissnitzer M, Cunha TM. Update on Imaging of Ovarian Cancer. *Curr Radiol Rep* 2016;4:31.
- [14] Kobel M, Kalloger SE, Boyd N, et al. Ovarian carcinoma subtypes are different diseases: implications for biomarker studies. *PLoS Med* 2008;5:e232.
- [15] Kobel M, Rahimi K, Rambau PF, et al. An Immunohistochemical Algorithm for Ovarian Carcinoma Typing. *Int J Gynecol Pathol* 2016; 35(5):430-41.
- [16] Argento M, Hoffman P, Gauchez AS. Ovarian cancer detection and treatment: current situation and future prospects. *Anticancer Res* 2008;28:3135-8.
- [17] Hippisley-Cox J, Coupland C. Identifying women with suspected ovarian cancer in primary care: derivation and validation of algorithm. *BMJ* 2012;344:d8009.
- [18] Goff BA, Mandel LS, Drescher CW, et al. Development of an ovarian cancer symptom index: possibilities for earlier detection. *Cancer* 2007;109:221-7.
- [19] Hamilton W, Peters TJ, Bankhead C, Sharp D. Risk of ovarian cancer in women with symptoms in primary care: population based case-control study. *BMJ* 2009;339:b2998.
- [20] Parente-Pereira AC, Whilding LM, Brewig N, et al. Synergistic Chemoimmunotherapy of Epithelial Ovarian Cancer Using ErbB-Retargeted T Cells Combined with Carboplatin. *J Immunol* 2013;191:2437-45.
- [21] Hayes DF, Bast RC, Desch CE, et al. Tumor marker utility grading system: a framework to evaluate clinical utility of tumor markers. *J Natl Cancer Inst* 1996;88:1456-66.
- [22] Werner M, Faser C, Silverberg M. Clinical utility and validation of emerging biochemical markers for mammary adenocarcinoma. *Clin Chem* 1993;39:2386-96.
- [23] Duffy MJ. Tumor markers in clinical practice: a review focusing on common solid cancers. *Med Princ Pract* 2013;22:4-11.
- [24] Duffy MJ, Sturgeon CM, Soletormos G, et al. Validation of new cancer biomarkers: a position statement from the European group on tumor markers. *Clin Chem* 2015;61:809-20.
- [25] Hayes DF. Determination of clinical utility of tumor markers: a tumor marker utility grading system. *Recent Results Cancer Res* 1998;152:71-85.
- [26] Aggarwal P, Kehoe S. Serum tumour markers in gynaecological cancers. *Maturitas* 2010;67:46-53.
- [27] Montagnana M, Lippi G, Ruzzenente O, et al. The utility of serum human epididymis protein 4 (HE4) in patients with a pelvic mass. *J Clin Lab Anal* 2009;23:331-5.
- [28] Gao L, Cheng HY, Dong L, et al. The role of HE4 in ovarian cancer: inhibiting tumour cell proliferation and metastasis. *J Int Med Res* 2011;39:1645-60.
- [29] Molina R, Escudero JM, Auge JM, et al. HE4 a novel tumour marker for ovarian cancer: comparison with CA 125 and ROMA algorithm in patients with gynaecological diseases. *Tumour Biol* 2011;32:1087-95.
- [30] Lu R, Sun X, Xiao R, Zhou L, Gao X, Guo L. Human epididymis protein 4 (HE4) plays a key role in ovarian cancer cell adhesion and motility. *Biochem Biophys Res Commun* 2012;419:274-80.
- [31] Hogdall EV, Christensen L, Kjaer SK, et al. Protein expression levels of carcinoembryonic antigen (CEA) in Danish ovarian cancer patients: from the Danish 'MALOVA' ovarian cancer study. *Pathology* 2008;40:487-92.
- [32] Kelly PJ, Archbold P, Price JH, Cardwell C, McCluggage WG. 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.
- [33] Kim HS, Park CY, Lee JM, et al. Evaluation of serum CA-125 levels for preoperative counseling in endometrioid endometrial cancer: a multi-center study. *Gynecol Oncol* 2010;118:283-8.
- [34] Gagnon A, Ye B. Discovery and application of protein biomarkers for ovarian cancer. *Curr Opin Obstet Gynecol* 2008;20:9-13.

- [35] Tuxen MK, Sölétormos G, Dombernowsky P. Tumor markers in the management of patients with ovarian cancer. *Cancer Treat Rev* 1995;21:215-45.
- [36] Holcomb K, Vucetic Z, Miller MC, Knapp RC. Human epididymis protein 4 offers superior specificity in the differentiation of benign and malignant adnexal masses in premenopausal women. *Am J Obstet Gynecol* 2011;205:358-6.
- [37] Bast RC, Jr., Klug TL, St JE, et al. A radioimmunoassay using a monoclonal antibody to monitor the course of epithelial ovarian cancer. *N Engl J Med* 1983;309:883-7.
- [38] Donach M, Yu Y, Artioli G, et al. Combined use of biomarkers for detection of ovarian cancer in high-risk women. *Tumour Biol* 2010;31:209-15.
- [39] Bonfrer JM, Korse CM, Verstraeten RA, Van Kamp GJ, Hart GA, Kenemans P. Clinical evaluation of the Byk LIA-mat CA125 II assay: discussion of a reference value. *Clin Chem* 1997;43:491-7.
- [40] Duffy MJ, Bonfrer JM, Kulpa J, et al. CA125 in ovarian cancer: European Group on Tumor Markers guidelines for clinical use. *Int J Gynecol Cancer* 2005;15:679-91.
- [41] Liu J, Matulonis UA. Anti-angiogenic agents in ovarian cancer: dawn of a new era? *Curr Oncol Rep* 2011;13:450-8.
- [42] Molina R, Auge JM, Escudero JM, et al. Mucins CA 125, CA 19.9, CA 15.3 and TAG-72.3 as tumor markers in patients with lung cancer: comparison with CYFRA 21-1, CEA, SCC and NSE. *Tumour Biol* 2008;29:371-80.
- [43] Hogdall EV, Christensen L, Kjaer SK, et al. Protein expression levels of carcinoembryonic antigen (CEA) in Danish ovarian cancer patients: from the Danish 'MALOVA' ovarian cancer study. *Pathology* 2008;40:487-92.
- [44] Molina R, Ojeda B, Filella X, et al. A prospective study of tumor markers CA 125 and CA 19.9 in patients with epithelial ovarian carcinomas. *Tumour Biol* 1992;13:278-86.
- [45] Jacobs I, Bast RC, Jr. The CA 125 tumour-associated antigen: a review of the literature. *Hum Reprod* 1989;4:1-12.
- [46] Skates SJ, Mai P, Horick NK, et al. Large prospective study of ovarian cancer screening in high-risk women: CA125 cut-point defined by menopausal status. *Cancer Prev Res (Phila)* 2011;4:1401-8.
- [47] Skates SJ. Ovarian cancer screening: development of the risk of ovarian cancer algorithm (ROCA) and ROCA screening trials. *Int J Gynecol Cancer* 2012;22 Suppl 1:S24-S26.
- [48] Jacobs IJ, Menon U, Ryan A, et al. Ovarian cancer screening and mortality in the UK Collaborative Trial of Ovarian Cancer Screening (UKCTOCS): a randomised controlled trial. *Lancet* 2016;387:945-56.
- [49] The role of the generalist obstetrician-gynecologist in the early detection of ovarian cancer. *Int J Gynaecol Obstet* 2003;80:235-8.
- [50] Jacobs I, Oram D, Fairbanks J, Turner J, Frost C, Grudzinskas JG. A risk of malignancy index incorporating CA 125, ultrasound and menopausal status for the accurate preoperative diagnosis of ovarian cancer. *Br J Obstet Gynaecol* 1990;97:922-9.
- [51] Tingulstad S, Hagen B, Skjeldestad FE, et al. Evaluation of a risk of malignancy index based on serum CA125, ultrasound findings and menopausal status in the pre-operative diagnosis of pelvic masses. *Br J Obstet Gynaecol* 1996;103:826-31.
- [52] Hogdall E. Cancer antigen 125 and prognosis. *Curr Opin Obstet Gynecol* 2008;20:4-8.
- [53] Meyer T, Rustin GJ. Role of tumour markers in monitoring epithelial ovarian cancer. *Br J Cancer* 2000;82:1535-8.
- [54] Gadducci A, Cosio S, Tana R, Genazzani AR. Serum and tissue biomarkers as predictive and prognostic variables in epithelial ovarian cancer. *Crit Rev Oncol Hematol* 2009;69:12-27.
- [55] Sölétormos G, Duffy MJ, Othman Abu HS, et al. Clinical Use of Cancer Biomarkers in Epithelial Ovarian Cancer: Updated Guidelines From the European Group on Tumor Markers. *Int J Gynecol Cancer* 2016;26:43-51.
- [56] Morgante G, la MA, Ditto A, De L, V. Comparison of two malignancy risk indices based on serum CA125, ultrasound score and menopausal status in the diagnosis of ovarian masses. *Br J Obstet Gynaecol* 1999;106:524-7.
- [57] Aslam N, Tailor A, Lawton F, Carr J, Savvas M, Jurkovic D. Prospective evaluation of three different models for the pre-operative diagnosis of ovarian cancer. *BJOG* 2000;107:1347-53.
- [58] Therasse P, Arbuck SG, Eisenhauer EA, et al. New guidelines to evaluate the response to treatment in solid tumors. European Organization for Research and Treatment of Cancer, National Cancer Institute of the United States, National Cancer Institute of Canada. *J Natl Cancer Inst* 2000;92:205-16.

- [59] Eisenhauer EA, Therasse P, Bogaerts J, et al. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). *Eur J Cancer* 2009;45:228-47.
- [60] Rustin GJ, Vergote I, Eisenhauer E, et al. Definitions for response and progression in ovarian cancer clinical trials incorporating RECIST 1.1 and CA 125 agreed by the Gynecological Cancer Intergroup (GCIg). *Int J Gynecol Cancer* 2011;21:419-23.
- [61] Partovi S, Kohan A, Rubbert C, et al. Clinical oncologic applications of PET/MRI: a new horizon. *Am J Nucl Med Mol Imaging* 2014;4:202-12.
- [62] Escudero JM, Auge JM, Filella X, Torne A, Pahisa J, Molina R. Comparison of serum human epididymis protein 4 with cancer antigen 125 as a tumor marker in patients with malignant and nonmalignant diseases. *Clin Chem* 2011;57:1534-44.
- [63] Drapkin R, von Horsten HH, Lin Y, et al. Human epididymis protein 4 (HE4) is a secreted glycoprotein that is overexpressed by serous and endometrioid ovarian carcinomas. *Cancer Res* 2005;65:2162-9.
- [64] Jacob F, Meier M, Caduff R, et al. No benefit from combining HE4 and CA125 as ovarian tumor markers in a clinical setting. *Gynecol Oncol* 2011;121:487-91.
- [65] Karlsen MA, Sandhu N, Hogdall C, et al. Evaluation of HE4, CA125, risk of ovarian malignancy algorithm (ROMA) and risk of malignancy index (RMI) as diagnostic tools of epithelial ovarian cancer in patients with a pelvic mass. *Gynecol Oncol* 2012;127:379-83.
- [66] Van GT, Cadron I, Despierre E, et al. HE4 and CA125 as a diagnostic test in ovarian cancer: prospective validation of the Risk of Ovarian Malignancy Algorithm. *Br J Cancer* 2011;104:863-70.
- [67] Bast RC, Jr., Klug TL, Schaetzl E, et al. Monitoring human ovarian carcinoma with a combination of CA 125, CA 19-9, and carcinoembryonic antigen. *Am J Obstet Gynecol* 1984;149:553-9.
- [68] Krebs HB, Goplerud DR, Kilpatrick SJ, Myers MB, Hunt A. Role of Ca 125 as tumor marker in ovarian carcinoma. *Obstet Gynecol* 1986;67:473-7.
- [69] Fioretti P, Gadducci A, Ferdeghini M, et al. Combined evaluation of some tumor associated antigens in the monitoring of integrated surgical and chemotherapeutic treatment of epithelial ovarian cancer. *Eur J Gynaecol Oncol* 1986;7:200-5.
- [70] Fioretti P, Gadducci A, Ferdeghini M, Bartolini T, Bianchi R, Facchini V. Correlation of CA125 and CA19-9 serum levels with clinical course and second-look findings in patients with ovarian carcinoma. *Gynecol Oncol* 1987;28:278-83.
- [71] Alvarez RD, To A, Boots LR, et al. CA125 as a serum marker for poor prognosis in ovarian malignancies. *Gynecol Oncol* 1987;26:284-9.
- [72] Vergote IB, Bormer OP, Abeler VM. Evaluation of serum CA 125 levels in the monitoring of ovarian cancer. *Am J Obstet Gynecol* 1987;157:88-92.
- [73] Panza N, Pacilio G, Campanella L, et al. Cancer antigen 125, tissue polypeptide antigen, carcinoembryonic antigen, and beta-chain human chorionic gonadotropin as serum markers of epithelial ovarian carcinoma. *Cancer* 1988;61:76-83.
- [74] Altaras MM, Goldberg GL, Levin W, Bloch B, Darge L, Smith JA. The role of cancer antigen 125 (CA 125) in the management of ovarian epithelial carcinomas. *Gynecol Oncol* 1988;30:26-34.
- [75] Gadducci A, Ferdeghini M, Ceccarini T, et al. A comparative evaluation of the ability of serum CA 125, CA 19-9, CA 15-3, CA 50, CA 72-4 and TATI assays in reflecting the course of disease in patients with ovarian carcinoma. *Eur J Gynaecol Oncol* 1990;11:127-33.
- [76] Hording U, Toftager-Larsen K, Dreisler A, et al. CA 125, placental alkaline phosphatase, and tissue polypeptide antigen in the monitoring of ovarian carcinoma. A comparative study of three different tumor markers. *Gynecol Obstet Invest* 1990;30:178-83.
- [77] Cruickshank DJ, Terry PB, Fullerton WT. CA125-response assessment in epithelial ovarian cancer. *Int J Cancer* 1992;51:58-61.
- [78] Fioretti P, Gadducci A, Ferdeghini M, et al. The concomitant determination of different serum tumor markers in epithelial ovarian cancer: relevance for monitoring the response to chemotherapy and follow-up of patients. *Gynecol Oncol* 1992;44:155-60.
- [79] Vergote IB, Abeler VM, Bormer OP, Stigbrand T, Trope C, Nustad K. CA125 and placental alkaline phosphatase as serum tumor markers in epithelial ovarian carcinoma. *Tumour Biol* 1992;13:168-74.
- [80] Rustin GJ, Nelstrop A, Stilwell J, Lambert HE. Savings obtained by CA-125 measurements during therapy for ovarian carcinoma. The North Thames Ovary Group. *Eur J Cancer* 1992;28:79-82.
- [81] Rustin GJ, van der Burg ME, Berek JS. Advanced ovarian cancer. Tumor markers. *Ann Oncol* 1993;4 Suppl 4:71-7.

- [82] Rustin GJ, Nelstrop AE, Tuxen MK, Lambert HE. Defining progression of ovarian carcinoma during follow-up according to CA 125: a North Thames Ovary Group Study. *Ann Oncol* 1996;7:361-4.
- [83] Rustin GJ, Marples M, Nelstrop AE, Mahmoudi M, Meyer T. Use of CA-125 to define progression of ovarian cancer in patients with persistently elevated levels. *J Clin Oncol* 2001;19:4054-7.
- [84] Tuxen MK, Sölétormos G, Dombernowsky P. Serum tumour marker CA 125 in monitoring of ovarian cancer during first-line chemotherapy. *Br J Cancer* 2001;84:1301-7.
- [85] Tuxen MK, Sölétormos G, Dombernowsky P. Serum tumor marker CA 125 for monitoring ovarian cancer during follow-up. *Scand J Clin Lab Invest* 2002;62:177-88.
- [86] Liu PY, Alberts DS, Monk BJ, Brady M, Moon J, Markman M. An early signal of CA-125 progression for ovarian cancer patients receiving maintenance treatment after complete clinical response to primary therapy. *J Clin Oncol* 2007;25:3615-20.
- [87] Gadducci A, Ferdeghini M, Rispoli G, Prontera C, Bianchi R, Fioretti P. Comparison of tumor-associated trypsin inhibitor (TATI) with CA125 as a marker for diagnosis and monitoring of epithelial ovarian cancer. *Scand J Clin Lab Invest Suppl* 1991;207:19-24.
- [88] Rustin GJ. Tumour markers for ovarian cancer. *Eur J Cancer* 1992;28:2-3.
- [89] Rustin GJ, Nelstrop AE, Bentzen SM, Piccart MJ, Bertelsen K. Use of tumour markers in monitoring the course of ovarian cancer. *Ann Oncol* 1999;10 Suppl 1:21-7.
- [90] Rustin GJ, Timmers P, Nelstrop A, et al. Comparison of CA-125 and standard definitions of progression of ovarian cancer in the intergroup trial of cisplatin and paclitaxel versus cisplatin and cyclophosphamide. *J Clin Oncol* 2006;24:45-51.
- [91] Rustin GJ, van der Burg ME, Griffin CL, et al. Early versus delayed treatment of relapsed ovarian cancer (MRC OV05/EORTC 55955): a randomised trial. *Lancet* 2010;376:1155-63.
- [92] Rustin G, Tuxen M. Use of CA 125 in follow-up of ovarian cancer. *Lancet* 1996;348:191-2.
- [93] Tuxen MK, Sölétormos G, Petersen PH, Schioler V, Dombernowsky P. Assessment of biological variation and analytical imprecision of CA 125, CEA, and TPA in relation to monitoring of ovarian cancer. *Gynecol Oncol* 1999;74:12-22.
- [94] Tuxen MK, Sölétormos G, Rustin GJ, Nelstrop AE, Dombernowsky P. Biological variation and analytical imprecision of CA 125 in patients with ovarian cancer. *Scand J Clin Lab Invest* 2000;60:713-21.
- [95] Tuxen MK, Sölétormos G, Petersen PH, Dombernowsky P. Interpretation of sequential measurements of cancer antigen 125 (CA 125), carcinoembryonic antigen (CEA), and tissue polypeptide antigen (TPA) based on analytical imprecision and biological variation in the monitoring of ovarian cancer. *Clin Chem Lab Med* 2001;39:531-8.
- [96] Fraser CG. Biological variation: From principles to practice. Washington DC: AACCC Press; 2001.
- [97] Sölétormos G, Hyltoft PP, Dombernowsky P. Progression criteria for cancer antigen 15.3 and carcinoembryonic antigen in metastatic breast cancer compared by computer simulation of marker data. *Clin Chem* 2000;46:939-49.
- [98] Sölétormos G, Schioler V. Description of a computer program to assess cancer antigen 15.3, carcinoembryonic antigen, and tissue polypeptide antigen information during monitoring of metastatic breast cancer. *Clin Chem* 2000;46:1106-13.
- [99] Sölétormos G, Petersen PH, Nielsen D. Computer-simulated tumor-marker data used to compare progression criteria for cytokeratin tissue polypeptide antigen in metastatic breast cancer. *Clin Chem* 2001;47:2035-7.
- [100] Petersen PH, Sölétormos G, Pedersen MF, Lund F. Interpretation of increments in serial tumour biomarker concentrations depends on the distance of the baseline concentration from the cut-off. *Clin Chem Lab Med* 2011;49:303-10.
- [101] Lund F, Petersen PH, Pedersen MF, Hassan SO, Sölétormos G. Criteria to interpret cancer biomarker increments crossing the recommended cut-off compared in a simulation model focusing on false positive signals and tumour detection time. *Clin Chim Acta* 2014;431:192-7.
- [102] Lund F, Petersen PH, Fraser CG, Sölétormos G. Calculation of limits for significant unidirectional changes in two or more serial results of a biomarker based on a computer simulation model. *Ann Clin Biochem* 2014.
- [103] Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *J Clin Epidemiol* 2009;62:1006-12.

- [104] Moher D, Shamseer L, Clarke M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Syst Rev* 2015;4:1.
- [105] Whiting PF, Rutjes AW, Westwood ME, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Ann Intern Med* 2011;155:529-36.
- [106] Linnet K, Bossuyt PM, Moons KG, Reitsma JB. Quantifying the accuracy of a diagnostic test or marker. *Clin Chem* 2012;58:1292-301.
- [107] Abu Hassan SO, Petersen PH, Lund F, Nielsen DL, Tuxen MK, Sölétormos G. Monitoring performance of progression assessment criteria for cancer antigen 125 among patients with ovarian cancer compared by computer simulation. *Biomark Med* 2015;9:911-22.
- [108] Han LY, Karavasilis V, Hagen T, et al. Doubling time of serum CA125 is an independent prognostic factor for survival in patients with ovarian cancer relapsing after first-line chemotherapy. *Eur J Cancer* 2010;46:1359-64.
- [109] Fokkema MR, Herrmann Z, Muskiet FA, Moecks J. Reference change values for brain natriuretic peptides revisited. *Clin Chem* 2006;52:1602-3.
- [110] Limpert E, Stahel WA. Problems with using the normal distribution--and ways to improve quality and efficiency of data analysis. *PLoS One* 2011;6:e21403.
- [111] Bossuyt PM, Reitsma JB, Bruns DE, et al. The STARD statement for reporting studies of diagnostic accuracy: explanation and elaboration. The Standards for Reporting of Diagnostic Accuracy Group. *Croat Med J* 2003;44:639-50.
- [112] Bossuyt PM, Reitsma JB, Bruns DE, et al. STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies. *BMJ* 2015;351:h5527.
- [113] Miller AB, Hoogstraten B, Staquet M, Winkler A. Reporting results of cancer treatment. *Cancer* 1981;47:207-14.
- [114] Sölétormos G, Duffy MJ, Hayes DF, et al. Design of tumor biomarker-monitoring trials: a proposal by the European Group on Tumor Markers. *Clin Chem* 2013;59:52-9.
- [115] Mathematics and Statistics. Ciba-Geigy; 1975.
- [116] Sölétormos G, Nielsen D, Schioler V, Skovsgaard T, Dombernowsky P. Tumor markers cancer antigen 15.3, carcinoembryonic antigen, and tissue polypeptide antigen for monitoring metastatic breast cancer during first-line chemotherapy and follow-up. *Clin Chem* 1996;42:564-75.
- [117] Abu Hassan SO, Nielsen DL, Tuxen MK, Bast RC, Jr., Sölétormos G. Systematic Review of Monitoring Criteria to Interpret CA125 Increments during First-Line Chemotherapy and the Subsequent Follow-Up Period among Patients with Advanced Epithelial Ovarian Cancer. *Journal of Oncology Medicine and Practice, Open Access*; 2015; 1:1.
- [118] Abu Hassan SO, Nielsen DL, Tuxen MK, Petersen PH, Sölétormos G. Performance of seven criteria to assess CA125 increments among ovarian cancer patients monitored during first-line chemotherapy and the post-therapy follow-up period. Submitted Gyencol Oncol 2016.
- [119] Diaz-Gil D, Fintelmann FJ, Molaei S, Elmi A, Hedgire SS, Harisinghani MG. Prediction of 5-year survival in advanced-stage ovarian cancer patients based on computed tomography peritoneal carcinomatosis index. *Abdom Radiol (NY)* 2016. Jun 23 [Epub ahead of print].
- [120] Risum S, Hogdall C, Loft A, et al. The diagnostic value of PET/CT for primary ovarian cancer--a prospective study. *Gynecol Oncol* 2007;105:145-9.
- [121] Chandrashekhara SH, Thulkar S, Srivastava DN, et al. Pre-operative evaluation of peritoneal deposits using multidetector computed tomography in ovarian cancer. *Br J Radiol* 2011;84:38-43.
- [122] De IP, Musto A, Orazi L, et al. FDG-PET/CT in advanced ovarian cancer staging: value and pitfalls in detecting lesions in different abdominal and pelvic quadrants compared with laparoscopy. *Eur J Radiol* 2011;80:e98-103.
- [123] Pearce MS, Salotti JA, Little MP, et al. Radiation exposure from CT scans in childhood and subsequent risk of leukaemia and brain tumours: a retrospective cohort study. *Lancet* 2012;380:499-505.
- [124] Friedlander M, Trimble E, Tinker A, et al. Clinical trials in recurrent ovarian cancer. *Int J Gynecol Cancer* 2011;21:771-5.
- [125] Hall M, Gourley C, McNeish I, et al. Targeted anti-vascular therapies for ovarian cancer: current evidence. *Br J Cancer* 2013;108:250-8.
- [126] Kaye SB, Poole CJ, Danska-Bidzinska A, et al. A randomized phase II study evaluating the combination of carboplatin-based chemotherapy with pertuzumab versus carboplatin-based therapy alone in patients with relapsed, platinum-sensitive ovarian cancer. *Ann Oncol* 2013;24:145-52.
- [127] Ledermann JA, Hackshaw A, Kaye S, et al. Randomized phase II placebo-controlled trial of maintenance therapy using the oral triple angiokinase inhibitor BIBF 1120

- after chemotherapy for relapsed ovarian cancer. *J Clin Oncol* 2011;29:3798-804.
- [128] Pujade-Lauraine E, Wagner U, Aavall-Lundqvist E, et al. Pegylated liposomal Doxorubicin and Carboplatin compared with Paclitaxel and Carboplatin for patients with platinum-sensitive ovarian cancer in late relapse. *J Clin Oncol* 2010;28:3323-9.
- [129] Cognetti F, Bagnato A, Colombo N, et al. A Phase II, randomized, double-blind study of zibotentan (ZD4054) in combination with carboplatin/paclitaxel versus placebo in combination with carboplatin/paclitaxel in patients with advanced ovarian cancer sensitive to platinum-based chemotherapy (AGO-OVAR 2.14). *Gynecol Oncol* 2013;130:31-7.
- [130] Barber EL, Zsiros E, Lurain JR, Rademaker A, Schink JC, Neubauer NL. The combination of intravenous bevacizumab and metronomic oral cyclophosphamide is an effective regimen for platinum-resistant recurrent ovarian cancer. *J Gynecol Oncol* 2013;24:258-64.
- [131] Alexandre J, Brown C, Coeffic D, et al. CA-125 can be part of the tumour evaluation criteria in ovarian cancer trials: experience of the GCIg CALYPSO trial. *Br J Cancer* 2012;106:633-7.
- [132] Levy T, Weiser R, Boaz M, Ben SE, Golan A, Menczer J. The significance of the pattern of serum CA125 level ascent to above the normal range in epithelial ovarian, primary peritoneal and tubal carcinoma patients. *Gynecol Oncol* 2013;129:165-8.
- [133] Lindemann K, Kristensen G, Mirza MR, et al. Poor concordance between CA-125 and RECIST at the time of disease progression in patients with platinum-resistant ovarian cancer: analysis of the AURELIA trial. *Ann Oncol* 2016;27:1505-10.
- [134] Prat A, Parera M, Adamo B, et al. Risk of recurrence during follow-up for optimally treated advanced epithelial ovarian cancer (EOC) with a low-level increase of serum CA-125 levels. *Ann Oncol* 2009;20:294-7.
- [135] Levy T, Weiser R, Boaz M, Shem EB, Golan A, Menczer J. Prognostic significance of rising serum CA-125 levels within the normal range in patients with epithelial ovarian, primary peritoneal, and tubal cancers, who, after initial treatment, had a complete clinical response. *Int J Gynecol Cancer* 2012;22:1344-8.
- [136] Bast RC, Jr. CA 125 and the detection of recurrent ovarian cancer: a reasonably accurate biomarker for a difficult disease. *Cancer* 2010;116:2850-3.
- [137] Pignata S, Cannella L, Leopardo D, Bruni GS, Facchini G, Pisano C. Follow-up with CA125 after primary therapy of advanced ovarian cancer: in favor of continuing to prescribe CA125 during follow-up. *Ann Oncol* 2011;22 Suppl 8:viii40-viii44.
- [138] Verheijen RH, Cibula D, Zola P, Reed N. Cancer Antigen 125: Lost to Follow-Up?: A European Society of Gynaecological Oncology Consensus Statement. *Int J Gynecol Cancer* 2012;22:170-4.
- [139] Moore RG, Miller MC, Steinhoff MM, et al. Serum HE4 levels are less frequently elevated than CA125 in women with benign gynecologic disorders. *Am J Obstet Gynecol* 2012;206:351-8.
- [140] Hallamaa M, Suvitie P, Huhtinen K, Matomaki J, Poutanen M, Perheentupa A. Serum HE4 concentration is not dependent on menstrual cycle or hormonal treatment among endometriosis patients and healthy premenopausal women. *Gynecol Oncol* 2012; 125(3):667-72.
- [141] Moore RG, McMeekin DS, Brown AK, et al. A novel multiple marker bioassay utilizing HE4 and CA125 for the prediction of ovarian cancer in patients with a pelvic mass. *Gynecol Oncol* 2009;112:40-6.
- [142] Bailey J, Taylor A, Naik R, et al. Risk of malignancy index for referral of ovarian cancer cases to a tertiary center: does it identify the correct cases? *Int J Gynecol Cancer* 2006;16 Suppl 1:30-4.
- [143] Montagnana M, Danese E, Ruzzenente O, et al. The ROMA (Risk of Ovarian Malignancy Algorithm) for estimating the risk of epithelial ovarian cancer in women presenting with pelvic mass: is it really useful? *Clin Chem Lab Med* 2011;49:521-5.
- [144] Karlsen MA, Hogdall EV, Christensen IJ, et al. A novel diagnostic index combining HE4, CA125 and age may improve triage of women with suspected ovarian cancer - An international multicenter study in women with an ovarian mass. *Gynecol Oncol* 2015;138:640-6.
- [145] Cho HY, Kyung MS. Serum CA19-9 as a predictor of malignancy in primary ovarian mucinous tumors: a matched case-control study. *Med Sci Monit* 2014;20:1334-9.
- [146] Leung F, Bernardini MQ, Brown MD, et al. Validation of a Novel Biomarker Panel for the Detection of Ovarian Cancer. *Cancer Epidemiol Biomarkers Prev* 2016;25:1333-40.
- [147] Simmons AR, Clarke CH, Badgwell DB, et al. Validation of a Biomarker Panel and Longitudinal Biomarker Performance for Early Detection of Ovarian Cancer. *Int J Gynecol Cancer* 2016;26:1070-7.