# Base rate of ovarian cancer on algorithms in patients with a pelvic mass

Anne Lone Denny Rolfsen,[1] Alv A Dahl,[2,3] Are Hugo Pripp,[4] Anne Dørum [ID] [5]

**OPEN ACCESS**

Check for updates

## HIGHLIGHTS

- Comparison of three algorithms on the same pelvic tumor sample with 74% of ovarian cancer had poor case finding abilities.
- The base rate of ovarian cancer in the samples tested is essential for the predictive ability of algorithms.
- 'Base rate neglect' in ovarian cancer algorithms results in imprecise case finding ability and limits their clinical value.

## ABSTRACT

**Objective** Algorithms have been developed to identify ovarian cancer in women with a pelvic mass. The aim of this study was to determine how the base rates of ovarian cancer influence the case finding abilities of recently developed algorithms applicable to pelvic tumors. We used three ovarian cancer algorithms and the principle of Bayes' theorem for risk estimation.

**Methods** First, we evaluated the case finding abilities of the Risk of Malignancy Algorithm, the Rajavithi–Ovarian Predictive Score, and the Copenhagen Index in a prospectively collected sample at Oslo University Hospital of 227 postmenopausal women with a 74% base rate of ovarian cancer. Second, we examined the case finding abilities of the Risk of Malignancy Algorithm in three published studies with different base rates of ovarian cancer. We applied Bayes' theorem in these examinations.

**Results** In the Oslo sample, all three algorithms functioned poorly as case finders for ovarian cancer. When the base rate changed from 8.2% to 43.8% in the three studies using the Risk of Malignancy Algorithm, the proportion of false negative ovarian cancer diagnoses increased from 1.2% to 3.4%, and the number of false positive diagnosis increased from 4.6% to 14.2%.

**Conclusion** This study demonstrated that the base rate of ovarian cancer in the samples tested was important for the case finding abilities of algorithms.

## INTRODUCTION

Correct early identification of ovarian cancer is important for referral of patients to the right level of care, as the morbidity and survival of ovarian cancer patients is highly dependent on the primary surgical and oncological treatment. Most gynecologists use an algorithm to confirm a clinical suspicion of ovarian cancer. The national Norwegian guideline recommends the use of algorithms to select the patient to the right level of care (local hospital or gynecologic cancer center) without further delay. The Risk of Malignancy Index described by Jacobs et al[1] is the most frequently used algorithm for ovarian cancer identification, and is currently used in Norway to identify patients who should be referred to oncological centers for treatment.[2] To improve the case finding property of the Risk of Malignancy Index, several new ovarian cancer algorithms include the promising tumor marker, human epididymis protein 4 (HE4), which is not covered by the Risk of Malignancy Index.[3] In contrast with cancer antigen 125 (CA125), HE4 is not elevated in many common benign gynecologic and medical conditions. In addition, HE4 is elevated in >50% of tumors that do not express CA125. A combination of HE4 and CA125 or HE4 alone has greater sensitivity in patients with early stage ovarian cancer compared with CA125 alone.[3]

Base rate is a statistical term used to describe the percentage of a population that demonstrates some characteristics. In our case, the base rate concerns the rate of ovarian cancer in a population of women with pelvic tumors in the general population or at a certain level in the healthcare system. There are no standard definitions of high, intermediate, or low base rate, only relative ones. In gynecological private practices, the base rate of ovarian cancer among women with pelvic tumors is low compared with the base rate at a university clinic for gynecological cancer.

Most of the new algorithms have been developed at university centers with high relative base rates of ovarian cancer. The predictive probability of these algorithms is regularly tested by their sensitivity and specificity without consideration of the base rate (point prevalence) of ovarian cancer in the samples tested. The algorithms and samples shown include both premenopausal and postmenopausal patients (Table 1). The relevance of considering the base rate for diagnostic accuracy is explained by Bayes' theorem, which is an equation estimating the probability of an event based on the prior knowledge of conditions that are related to that event. Accordingly, the predictive probability of an ovarian cancer algorithm depends on the base rate of ovarian cancer since the specificity and sensitivity vary with the base rate.

The aims of this study were twofold: (1) to examine the predictive power of three new algorithms in a

**Table 1** Overview of modern algorithms with characteristics

| Algorithm* | Components† | Sample‡ | Base rate (%) | Sensitivity (%)§/ specificity (%)¶ |
|---|---|---|---|---|
| ROMA, 2009 USA[7] | CA125, HE4, menopause | 352 Benign 179 OC | 33.7 | 74.7/92.3 |
| CPH-I, 2015 International[9] | CA125, HE4, age | 809 Benign 246 OC | 23.3 | 95.0/78.4 |
| R-OPS, 2016 Thailand[10] | CA125, HE4, menopause, ultrasound | 158 Benign 102 OC | 39.2 | 93.9/79.9 |

Benign=benign tumors. OC=Ovarian cancer incl borderline tumors. For CPH-I and R-OPS, the development samples were used.
*Algorithm abbreviation, year of publication, country, and reference.
†Components of the algorithm.
‡Samples of pelvic masses.
§Sensitivity of ovarian cancer identified according to the algorithm.
¶Specificity of benign tumors identified according to the algorithm.
CA125, Cancer antigen 125; CPH-I, Copenhagen Index; HE4, human epididymis protein 4; OC, ovarian cancer, including borderline tumors; ROMA, Risk of Malignancy Algorithm ; R-OPS, Rajavithi–Ovarian Predictive Score.

sample with a high base rate of ovarian cancer; and (2) to examine the predictive power of the Risk of Malignancy Algorithm in three pelvic tumor samples with different base rates of ovarian cancer.

## METHODS

### Ethics

The Regional Committee for Medicine and Health related Research Ethics of South-Eastern Norway and the Protection Officer at Oslo University Hospital approved the study (project No 2013/141). Written informed consent was obtained from all patients at admission.

### Study Samples

Data Set 1

*The Oslo University Hospital Sample*

From 2012 to 2016, 465 women aged >18 years with a pelvic mass requiring surgery for a definite histologic diagnosis were prospectively recruited at Oslo University Hospital–The Norwegian Radium Hospital, which is the largest gynecological cancer center in Norway. All patients had the same diagnostic work-up. Experienced gynecologists performed the ultrasound examinations and entered clinical data and ultrasound findings in a predefined form. All patients had transvaginal ultrasound, but in the case of larger pelvic masses, transabdominal sonography was used as a supplement. Ultrasound features were registered as unilocular, multilocular, solid tumor, septa, excrescences, solid areas, cystic areas, and ascites. Serum HE4, CA125, and creatinine values were measured at the first outpatient appointment prior to surgery. Definite histologic diagnosis was obtained by surgery. All specimens were examined by a pathologist experienced in gynecological oncology. Staging was based on the International Federation of Gynecology and Obstetrics classification.[4] Our sample consisted of patients with definite ovarian/tubal/peritoneal cancer, including borderline tumors. Postmenopausal status implied no menstruation for the last 12 months before inclusion. To simplify our presentation, we only studied postmenopausal patients, since the base rate of ovarian cancer is different in premenopausal women. Our sample consisted of 227 postmenopausal women, 59 had benign tumors (26%) and 168 were classified as 'positive ovarian cancer cases' (74% base rate), which included 19 borderline ovarian tumors, as these tumors

need correct surgical staging for identification of invasive meta-static implants or upstaging to low grade ovarian cancer.[5] Of the 149 invasive ovarian/tubal/peritoneal cancer cases, 40 had stage I (14 high grade and 3 low grade serous, 6 endometroid, 6 mucinous, 2 clearcell, and 9 other histologies), 11 had stage II (6 high grade and 2 low grade serous, 2 endometroid, and 1 other histology), 88 had stage III (70 high grade and 7 low grade serous, 4 endometroid, 1 mucinous, 3 clearcell, and 3 other histologies), and 10 had stage IV (6 high grade serous, 2 endometroid, and 2 clearcell). Of the 19 cases with borderline ovarian tumors, 16 had stage I (11 serous and 5 mucinous) and 3 had stage II serous type.

Data Set 2

We examined the specificities and sensitivities of the Risk of Malignancy Algorithm in three previously published samples of postmenopausal patients with various base rates of ovarian cancer. One sample was published by Novotny et al[6] and two were published by Moore et al.[7 8] Their published base rates of ovarian cancer were 8.2%, 29.9%, and 43.8%, and we selected these samples reflecting low, intermediate, and high base rates of ovarian cancer, respectively.

### Cancer Algorithms

The Risk of Malignancy Algorithm was introduced by Moore et al in 2009 and is frequently referred to in the ovarian cancer literature.[7] The algorithm is based on CA125, HE4, and menopausal status. The goal of the Risk of Malignancy Algorithm is to stratify patients with pelvic masses into low and high risk of malignancy groups using the designated predictive probability thresholds for premenopausal and postmenopausal women. The Predictive Index (PI) has this formula for postmenopausal women: Postmenopausal $PI = -8.09 + 1.04 \times LN(HE4) + 0.732 \times LN(CA125)$ and predicted probability $(PP) = \exp (PI)/[1+\exp (PI)] \times 100$. High risk of ovarian cancer group was defined as PP >27.7%. The AUC statistic was not given. Patients were included from 12 different geographical sites, but the levels of healthcare and their base rates were unspecified (Table 1).

The Copenhagen Index algorithm, published by Karlsen et al in 2015, represented a new algorithm including three elements: measurements of CA125 and HE4, and the patient's age.[9] Age was considered as more precise than menopausal status. Their formula

was: Copenhagen Index (CPH-I)=$-14.0647 + 1.0649 \times \log2$ (HE4) $\times \pm 0.6050 \times \log2$ (CA125) $+ 0.2672 \times$ age / 10, and the predicted probability (PP) was as follows: PP=$e^{(CPH-I)}/(1+e^{(CPH-I)})$. An optimal cut-off PP of $\geq 0.070$ was established in their Danish development sample of 1055 patients (Table 1).

The Rajavithi–Ovarian Predictive Score algorithm was published by Yanaranop et al in 2016 for identification of ovarian cancer in women presenting with pelvic masses.[10] Like the Risk of Malignancy Algorithm, the Rajavithi–Ovarian Predictive Score included measurements of CA125, HE4, and menopausal status, but added a fourth element, solid findings on ultrasound examination. Their algorithm was: Rajavithi–Ovarian Predictive Score (R-OPS) = M $\times$ U $\times$ (CA125 $\times$ HE4)$^{1/2}$, where M had the value 3 for postmenopausal women, and U was coded 1 for no solid lesion and 6 for presence of solid lesion on ultrasound with an optimal cut-off value of $\geq 330$ in their development sample. The sample was collected from the Rajavithi Hospital, a tertiary university hospital in Bangkok, Thailand (Table 1).

## Statistical Analyses

### Base Rate Calculations

We calculated base rates of the samples by the formula[11]:

$$\text{Prevalence/base rate} = \frac{\text{cases with condition of interest}}{\text{cases in a sample}}$$

The formula was based on the bayesian statistical approach to clinical diagnosis posed as: what is the probability of a woman having ovarian cancer given the results of an algorithm applied to a sample of women with pelvic masses with an ovarian cancer base rate of x.[12] The theorem can be reformulated as an equation[13]:

(Odds before test) $\times$ (test odds) = (Post test odds)

The 'odds before test' is the base rate of ovarian cancer in the sample before the test, and the 'test odds' is the sensitivity and specificity of an ovarian cancer algorithm. When multiplied, they give the new odds ('post test odds') of ovarian cancer in the sample tested.

### Sensitivities and Specificities

Sensitivity measures how often a test correctly generates a positive result for people who have the condition that is being tested for (also known as the true positive rate). Specificity measures a test's ability to correctly generate a negative result for people who do not have the condition that is being tested for (also known as the true negative rate).

In data set 1, sensitivities and specificities of the Risk of Malignancy Algorithm, the Copenhagen Index, and the Rajavithi–Ovarian Predictive Score were calculated for postmenopausal women/age at survey.

In data set 2, we used the published sensitivity and specificity findings of the Risk of Malignancy Algorithm in each sample and calculated the base rate of ovarian cancer according to the formula previously described. We used these data to find the number of false positive and negative patients in a sample of 1000 patients according to Bayes' theorem. The parameters and calculations are shown in Box 1.

We also evaluated changes in the algorithms' abilities to identify false positive and false negative patients. The differences in sensitivities and specificities observed by the Risk of Malignancy Algorithm between each of the three studies were tested and compared

---

**Box 1    Postmenopausal women in a low base rate ovarian cancer sample (Novotny et al 2012)[6]**

Base rate of ovarian cancer in women with pelvic masses: 8.2%, benign masses: 91.8%
Risk of Malignancy Algorithm sensitivity 0.857 and specificity 0.95
Sensitivity of ovarian cancer: 8.2% $\times$ 0.857=7% correctly identified as cancer
8.2% $\times$ 0.143=1.2% identified as benign
Specificity of ovarian cancer: 91.8% $\times$ 0.95=87.2% correctly identified as benign
91.8% $\times$ 0.05=4.6% identified as cancer
Identified as ovarian cancer: 6.9% (correctly) +4.6% (false positives) =11.6%
Identified as benign: 87.2% (correctly) + 1.2% (false negatives) =88.4%
Among 1000 women with pelvic masses: 82 true ovarian cancers and 918 benign
Identified as ovarian cancer: 70 (correctly) + 46 (false positives) =116 patients
Identified as benign: 872 (correctly) +12 (false negatives) =884 patients

---

with Fisher's exact tests. Descriptive statistics were calculated using the IBM SPSS for PCs version 24 (IBM, Armonk, New York, USA). The p value was set at $<0.05$ and all tests were two sided.

## RESULTS

### Data Set 1

In the Oslo University Hospital sample, the Risk of Malignancy Algorithm, the Rajavithi–Ovarian Predictive Score, and the Copenhagen Index showed adequate strength for sensitivities, however, specificities were low. The proportions of false negative and positive patients were high for all three algorithms (Table 2).

### Data Set 2

In the low base rate group, the Risk of Malignancy Algorithm identified 1.2% false negative and 4.6% false positive women with ovarian cancer (Box 1 and Table 3). The Risk of Malignancy Algorithm identified 1.9% false negative and 17.2% false positive women with ovarian cancer in the intermediate base rate group. Compared with the low base rate group, the Risk of Malignancy Algorithm identified more false positives (p=0.14) and false negatives (p<0.001) in the intermediate base rate group.

In the high base rate group, the Risk of Malignancy Algorithm identified 3.4% false negative and 14.2% false positive women with ovarian cancer (Table 3). The Risk of Malignancy Algorithm algorithm identified less false positives (p=0.05) and more false negatives (p=0.07) compared with the intermediate base rate group. However, compared with the low base rate group, the Risk of Malignancy Algorithm identified more false positives (p=0.002) and more false negatives (p<0.001). The main findings of the comparisons of the case finding abilities of the Risk of Malignancy Algorithm related to various base rates are summarized in Table 3.

**Table 2** Findings of the three algorithms using the Oslo University Hospital sample

| Algorithm | Sensitivity | Specificity | False positives (%) | False negatives (%) |
|---|---|---|---|---|
| ROMA | 0.81 | 0.24 | 19.8 | 14.1 |
| R-OPS | 0.86 | 0.19 | 21.1 | 10.4 |
| CPH-I | 0.82 | 0.22 | 20.3 | 13.3 |

CPH-I, Copenhagen Index; ROMA, Risk of Malignancy Algorithm; R-OPS, Rajavithi–Ovarian Predictive Score.

## DISCUSSION

Comparison of three algorithms using data set 1 with 74% base rate of ovarian cancer showed minimal difference in their case finding properties, but the proportion of false positive and negative cases were high (Table 2). Testing of the Risk of Malignancy Algorithm in three samples with different base rates of ovarian cancer showed significant differences in case finding ability of ovarian cancer when the base rate was incorporated into the equation (Table 3). Our findings reflect the 'base rate neglect' in the development of new ovarian cancer algorithms, resulting in imprecise case finding ability and limiting their clinical value. The Risk of Malignancy Algorithm, the Rajavithi–Ovarian Predictive Score, and the Copenhagen Index are new algorithms intended to improve ovarian cancer case identification and thereby correct treatment and improve survival rates for women with ovarian cancer. The algorithms are needed by gynecologists at the primary level of care. At the tertiary level of care, the algorithms are used to select patients for admittance, however, they are of less importance within the hospital as many gynecologic oncologists are specialized in ultrasound and have access to advanced radiologic imaging in addition to their expert clinical judgment. All the algorithms were developed in samples with an ovarian cancer base rate of between 23% and 39%, and at this range they showed adequate case finding properties of ovarian cancer, but still they were recommended for primary healthcare with a substantially lower base rate of ovarian cancer. They functioned poorly in the Oslo University Hospital sample with a very high base rate, and their clinical value at very low base rates is uncertain. Regarding pelvic tumors in premenopausal women, they could function worse since the ovarian cancer base rate is lower than in postmenopausal women. In Norway and in many other countries, a base rate of ovarian cancer <5% is expected in unselected population samples.

We have demonstrated that changes in the base rate of ovarian cancer imply significant differences in the number of false positive and false negative cases predicted by recommended algorithms. False positive results lead to unnecessary referral to gynecologic cancer centers, mental distress in patients, and surgery with benign findings. False negative results could reduce survival due to surgery at non-specialized hospitals or delayed referral to a cancer center or worse, no explorative surgery. A strength of our study was the utilization of Bayes' theorem. The theorem is crucial for the development of algorithms in all fields of medicine, and the knowledge and understanding of the theorem is therefore highly important to both clinicians and biostatisticians. It is also a strength that we proved our thesis in both our own clinical sample and in publicly available data from previous studies. A limitation of this study is the selection of postmenopausal women, however, that selection was made to simplify our base rate documentation. The performance of the newly emerged ovarian cancer algorithms in the primary care setting is yet to be determined in unselected samples. HE4 is not a part of the Risk of Malignancy Index, and on this basis we suggest studies in primary care to assess the case finding ability of new ovarian cancer algorithms. The findings should be compared with the case finding ability of the already implemented Risk of Malignancy Index. In general, clinicians in any field of medicine should consider the base rate of disease when reviewing diagnostic algorithms, and then indicate at what base rate they have adequate case finding abilities.

**Author affiliations**
[1]Gynecology and Obstetrics, Oslo University Hospital, Oslo, Norway
[2]National Advisory Center for Late Effects after Cancer Treatment, Oslo University Hospital, The Norwegian Radium Hospital, Oslo, Norway
[3]Faculty of Medicine, University of Oslo, Oslo, Norway
[4]Oslo Center for Biostatistics and Epidemiology, Oslo University Hospital, Oslo, Norway
[5]Gynecologic Oncology, Oslo University Hospital, The Norwegian Radiumhospital, Oslo, Norway

**Twitter** Anne Lone Denny Rolfsen @RolfsenAnne

**Table 3** Summary of the Risk of Malignancy Algorithm findings in three samples with different base rates of ovarian cancer

| Variable | Novotny (2012) | Moore (2014) | Moore (2009) |
|---|---|---|---|
| Base rate of ovarian cancer (%) | 8.2 | 29.9 | 43.8 |
| Benign cases (%) | 91.8 | 70.1 | 56.2 |
| Sensitivity | 0.86 | 0.94 | 0.92 |
| Specificity | 0.95 | 0.76 | 0.75 |
| False positives (%) | 4.6 | 17.2 | 14.2 |
| False negatives (%) | 1.2 | 1.9 | 3.4 |

**ORCID iD**

Anne Dørum http://orcid.org/0000-0002-6274-2398

## REFERENCES

1  Jacobs I, Oram D, Fairbanks J, *et al*. 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.
2  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. *BJOG* 1996;103:826–31
3  Dochez V, Caillon H, Vaucel E, *et al*. Biomarkers and algorithms for diagnosis of ovarian cancer: CA125, HE4, RMI and Roma, a review. *J Ovarian Res* 2019;12:28.
4  FIGO. Ovarian cancer staging 2014. Available: https://www.sgo.org/wp-content/uploads/2012/09/FIGO-Ovarian-Cancer-Staging_1.10.14.pdf [Accessed 13 Mar 2020].
5  Trillsch F, Mahner S, Vettorazzi E, *et al*. Surgical staging and prognosis in serous borderline ovarian tumours (BOT): a subanalysis of the ago ROBOT study. *Br J Cancer* 2015;112:660–6.
6  Novotny Z, Presl J, Kucera R, *et al*. He4 and Roma index in Czech postmenopausal women. *Anticancer Res* 2012;32:4137–40.
7  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.
8  Moore RG, Hawkins DM, Miller MC, *et al*. Combining clinical assessment and the risk of ovarian malignancy algorithm for the prediction of ovarian cancer. *Gynecol Oncol* 2014;135:547–51.
9  Karlsen MA, Høgdall EVS, 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.
10  Yanaranop M, Tiyayon J, Siricharoenthai S, *et al*. Rajavithi-ovarian cancer predictive score (R-OPS): a new scoring system for predicting ovarian malignancy in women presenting with a pelvic mass. *Gynecol Oncol* 2016;141:479–84.
11  Altman DG. *Practical statistics for medical research*. Boca Raton, FL: Chapman & Hall/CRC, 1991.
12  Gill CJ, Sabin L, Schmid CH. Why clinicians are natural bayesians. *BMJ* 2005;330:1080–3.
13  Brakedal B. En kliniker OG en bayesianer (a clinician and a Bayesian) [in Norwegian]. *Tidsskr Nor Legeforen* 2015;135:1468–70.