

RESEARCH

Identifying women with suspected ovarian cancer in primary care: derivation and validation of algorithm

 OPEN ACCESS

Julia Hippisley-Cox *professor of clinical epidemiology and general practice*, Carol Coupland *associate professor in medical statistics*

Division of Primary Care, University of Nottingham, Nottingham NG2 7RD, UK

Abstract

Objective To derive and validate an algorithm to estimate the absolute risk of having ovarian cancer in women with and without symptoms.

Design Cohort study with data from 375 UK QResearch general practices for development and 189 for validation.

Participants Women aged 30-84 without a diagnosis of ovarian cancer at baseline and without appetite loss, weight loss, abdominal pain, abdominal distension, rectal bleeding, or postmenopausal bleeding recorded in previous 12 months.

Main outcome The primary outcome was incident diagnosis of ovarian cancer recorded in the next two years.

Methods Risk factors examined included age, family history of ovarian cancer, previous cancers other than ovarian, body mass index (BMI), smoking, alcohol, deprivation, loss of appetite, weight loss, abdominal pain, abdominal distension, rectal bleeding, postmenopausal bleeding, urinary frequency, diarrhoea, constipation, tiredness, and anaemia. Cox proportional hazards models were used to develop the risk equation. Measures of calibration and discrimination assessed performance in the validation cohort.

Results In the derivation cohort there were 976 incident cases of ovarian cancer from 2.03 million person years. Independent predictors were age, family history of ovarian cancer (9.8-fold higher risk), anaemia (2.3-fold higher), abdominal pain (sevenfold higher), abdominal distension (23-fold higher), rectal bleeding (twofold higher), postmenopausal bleeding (6.6-fold higher), appetite loss (5.2-fold higher), and weight loss (twofold higher). On validation, the algorithm explained 57.6% of the variation. The receiver operating characteristics curve (ROC) statistic was 0.84, and the D statistic was 2.38. The 10% of women with the highest predicted risks contained 63% of all ovarian cancers diagnosed over the next two years.

Conclusion The algorithm has good discrimination and calibration and, after independent validation in an external cohort, could potentially be used to identify those at highest risk of ovarian cancer to facilitate early referral and investigation. Further research is needed to assess how best to implement the algorithm, its cost effectiveness, and whether, on implementation, it has any impact on health outcomes.

Introduction

Ovarian cancer is the seventh most common cancer in women worldwide, affecting 225 000 new patients each year.¹ Of these, about 6700 women are in the United Kingdom, giving the UK one of the highest rates in Europe.² Most women are diagnosed with stage III or stage IV cancer, for which the five year survival is 20% and 6%, respectively.³ Less than 30% of women are diagnosed with stage I ovarian cancer, and, of these, 90% will survive to five years. While ovarian cancer is the leading cause of death in the UK from gynaecological malignancies, there have been improvements in survival in the past two decades, which might reflect earlier diagnosis and more effective treatments.² In general terms, the earlier the cancer is diagnosed, the more treatment options are available and the better the prognosis.

As there are few established risk factors, targeted screening of asymptomatic patients at risk of developing ovarian cancer is unlikely to be cost effective at present (although further information is likely to become available when the UK ovarian cancer screening trial reports in 2015-6). The challenge presented by ovarian cancer, therefore, is to make the correct diagnosis as early as possible, despite the non-specific nature of symptoms and signs.⁴ This is particularly the case in primary care, where general practitioners need to differentiate those patients for whom further investigation is warranted from those who require reassurance or a "watch and wait" policy. Moreover, primary care clinicians need to decide which patients require urgent investigation or referral and which require routine tests or referral. Earlier diagnosis, however, could improve with more targeted investigation of symptomatic patients^{5 6} and increased public awareness of symptoms as encouraged by the National Awareness and Early Diagnosis Initiative (NAEDI).⁷ It has been estimated that 10% of deaths from ovarian cancers might be avoidable.⁸ Other guidelines and policies aim to increase access to diagnostic investigations for general practitioners, and tools to help assess absolute risk of different types of cancer are needed to help ensure the right patients are investigated as well as to optimise the use of scarce resources including abdominal

and transvaginal ultrasonography, computed tomography, or magnetic resonance imaging. For ovarian cancer, the current guidance from the National Institute for Health and Clinical Excellence² encourages the use of blood tests to measure CA125 concentration for symptomatic women as a prelude to ultrasound scanning, although this has not been validated in a primary care setting. CA125 concentration is raised in half the women who have early stage ovarian cancer and 90% of those with more advanced disease.³

We developed and validated a risk prediction algorithm to estimate the individualised absolute risk of having ovarian cancer, incorporating both symptoms and other risk factors, to help identify those at highest risk for further investigation or referral. We used QResearch (a large UK primary care database) to develop the risk prediction models as it contains robust data on many of the relevant exposures and outcomes. It is also representative of the population in which such a model is likely to be used and has been used successfully to develop and validate a range of prognostic models for use in primary care⁹⁻¹⁰ as well as models designed to help earlier detection of other cancers.¹¹⁻¹²

Methods

Study design and data source

We did a prospective cohort study in a large population of primary care patients from an open cohort study using the QResearch database (version 30). We included all practices in England and Wales who had been using their EMIS (Egton Medical Information System) computer system for at least a year. We randomly allocated two thirds of practices to the derivation dataset and the remaining third to a validation dataset. We identified an open cohort of women aged 30-84 drawn from patients registered with practices between 1 January 2000 and 30 September 2010. We excluded patients without a postcode related Townsend score, patients with a history of bilateral oophorectomy or ovarian cancer, and those with a recorded "red flag symptom" in the 12 months before the study entry date—that is, symptoms of loss of appetite, weight loss, abdominal pain, abdominal distension, rectal bleeding, or postmenopausal bleeding—that might indicate ovarian cancer. Entry to the cohort was the latest of study start date (1 January 2000), 12 months after the patient registered with the practice, and, for those patients with one or more red flag symptom, the date of first recorded onset within the study period. When patients had new onset of multiple symptoms recorded, the entry date was the earliest recorded date of the new symptom in the study period. Other symptoms were included if they occurred within 60 days of the entry date and before the diagnosis of ovarian cancer or the date on which the patient left, died, or the study ended.

Clinical outcome definition

Our outcome was ovarian cancer, which we defined as incident diagnosis of ovarian cancer during the two years after study entry recorded either in the patient's GP record using the relevant UK diagnostic Read codes or on their linked Office for National Statistics (ONS) cause of death record with the relevant ICD-9 (international classification of diseases, ninth revision) codes (183) or ICD-10 (10th revision) diagnostic codes (C56). The ONS data are currently linked deterministically within the NHS clinical computer system with NHS number, postcode, date of birth, and date of death. We used a two year period as this represents the period of time during which existing cancers are likely to become clinically manifest.¹³⁻¹⁴ We assumed that

when deaths from ovarian cancer occurred within two years, without a recorded diagnostic code in the GP record, the cancer would have been present at the start of the two year period.

Predictor variables

We examined established predictor variables, focusing on those that are likely to be recorded in the patient's electronic record and that the patient is likely to know. We also included symptoms that might herald a diagnosis of ovarian cancer based on recent studies.⁶⁻¹⁵ We included both chronic risk factors (such as age and family history) and symptoms to determine the absolute risk of ovarian cancer. The predictor variables examined were:

- Currently consulting general practitioner with first onset of loss of appetite (yes/no)
- Currently consulting general practitioner with first onset of weight loss symptom (yes/no)
- Currently consulting general practitioner with first onset of abdominal pain (yes/no)
- Currently consulting general practitioner with first onset of abdominal distension (yes/no)
- Currently consulting general practitioner with first onset of rectal bleeding (yes/no)
- Currently consulting general practitioner with first onset of postmenopausal bleeding (yes/no)
- Recently consulted general practitioner with constipation in past 12 months (yes/no)
- Recently consulted general practitioner with diarrhoea in past 12 months (yes/no)
- Recently consulted general practitioner with tiredness in past 12 months (yes/no)
- Recently consulted general practitioner with increased urinary frequency in past 12 months (yes/no)
- Age at baseline (continuous, range 30-84)
- Body mass index (BMI) (continuous)
- Smoking status (non-smoker; ex-smoker; light (1-9 cigarettes/day); moderate (10-19 cigarettes/day); heavy smoker (≥ 20 cigarettes/day))
- Alcohol use (none, trivial (< 1 unit/day); light (1-2 units/day); moderate or heavy (≥ 3 units/day))
- Townsend deprivation score (continuous)
- Previous diagnosis of cancer apart from ovarian cancer
- Anaemia defined as recorded haemoglobin < 110 g/L in past 12 months (yes/no).

Derivation and validation of the models

We developed and validated the risk prediction algorithm using established methods.^{9-10, 16-20} We used multiple imputation to replace missing values for BMI, alcohol use, and smoking status and used these values in our main analyses.²¹⁻²⁴ We carried out five imputations. We used Cox's proportional hazards models to estimate the coefficients for each risk factor using robust variance estimates to allow for the clustering of patients within general practices. We used Rubin's rules to combine the results across the imputed datasets.²⁵ We used fractional polynomials to model non-linear risk relations with continuous variables.²⁶ We fitted a full model initially and retained variables if they had a hazard ratio of < 0.80 or > 1.20 (for binary variables) and were significant at the 0.01 level. We examined interactions

between predictor variables and age and included them in the final models if they were significant at the 0.01 level.

We used the regression coefficients for each variable from the final model as weights, which we combined with the baseline survivor function evaluated at two years to derive absolute risk equations for two years of follow-up.²⁷ We estimated the baseline survivor function based on zero values of centred continuous variables, with all binary predictor values set to zero, using the methods implemented in Stata.

We used multiple imputation in the validation cohort to replace missing values for BMI, alcohol, and smoking. We then applied the risk equations obtained from the derivation cohort to the validation cohort and calculated measures of discrimination. We calculated R^2 (estimated variation in time to ovarian cancer²⁸), the D statistic²⁹ (a measure of discrimination where higher values indicate better discrimination), and the area under the receiver operating characteristic curve (Receiver Operating Curve statistic) at two years. We assessed calibration (comparing the mean predicted risk at two years with the observed risk by 10th of predicted risk). The observed risks were obtained by using Kaplan-Meier estimates evaluated at two years.

We used the validation cohort to define the thresholds for the 0.1%, 0.5%, 1%, 5%, and 10% of women at highest estimated risk of ovarian cancer at two years. We calculated sensitivity, specificity, and positive and negative predictive values using these thresholds, restricting the analyses to women who had the outcome within two years or had at least two years of follow-up. We used all the available data on the database to maximise the power and also generalisability of the results. We used Stata (version 11) for all analyses.

Results

Overall study population

Overall, 564 QResearch practices in England and Wales met our inclusion criteria and 375 were randomly assigned to the derivation dataset with the remainder assigned to a validation cohort. We identified 1 272 186 women aged 30-84 in the derivation cohort. We excluded 62 392 women (4.9%) without a recorded Townsend deprivation score, 13 748 (1.1%) with bilateral oophorectomy, 1330 (0.1%) with a history of ovarian cancer, and 35 993 (2.8 %) with at least one red flag symptom recorded in the 12 months before entry to the study, leaving 1 158 723 patients for analysis

We identified 672 661 women aged 30-84 in the validation cohort. We excluded 35 868 patients (5.3%) without a recorded Townsend score, 7351 (1.1%) with bilateral oophorectomy, 749 (0.1 %) with a history of ovarian cancer, and 19 831 (2.9%) with at least one red flag symptom recorded in the 12 months before study entry, leaving 608 862 patients for analysis.

The baseline characteristics of each cohort were similar (table 1). As in previous studies,^{9 16 30} the patterns of missing data supported the use of multiple imputation to replace missing values for smoking status, alcohol, and BMI (not shown, available from the authors).

Incidence of red flag symptoms

In the derivation cohort, we identified 132 576 women with incident abdominal pain, 5140 with abdominal distension, 5920 with appetite loss, 25 274 with rectal bleeding, 18 244 with postmenopausal bleeding, and 9081 with weight loss. Overall, 196 466 women (17%) had one red flag symptom, 2223 (0.2%) had two, and 33 had three or more recorded symptoms.

Incidence rates of ovarian cancer

In the derivation cohort, during the two year follow-up period we identified a total of 976 incident cases of ovarian cancer arising from 2 025 812 person years of observation, giving a crude rate of 48 per 100 000 person years. There were 853 cases (87% of 976) identified using the GP record and an additional 123 (13% of 976) identified solely from the linked death record.

In the validation cohort we identified 538 incident cases of ovarian cancer arising from 1 065 490 person years of observation giving a crude rate of 50 per 100 000 person years. There were 479 cases (89% of 538) identified with the GP record and an additional 59 (11%) solely from the linked death record.

Predictor variables

Table 2 shows the predictor variables selected for the final model. Independent predictors were age, family history of ovarian cancer (9.8-fold higher risk), anaemia (2.3-fold higher), abdominal pain (sevenfold higher), abdominal distension (23-fold higher), rectal bleeding (twofold higher), postmenopausal bleeding (6.6-fold higher), appetite loss (5.2-fold higher), and weight loss (twofold higher). The other variables examined were not independent risk factors so were not included in the final model. There were no significant interactions with age.

Validation

The validation statistics (table 3) showed that the risk prediction equation explained 57.6% of the variation in time to diagnosis. The D statistic was 2.38, and the ROC statistic was 0.84.

The figure shows the mean predicted scores and the observed risks at two years within each 10th of predicted risk to assess the calibration of the model in the validation cohort. Overall, the model was well calibrated with close correspondence between predicted and observed two year risks within each model 10th.

Individual risk assessment and thresholds

One potential use for this algorithm is within consultations with individual patients, particularly if they present with new onset of an alarm symptom such as abdominal distension, abdominal pain, weight loss, or appetite loss. The results could help inform the decision to undertake further investigations such as a CA125 blood test or abdominal ultrasonography. Some clinical examples are shown in the box.

The algorithm could also be used for systematic risk stratification for a population of patients aged 30-84. Software implementing the algorithm could calculate the risk of a patient having an existing but as yet undiagnosed ovarian cancer based on information already recorded in the patient's electronic health record. Patients at highest risk could be identified for a clinical assessment.

As this is a new algorithm, there are no established thresholds for defining high risk groups. We calculated a range of centiles of predicted risk from the validation population to define a high risk group (that is, the top 0.1%, 0.5%, 1%, 5%, and 10% at highest risk) of women. We then determined the numbers and proportion of incident cases in the validation cohort that fell within each category of risk.

The 90th centile defined a high risk group with a two year risk score of >0.2 % (table 4). There were 340 new cases of ovarian cancer within this group out of 538 new cases identified in the validation cohort, which accounted for 63.2% of all new cases

Clinical examples of algorithm for ovarian cancer

A 70 year old woman consulting with abdominal pain has an estimated risk of ovarian cancer of 0.6%. If she also has had anaemia in the past year her estimated risk is 1.4%. If she also has abdominal distension her estimated risk of ovarian cancer is 28%

A 55 year old woman with a family history of ovarian cancer and consulting with loss of appetite has a 2.6% estimated risk of ovarian cancer. If she also has abdominal distension her estimated risk is 46%. If she has loss of appetite and abdominal distension but no family history of ovarian cancer her estimated risk of ovarian cancer is 6.1%

A 40 year old woman consulting with weight loss and abdominal pain and with anaemia in the past year has an estimated risk of ovarian cancer of 0.3%. If she also has abdominal distension, her estimated risk is 7%

of ovarian cancer (sensitivity). The positive predictive value with this threshold was 0.8 %. Alternatively, a threshold based on the top 5% of risk (a two year risk score >0.5%) had a sensitivity of 42% and a positive predictive value of 1.1%. In contrast, the positive predictive value of single symptoms ranged between 0.1% for rectal bleeding to 1.8% for abdominal distension. The sensitivity of an approach based on single symptoms ranged from 2% for appetite loss to 49.4% for abdominal pain.

Discussion

Summary of key findings

We have developed and validated a new algorithm designed to estimate the absolute risk of having existing but as yet undiagnosed ovarian cancer based on a combination of symptoms and simple variables such as age and family history of ovarian cancer, which the patient is likely to know and which will increase the baseline absolute risk. The algorithm could be used to assess risk at the point of care in those patients presenting to general practitioners with these symptoms, many of which are non-specific. The algorithm does not actually result in a diagnosis of ovarian cancer, rather it can be used to identify a subset of high risk women suitable for targeted investigation.

The algorithm performed well in a separate validation sample with good discrimination and calibration. After external validation this new algorithm could potentially be used to identify those at highest risk of having ovarian cancer to facilitate early referral and investigation and so help earlier identification. Further research is needed to assess how to implement the algorithm, its cost effectiveness, and whether, on implementation, it has any impact on the stage of ovarian cancer at diagnosis and subsequent survival.

Implications for clinical guidelines

Our study is topical given the recent guidelines published by NICE in April 2011 on the recognition of ovarian cancer.² This recommends carrying out tests in primary care for women (especially those aged 50 or over) if they have any of the following symptoms particularly more than 12 times a month: abdominal distension; feeling full or loss of appetite, or both; pelvic or abdominal pain; increased urinary frequency or urgency, or both; or symptoms suggestive of irritable bowel syndrome in the past 12 months (on the basis that irritable bowel syndrome rarely presents for the first time in women aged 50 and over); or unexplained weight loss, fatigue, or changes in bowel habit. NICE guidelines recommend that women with symptoms suggestive of ovarian cancer should have a CA125 test and if the concentration is 35 IU/mL or more, an ultrasound scan of the abdomen and pelvis should be undertaken. After the scan a risk of malignancy score should be calculated, based on menopausal status, CA125 concentration, and ultrasound findings. Those with a score of 250 or more should be referred to a specialist team. NICE also acknowledges, however, that research is needed to determine the specificities and sensitivities

of the risk of malignancy score at different thresholds as well as evidence for the performance of CA125 in a primary care setting.

Our study lends some support to NICE guidelines, as we have confirmed that abdominal distension, unintentional weight loss, loss of appetite, and abdominal pain all independently predict ovarian cancer. Other symptoms such as urinary frequency, however, are mentioned in the NICE guideline but were not significant predictors in our study on multivariate analysis. Similarly, we found additional symptoms such as anaemia, postmenopausal bleeding, and rectal bleeding, which were independently predictive on multivariate analysis, that were not included in the NICE guideline. Importantly, our algorithm takes better account of age than the NICE guideline, which simply dichotomises patients into those aged under 50 or 50 and older. This is relevant as the risk of ovarian cancer increases with age. We have also quantified the risk associated with family history of ovarian cancer and incorporated it into the underlying algorithm so that it is possible to calculate a woman's absolute risk of ovarian cancer. We have provided information on the sensitivity, specificity, and positive and negative predictive powers at different thresholds of risk so that this can be used for cost effectiveness modelling, which is outside the scope of the present study. Such modelling, along with an evaluation of the performance of CA125 testing in symptomatic women in a primary care setting, has the potential to inform future revisions of the NICE guideline.

Comparison with previous studies

Our study has good face validity as the direction and magnitude of the hazard ratios and predictive value of individual symptoms in our study are comparable with those reported elsewhere.^{2 6 15} In particular, the symptom that had the largest positive predictive power in our study (abdominal distension) was also the strongest predictor in a recent study by Hamilton et al based on 39 practices in Devon over a four year period.⁶ Abdominal distension was also associated with an odds ratio of 29.2 (95% confidence interval 16.5 to 51.8) in a recent systematic review⁴ and was the symptom with the highest odds ratio. The frequency of abdominal pain in patients with ovarian cancer in our study was 49%, which is similar to that reported in a recent systematic review⁴ and that reported by Hamilton et al.⁶ Overall, this acts as a useful cross validation of both studies, which have different strengths. The Hamilton study was able to validate outcomes against histological records, which was not possible in our study. Our study, however, was much larger and nationally rather than locally based and included additional variables such as age, family history, and presence of anaemia alongside symptoms and gives a combined individualised measure of absolute risk of ovarian cancer. The inclusion of symptoms potentially extends the utility of this algorithm to the point of care consultation with a symptomatic patient as family physicians could use it to assess the patient's absolute baseline risk as well as the probable increased risk from recent onset of symptoms.

Methodological strengths

Key strengths of our study include size, duration of follow-up, representativeness, and lack of selection, recall, and respondent bias. UK general practices have good levels of accuracy and completeness in recording clinical diagnoses and prescribed drugs.³¹ We think our study has good face validity as it has been conducted in the setting in which most patients in the UK are assessed, treated, and followed up. We developed the algorithm in one cohort and validated it in a separate cohort representative of the patients likely to be considered for referral and treatment. Comparison of published discrimination statistics suggests our model performs well (our ROC value was 0.84). Lastly, the algorithm can be built into clinical systems and the results generated automatically with suggestions on next steps (for example, suitability for CA125 testing or ultrasound scanning), which potentially has a greater utility than a paper based flow chart that might be difficult for busy clinicians to remember in routine primary care.

Limitations

Limitations include a lack of formally adjudicated outcomes, potential information bias, and missing data. Our database has linked cause of death from the UK Office for National Statistics, and our study is therefore likely to have picked up most cases of ovarian cancer, thereby minimising ascertainment bias. While QResearch does not currently have information on the type, grade, and stage of ovarian cancer, it is highly unlikely that the diagnosis would have been recorded without this being established in the clinical setting. The QResearch database is currently being linked to the cancer registry so that more information on type, grade, and stage of cancer at diagnosis will be available for future analyses and refinements of this model. Patients who die from ovarian cancer in hospital will be included through the linked cause of death data. Patients diagnosed with ovarian cancer in hospital will have the information recorded in hospital discharge letters, which are sent to the general practitioner and then entered into the patient's electronic record. The incidence rate in our population was higher than published national data based on cancer registries.² While we rely on accuracy of information recorded by primary care physicians, we think that the quality of information is probably good as previous studies have validated similar outcomes and exposures using questionnaire data and found levels of completeness and accuracy in similar general practice databases to be good.^{32, 33} For example, one systematic review reported that on average 89% of diagnoses recorded on the general practice electronic record are confirmed from other data sources.³²

Another limitation of our study is that recording of symptoms might be less complete or less accurate than diagnostic codes as women might not visit their general practitioner with mild symptoms, might not report all symptoms when they do consult, or general practitioners might not record all the symptoms in the electronic health record. The effect of this information or recording bias could be to overinflate the hazard ratios if they relate to more severe symptoms (such as abdominal distension) or underestimate the hazard ratios if patients with the symptoms don't have them recorded. Also, the design of our study meant it was not possible to rate severity of symptoms as in the study by Goff et al.¹⁵ The Goff study was designed to describe the pattern of self reported symptoms in women with and without ovarian cancer presenting to primary care rather than to develop and validate a prediction algorithm. Similarly, family history of ovarian cancer might be under-recorded as it is not routinely assessed and recorded in general practice records. One practical mechanism to help improve clinical recording of family history

and symptoms for future studies would be to introduce electronic templates into general practice systems that are displayed when a "red flag" symptom is recorded in the patient's record. The template would then help structured data entry of other related symptoms, including important negative findings. Over time this would improve the accuracy and completeness of the electronic record and hence the underlying data used for future versions of this algorithm.

While the validation cohort was derived from practices using the same clinical computer system (EMIS), they were physically discrete. Also, as this computer system is used in over half of general practices on the UK, our results are likely to generalise well. Nonetheless, it is possible that the validation has given overoptimistic results as the practices in the validation sample used the same computer system. A separate independent validation study using another general practice database is planned and hasn't been included in the present study so that it can be undertaken and published by an independent team.

Summary

In summary, we have developed and validated a model that can be used to estimate the absolute risk of patients having an existing but as yet undiagnosed ovarian cancer. The algorithm is based on simple clinical variables that can be ascertained in clinical practice. While the algorithm itself does not make a diagnosis of ovarian cancer, it performed well to identify high risk women in a separate validation sample with good discrimination and calibration. The early diagnosis of ovarian cancer, however, remains a challenge. Further research is needed to assess how best to implement the algorithm, its cost effectiveness, and whether, on implementation, it has any impact on the stage of ovarian cancer at diagnosis and subsequent survival.

We acknowledge the contribution of EMIS practices who contribute to QResearch and EMIS for expertise in establishing, developing, and supporting the database. A simple web calculator to implement the QCancer (ovary) algorithm is available at www.qcancer.org/ovary.

Contributors: JH-C initiated the study, undertook the literature review, data extraction, data manipulation, and primary data analysis, and wrote the first draft of the paper. CC contributed to the design, analysis, interpretation and drafting of the paper. JH-C is guarantor.

Funding: This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Competing interests: Both authors have completed the ICMJE uniform disclosure form at www.icmje.org/coi_disclosure.pdf (available on request from the corresponding author) and declare: JH-C is co-director of QResearch, a not-for-profit organisation that is a joint partnership between the University of Nottingham and EMIS (leading commercial supplier of IT for 60% of general practices in the UK); JH-C is also a paid director of ClinRisk, which produces software to ensure the reliable and updatable implementation of clinical risk algorithms within clinical computer systems to help improve patient care. CC is a paid consultant statistician for ClinRisk. This work and any views expressed within it are solely those of the co-authors and not of any affiliated bodies or organisations.

Ethical approval: All QResearch studies are independently reviewed in accordance with the QResearch agreement with Trent multicentre ethics committee (UK).

Data sharing: The algorithms presented in this paper will be released as Open Source Software under the GNU lesser GPL v3.

1 Ferlay J, Autier P, Boniol M, Heanue M, Colombet M, Boyle P. Estimates of the cancer incidence and mortality in Europe in 2006. *Ann Oncol* 2007;18:581-92.

What is already known on this topic

Ovarian cancer is the second most common gynaecological cancer and most women are diagnosed with late stage disease, which has a poor survival rate

Earlier diagnosis could improve with more targeted investigation of symptomatic patients and increased public awareness of symptoms, which is a major challenge given the non-specific nature of some of the symptoms

What this study adds

An algorithm based on simple clinical variables such as age, family history of ovarian cancer, anaemia, abdominal pain, abdominal distension, rectal bleeding, postmenopausal bleeding, appetite loss, and weight loss, which the patient is likely to know or which are routinely recorded in general practice computer systems, can estimate absolute risk of ovarian cancer in women with and without symptoms in primary care

The algorithm could be integrated into general practice clinical computer systems and used to assess risk in women presenting with and without symptoms

- 2 National Institute for Health and Clinical Excellence. Ovarian cancer: the recognition and initial management of ovarian cancer. NICE, 2011.
- 3 UK Cancer Research. Statistics and outlook for ovarian cancer. UK Cancer Research, 2011.
- 4 Bankhead CR, Kehoe ST, Austoker J. Symptoms associated with diagnosis of ovarian cancer: a systematic review. *BJOG* 2005;112:857-65.
- 5 Rufford BD, Jacobs IJ, Menon U. Feasibility of screening for ovarian cancer using symptoms as selection criteria. *BJOG* 2007;114:59-64.
- 6 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.
- 7 Richards MA. The national awareness and early diagnosis initiative in England: assembling the evidence. *Br J Cancer* 2009;101(suppl 2):S1-4.
- 8 Abdel-Rahman M, Stockton D, Rached B, Hakulinen T, Coleman MP. What if cancer survival in Britain were the same as in Europe: how many deaths are avoidable? *Br J Cancer* 2009;101(suppl 2):S115-24.
- 9 Hippisley-Cox J, Coupland C, Vinogradova Y, Robson J, Minhas R, Sheikh A, et al. Predicting cardiovascular risk in England and Wales: prospective derivation and validation of QRISK2. *BMJ* 2008;336:1475-82.
- 10 Hippisley-Cox J, Coupland C. Predicting risk of osteoporotic fracture in men and women in England and Wales: prospective derivation and validation of QFractureScores. *BMJ* 2009;339:b4229.
- 11 Hippisley-Cox J, Coupland C. Identifying patients with suspected colorectal cancer in primary care: derivation and validation of an algorithm. *Br J Gen Pract* (in press) 2011.
- 12 Hippisley-Cox J, Coupland C. Identifying patients with suspected lung cancer in primary care: derivation and validation of an algorithm. *Br J Gen Pract* 2011;61:715-23.
- 13 Jones R, Latinovic R, Charlton J, Gulliford MC. Alarm symptoms in early diagnosis of cancer in primary care: cohort study using general practice research database. *BMJ* 2007;334:1040.
- 14 Jones R, Charlton J, Latinovic R, Gulliford MC. Alarm symptoms and identification of non-cancer diagnoses in primary care: cohort study. *BMJ* 2009;339:b3094.
- 15 Goff BA, Mandel LS, Melancon CH, Muntz HG. Frequency of symptoms of ovarian cancer in women presenting to primary care clinics. *JAMA* 2004;291:2705-12.
- 16 Hippisley-Cox J, Coupland C, Robson J, Sheikh A, Brindle P. Predicting risk of type 2 diabetes in England and Wales: prospective derivation and validation of QDScore. *BMJ* 2009;338:b880.
- 17 Hippisley-Cox J, Coupland C. Predicting the risk of chronic kidney disease in men and women in England and Wales: prospective derivation and external validation of the QKidney(R) Scores. *BMC Fam Pract* 2010;11:49.
- 18 Collins GS, Altman DG. External validation of the QDScore for predicting the 10-year risk of developing type 2 diabetes. *Diabet Med* 2011;28:599-607.
- 19 Collins GS, Altman DG. An independent and external validation of QRISK2 cardiovascular disease risk score: a prospective open cohort study. *BMJ* 2010;340:c2442.
- 20 Collins GS, Altman DG. An independent external validation and evaluation of QRISK cardiovascular risk prediction: a prospective open cohort study. *BMJ* 2009;339:b2584.
- 21 Schafer J, Graham J. Missing data: our view of the state of the art. *Psychol Methods* 2002;7:147-77.
- 22 Group TAM. Academic medicine: problems and solutions. *BMJ* 1989;298:573-9.
- 23 Steyerberg EW, van Veen M. Imputation is beneficial for handling missing data in predictive models. *J Epidemiol Community Health* 2007;60:979.
- 24 Moons KGM, Donders RART, Stijnen T, Harrell FJ. Using the outcome for imputation of missing predictor values was preferred. *J Epidemiol Community Health* 2006;59:1092.
- 25 Rubin DB. Multiple imputation for non-response in surveys. John Wiley, 1987.
- 26 Royston P, Ambler G, Sauerbrei W. The use of fractional polynomials to model continuous risk variables in epidemiology. *Int J Epidemiol* 1999;28:964-74.
- 27 Hosmer D, Lemeshow S. Applied logistic regression. John Wiley, 1989.
- 28 Royston P. Explained variation for survival models. *Stata J* 2006;6:1-14.
- 29 Royston P, Sauerbrei W. A new measure of prognostic separation in survival data. *Stat Med* 2004;23:723-48.
- 30 Hippisley-Cox J, Coupland C. Predicting the risk of osteoporotic fracture in England and Wales: prospective derivation and validation of the QFractureScore. *BMJ* 2009;339:b4229.
- 31 Jick H, Jick SS, Derby LE. Validation of information recorded on general practitioner based computerised data resource in the United Kingdom. *BMJ* 1991;302:766-8.
- 32 Herrett E, Thomas SL, Schoonen WM, Smeeth L, Hall AJ. Validation and validity of diagnoses in the general practice research database: a systematic review. *Br J Clin Pharmacol* 2010;69:4-14.
- 33 Khan NF, Harrison SE, Rose PW. Validity of diagnostic coding within the general practice research database: a systematic review. *Br J Gen Pract* 2010;60:e128-36.

Accepted: 20 October 2011

Cite this as: *BMJ* 2012;344:d8009

This is an open-access article distributed under the terms of the Creative Commons Attribution Non-commercial License, which permits use, distribution, and reproduction in any medium, provided the original work is properly cited, the use is non commercial and is otherwise in compliance with the license. See: <http://creativecommons.org/licenses/by-nc/2.0/> and <http://creativecommons.org/licenses/by-nc/2.0/legalcode>.

Tables

Table 1 | Baseline characteristics of women in derivation and validation cohorts used to determine algorithm for identification of those with ovarian cancer. Patients were free from diagnosis of ovarian cancer at baseline. Figures are numbers (percentages) unless otherwise specified

	Derivation cohort (n=1 158 723)	Validation cohort (n=608 862)
Mean (SD) age (years)	51 (15.5)	50.9 (15.4)
Mean (SD) Townsend score (deprivation)	-0.4 (3.3)	-0.2 (3.5)
BMI recorded before study entry	956 594 (82.6)	512 388 (84.2)
Mean (SD) BMI	26.1 (5.0)	26.2 (5.0)
Smoking status:		
Non-smoker	666 968 (57.6)	348 702 (57.3)
Ex-smoker	190 871 (16.5)	102 369 (16.8)
Current smoker (number not recorded)	25 634 (2.2)	14 067 (2.3)
Light smoker (<10/day)	63 092 (5.4)	33 721 (5.5)
Moderate smoker (10-19/day)	88 433 (7.6)	47 486 (7.8)
Heavy smoker (≥20 day)	53 408 (4.6)	29 228 (4.8)
Not recorded	70 317 (6.1)	33 289 (5.5)
Alcohol status:		
None	329 078 (28.4)	176 830 (29.0)
Trivial (<1 unit/day)	381 599 (32.9)	207 388 (34.1)
Light (1-2 units/day)	196 164 (16.9)	100 966 (16.6)
Moderate or heavy (≥3 units/day)	28 396 (2.5)	14 740 (2.4)
Not recorded	223 486 (19.3)	108 938 (17.9)
Medical history:		
Previous cancer apart from ovarian cancer	29 333 (2.5)	15 404 (2.5)
Family history of ovarian cancer	1991 (0.2)	1297 (0.2)
Current symptoms:		
Abdominal pain	132 576 (11.4)	73 674 (12.1)
Abdominal distension	5140 (0.4)	3185 (0.5)
Appetite loss	5920 (0.5)	3176 (0.5)
Rectal bleeding	25 274 (2.2)	13 988 (2.3)
Postmenopausal bleeding	18 244 (1.6)	10 285 (1.7)
Weight loss	13 858 (1.2)	7725 (1.3)
Symptoms in preceding year:		
Constipation	9081 (0.8)	5138 (0.8)
Diarrhoea	12 259 (1.1)	6858 (1.1)
Tiredness	15 265 (1.3)	8608 (1.4)
Urinary frequency	931 (0.1)	745 (0.1)
Haemoglobin recorded	242 535 (20.9)	130 067 (21.4)
Haemoglobin <110 g/L	23 589 (2.0)	12 894 (2.1)

Table 2| Adjusted hazard ratios (95% CI) for final model* for ovarian cancer in derivation cohort. Hazard ratios adjusted for all other terms in table and for age

	HR (95% CI)
Family history of ovarian cancer†	9.8 (5.4 to 17.8)
Haemoglobin <110 g/L in past year†	2.3 (1.7 to 2.9)
Current symptoms:	
Abdominal pain†	7.0 (6.1 to 8.0)
Abdominal distension†	23.1 (18.2 to 29.4)
Appetite loss†	5.2 (3.4 to 7.9)
Rectal bleeding†	2.0 (1.4 to 2.8)
Postmenopausal bleeding†	6.6 (5.1 to 8.5)
Weight loss†	2.0 (1.3 to 3.1)

*Also included fractional polynomial terms for age, which were $\text{age}^{-0.5}$, $\text{age}^{-0.5} \ln(\text{age})$.

†Compared with women without this characteristic.

Table 3| Validation statistics for risk prediction algorithm for ovarian cancer in validation cohort

	Mean (95% CI)
R ² statistic* (%)	57.6 (54.8 to 60.4)
D statistic†	2.38 (2.24 to 2.51)
ROC statistic†	0.84 (0.83 to 0.86)

*Shows explained variation in time to diagnosis of ovarian cancer; higher values indicate more variation is explained.

†Measure of discrimination; higher values indicate better discrimination.

Table 4| Comparison of strategies to identify women at risk of diagnosis of ovarian cancer in next two years based on validation cohort

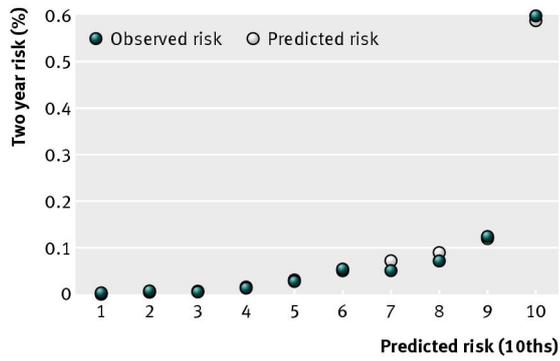
Criteria	Risk threshold (%)	True negative*	False negative*	False positive*	True positive*	Sensitivity (%)	Specificity (%)	Positive predictive value (%)	Negative predictive value (%)
Family history of ovarian cancer	NA	470 593	527	983	11	2.0	99.8	1.1	99.9
Rectal bleeding	NA	461 281	527	10 295	11	2.0	97.8	0.1	99.9
Postmenopausal bleeding	NA	463 624	489	7 952	49	9.1	98.3	0.6	99.9
Abdominal pain	NA	417 609	272	53 967	266	49.4	88.6	0.5	99.9
Abdominal distension	NA	469 312	496	2 264	42	7.8	99.5	1.8	99.9
Appetite loss	NA	469 572	527	2 004	11	2.0	99.6	0.5	99.9
Weight loss	NA	466 430	516	5 146	22	4.1	98.9	0.4	99.9
Any of six above symptoms†	NA	390 842	151	80 734	387	71.9	82.9	0.5	100.0
Top 10% risk	0.2	427 982	198	43 594	340	63.2	90.8	0.8	100.0
Top 5% risk	0.5	450 719	311	20 857	227	42.2	95.6	1.1	99.9
Top 1% risk	0.7	468 084	463	3492	75	13.9	99.3	2.1	99.9
Top 0.5% risk	1.4	469 790	479	1786	59	11.0	99.6	3.2	99.9
Top 0.1% risk	2.3	471 214	517	362	21	3.9	99.9	5.5	99.9

NA=not applicable.

*True negative=criterion not met, does not have disease; false negative=criterion not met, does have disease; false positive=criteria met, does not have disease; true positive=criterion met and has disease.

†This would include 17.2% of all women.

Figure



Mean predicted risk and observed risk of ovarian cancer over two years by 10th of predicted risk, applying risk prediction scores to validation cohort