ARTICLE

Predictive Value of Symptoms for Ovarian Cancer: Comparison of Symptoms Reported by Questionnaire, Interview, and General Practitioner Notes

Anita W. W. Lim, David Mesher, Aleksandra Gentry-Maharaj, Nyaladzi Balogun, Ian Jacobs, Usha Menon, Peter Sasieni

Manuscript received March 18, 2011; revised October 4, 2011; accepted October 14, 2011.

Correspondence to: Peter Sasieni, PhD, Wolfson Institute of Preventive Medicine, Queen Mary University of London, Charterhouse Square, London EC1M 6BQ, UK (e-mail: p.sasieni@qmul.ac.uk).

Background   Because of the poor survival outcomes associated with advanced ovarian cancer, early detection strategies are needed. Although several symptom indices have been described, their relationship with the potential lead time has been poorly documented.

Methods   Women aged 50–79 years who had newly diagnosed ovarian cancer (n = 194) and control subjects (n = 268) who attended ovarian cancer screening clinics were included in the analysis. Symptoms and their onset dates were obtained from three sources: a questionnaire (191 case patients and 268 control subjects), telephone interview (111 case patients and 125 control subjects), and general practitioner (GP) notes (171 case patients and 227 control subjects). Data from questionnaires and GP notes were used to derive two new symptom indices (Index 1 and Index 2). Sensitivity and specificity for these new indices and the previously reported Goff index were calculated for the periods of 0–11 and 3–14 months before diagnosis for all three data sources.

Results   For each data source and period, the two new symptom indices derived from questionnaire and GP notes were similar both qualitatively (symptoms included) and quantitatively (sensitivity and specificity) to the Goff index. When symptoms that started within 3 months before diagnosis were excluded, sensitivity was decreased for all indices and all data sources (eg, for telephone interviews, sensitivity for the period 0–11 vs 3−14 months before diagnosis: for Index 1 = 91.0% vs 69.4%, difference = 21.6%, 95% confidence interval [CI] = 13.6% to 29.7%; for Index 2 = 91.0% vs 60.4%, difference = 30.6%, 95% CI = 21.7% to 39.6%; and for the Goff index = 75.7% vs 51.4%, difference = 24.3%, 95% CI = 16.0% to 32.7%). Also, the specificity of all indices was consistently decreased for telephone interviews compared with questionnaires and GP notes (eg, 1 − specificity for the period of 3–14 months before diagnosis for telephone interviews vs questionnaires: for Index 1 = 19.2% vs 10.4%, difference = 8.8%, 95% CI = 1.0% to 16.6%; for Index 2 = 14.4% vs 6.7%, difference = 7.7%, 95% CI = 0.9% to 14.5%; and for the Goff Index = 7.2% vs 1.5%, difference = 5.7%, 95% CI = 0.9% to 10.5%).

Conclusions   Previous estimates of index performance have been overly optimistic because they did not take into account the time required to make a diagnosis on the basis of testing in response to symptoms. In addition, the specificity of a symptom index is lower when based on a telephone interview vs questionnaire or GP notes. Thus, the clinical utility of a symptom index depends on precisely how it is used and how index-positive women are managed.

J Natl Cancer Inst 2012;104:114–124

Historically, ovarian cancer has been perceived as a silent killer that rarely produces symptoms until it has spread beyond the ovaries. A growing body of evidence suggests otherwise, and this has led to several studies exploring the potential for using patient-reported symptoms as a screening tool to promote earlier diagnosis (1–6). The main challenges to this approach include the low specificity of ovarian cancer symptoms, resulting in an increased work-load for health-care systems and the potential for causing anxiety and interventions that may cause serious harm in women who do not have ovarian cancer, and the lack of evidence for decreased mortality or increased survival.

The most widely evaluated tool is the symptom index developed by Goff et al. (3), which used questionnaire data from case patients with primary ovarian cancer and control subjects undergoing ultrasound or who were at high risk of ovarian cancer and enrolled in a screening study. Goff et al. reported that the Goff index in women age 50 years and older had a sensitivity of 66.7% (in <75 women with cancer) and specificity of 90.0% (in <245 control subjects) for symptoms in the year before diagnosis in a confirmatory group of case patients. Three other groups have applied the Goff index in various settings and obtained comparable estimates of sensitivity and specificity (4,6,7). Women and physicians are

increasingly being advised to use this index or a similar measure, both in the United States (8) and in the United Kingdom (9). It is the foundation for the UK Department of Health's "Key messages on ovarian cancer" to women (10) and the 2011 UK National Institute for Clinical Excellence clinical guidelines on the recognition and initial management of ovarian cancer.

The hope is that a symptom index could be used in addition to current primary care strategies in women who are at low to moderate risk of ovarian cancer. For such a strategy to be effective, symptoms must be present sufficiently before diagnosis to allow for time to screen, evaluate, and intervene. However, the method of symptom assessment (eg, self-completed questionnaires vs patient interviews and symptom checklists vs open-ended questions, etc.) is known to affect the nature and threshold of symptoms reported (11) and will therefore influence index performance.

Here, we report on a multicenter study in which symptom data were collected from three different sources (questionnaire, telephone interview, and general practitioner [GP] notes). The goal was to provide an accurate estimate of the potential for expediting ovarian cancer diagnosis using a symptom index. Our objectives were 1) to derive new symptom indices and evaluate both these and the widely cited Goff index (3), 2) to evaluate the impact of the data source on index performance, and 3) to quantify the loss of sensitivity if symptoms that develop within a few months of diagnosis are discounted. Index performance was assessed for each data source by cross-validation for two 12-month periods: 0–11 months before diagnosis, similar to that used in previous reports; and 3–14 months before diagnosis, designed to take into account time to diagnosis in a (nonurgent) screening setting.

Methods

Study Population

Ethics approval was granted from the Joint University College London/University College London Hospital Ethics Committee (London, UK). All participants provided written informed consent and were enrolled in the UK Ovarian Cancer Population Study (12), a biobank case–control study that recruited from 10 centers across England, Wales, and Northern Ireland between February 16, 2006, and February 28, 2008. Patients with primary ovarian cancer (*International Classification of Diseases, Tenth Revision* C56), recruited before definitive diagnosis or treatment were included in the symptoms study. The date of diagnosis was defined as the date of the first histological/cytological report confirming cancer. Final staging and histology were confirmed by an independent review of the pathology reports and case notes by a gynecological oncologist. Control subjects were women who attended ovarian cancer screening clinics [for the UK Collaborative Trial of Ovarian Cancer Screening (13)], and recruitment took place at annual screening visits. For this symptoms study, we randomly selected noncancer participants in the UK Ovarian Cancer Population Study by frequency matching to balance for year of birth and agreement to a telephone interview (described further below) with the participants who had ovarian cancer. All participants were aged between 50 and 79 years.

CONTEXT AND CAVEATS

Prior knowledge
Reports have indicated that patient-reported symptoms may be a useful screening tool to detect ovarian cancer at early stages of disease. Previously, the Goff index has been shown to be an effective symptom index for identifying women who are at low to moderate risk of ovarian cancer who should undergo screening. However, it is unclear if the method of symptom assessment may influence index performance.

Study design
Symptom data from women newly diagnosed with ovarian cancer and healthy control subjects recorded by a questionnaire and general practitioner notes were used to derive two novel symptom indices, and their sensitivity and specificity were compared with those of the Goff index. Symptom data from questionnaires, telephone interviews, and general practitioner notes were also used to determine the effect of the source of symptom data on index performance. The sensitivity of symptoms that are reported within a few months of diagnosis was also investigated by comparing two 12-month periods (0–11 and 3–14 months before diagnosis).

Contribution
The novel indices derived from questionnaire and general practitioner notes were comparable to the Goff index in terms of the symptoms included, sensitivity, and specificity. The specificity of the indices was decreased for telephone interviews compared with questionnaires and general practitioner notes. When symptoms reported during the first 3 months before diagnosis were excluded from analysis, the sensitivities of all three indices for all data sources were decreased.

Implications
Both the method for ascertaining symptoms from patients and the timing influence the sensitivity and specificity of a symptom index. Because the Goff index has been validated and shown to perform similarly to two novel indices, there may be little to gain from research to derive new ovarian cancer symptom indices.

Limitations
The two novel indices were derived and validated using the same dataset. Recruitment bias may have been introduced by a possible healthy volunteer effect, which would result in a lower prevalence of symptoms among control subjects who were recruited from ovarian cancer screening clinics.

From the Editors

Symptom Ascertainment

Symptom data were collected using three different methods: a questionnaire, a structured telephone interview, and GP notes. Symptom onset dates, symptom duration, and GP visit dates were extracted from each of the three data sources. Onset dates for self-reported data (ie, from questionnaires and telephone interviews) were collected as month and year, and a midpoint (typically the 15th of the month) was used for analysis when a specific day was not provided. For self-reported data, symptom frequency (as days per month on an ordinal categorical scale: 1–4, 5–15, 16–31) and details of whether the symptom was ongoing at recruitment (or diagnosis, if earlier) were also collected.

Women were asked on questionnaires and by telephone interview if they had any symptoms from a checklist comprising 14 of

the most frequently reported symptoms in the literature including pelvic/abdominal pain or discomfort, back pain, indigestion, loss of appetite, nausea or vomiting, weight loss (unplanned) or appearance of weight loss, increase in abdominal size, abdomen feels bloated, able to feel a lump in the abdomen, urinary frequency or urgency, constipation, diarrhea, fatigue, and irregular vaginal bleeding (13–27) (Supplementary Figure 1, available online). Compared with the questionnaire used by Goff et al. (3), our questionnaire (Supplementary Table 1, available online) 1) used simplified symptom wording, 2) grouped closely related symptoms, 3) consisted of a shorter checklist of symptoms (22 vs 14 symptoms were listed), 4) contained minor differences in our categorization of the frequency and severity data, and 5) included details of whether or not the symptom was ongoing.

The questionnaire (Supplementary Figure 1, available online) was completed at the time of recruitment. Interviews were optional and took place within 3 months of diagnosis or recruitment for case patients and control subjects, respectively. All were conducted by a single researcher (A. W. W. Lim). Interviews provided the opportunity for the researcher to probe and clarify ambiguous answers. The wording of "loss of appetite" on the telephone interview checklist differed from that of the questionnaire and appeared as "loss of appetite or feeling full quickly." Symptoms reported on the questionnaire and by telephone interview were recorded regardless of whether or not the participant thought they were associated with ovarian cancer because patients with cancer may misinterpret cancer symptoms as normal changes and vice versa (28,29). Symptoms reported but not included on the checklist were also recorded.

Symptoms were extracted from GP notes for 2 years before diagnosis for case patients and for 2 years before consent for control subjects. Symptom onset was recorded as the first time during the 2 years that the symptom appeared in the notes. Details of symptoms and consultations were extracted by a single researcher (A. W. W. Lim) without blinding to case–control status. A systematic coding frame was drawn up with input from two gynecologists (U. Menon and A. Sharma, Department of Women's Cancer, Institute of Women's Health, University College London, London, UK) to ensure that symptoms were categorized consistently.

Development of Novel Symptom Indices

For each data source, symptoms were only included in the final analysis if they started within 15 months of diagnosis for case patients or consent for control subjects. This 15-month cut-off was on the basis of the observation that there were no case–control differences for any symptom beyond 15 months (data not shown).

Onset dates were only rarely missing for case patients; 86 (7%) of 1173 symptoms reported on questionnaire and 20 (3%) of 769 reported on telephone interview had missing onset dates. For control subjects, onset dates were missing more often: 169 (34%) of 491 questionnaire symptoms and 76 (16%) of 473 telephone interview symptoms had missing onset dates. We studied the symptom onset date collected from telephone interviews when it was missing on the individual's corresponding questionnaire. Onset dates appeared to be missing at random among the data collected from case patients. By contrast, control subjects missing onset dates on the questionnaire were typically reported via the telephone interview to have started more than 2 years before consent was given.

Therefore, symptoms with missing onset dates were treated as if they had started at least 2 years before consent and were excluded from the calculations.

Two symptom indices were derived separately using data from the questionnaire (Index 1) and GP notes (Index 2). Telephone interview data were not used to derive an index because of the smaller number of participants compared with questionnaire (n = 236 and 459 for telephone interview and questionnaire, respectively). Indices were derived using backward stepwise selection logistic regression on symptoms present 3–14 months before diagnosis.

The symptom with the greatest *P* value was removed from the model if its *P* value was greater than .1, and the *P* value for the restricted model was recalculated. Previously excluded symptoms, with a *P* value less than .05, were added back to each index. Symptoms that were statistically significantly associated with cancer, but were dropped by the software package because they predicted ovarian cancer perfectly, were added back into the final index. Only the 14 checklist symptoms (Supplementary Figure 1, available online) were included in the derivation of the questionnaire index because the number and type of symptoms detected can vary depending on whether they are elicited by a checklist or spontaneous reporting (30,31). In addition to the 14 checklist symptoms, symptoms in GP notes that displayed a statistically significant association with ovarian cancer in the study subjects by univariate analysis and had a difference in symptom prevalence between those with and without cancer of 5% or more within 0–14 months of diagnosis were included in the stepwise regression. These were leg swelling, change in bowel habit, vaginal discharge, and urinary symptoms other than frequency or urgency (eg, retention, dysuria, change in urine color or smell, hematuria, etc.). Symptoms that were retained by the backward stepwise selection logistic regression of questionnaire and GP note data formed Index 1 and Index 2, respectively.

Having derived the indices, they were applied to two 12-month periods: 0–11 and 3–14 months before diagnosis. A woman was considered positive for the index if she reported at least one index symptom that started within the period of interest. For each index, we ran a 10-fold cross-validation on the stepwise regression method used to derive it. The case patients and control subjects were randomly assigned to 10 groups. Leaving out one group, a backward stepwise regression was performed, as described above, on the remaining 9/10 of the data. This stepwise regression was used to generate a symptom index, which was validated using the group not used to generate that index: We counted the number of women in the group who did and did not have cancer and who had an index symptom. By leaving out each of the 10 groups in turn and repeating the process of index derivation and validation, every study subject was used once for validation. The sum of the number of correctly identified case patients and control subjects from all 10 groups was used to estimate sensitivity (ie, the percentage of case patients with a positive index) and specificity (ie, the percentage of control subjects with a negative index), respectively.

Index Performance

The performance of the Goff index was assessed separately for each data source 0–11 and 3–14 months before diagnosis. A woman

is positive on the Goff index if she had a new Goff symptom occurring more than 12 times per month within the previous 12 months. Goff symptoms are pelvic/abdominal pain, increased abdominal size/bloating, or difficulty eating/feeling full quickly (3). We considered the Goff index to be positive if a woman had any one of the same set of symptoms that was new in the respective period of interest and occurred 16–31 d/mo (on the questionnaire and telephone interview). For GP notes, symptom frequency and duration were not available, and so we considered the Goff Index to be positive if any Goff index symptom was recorded as a new symptom within the period of interest.

Statistical Analyses

Odds ratios and 95% confidence intervals (CIs) were calculated by use of unconditional logistic regression to evaluate the association between symptoms, treated as binary categorical variables, and case–control status. These odds ratios may be interpreted as a relative measure of the risk of a woman having ovarian cancer given that she has symptoms. No adjustments were made for race/ethnicity as there were only six case patients and five control subjects who were not of European descent. Odds ratios were calculated without adjustment (univariate) for each symptom and with adjustment for other symptoms in Index 1 and Index 2 (yes or no), which were derived by backward stepwise logistic regression. Confidence intervals rather than testing were used to summarize the role of random variation in the estimated statistics. Confidence intervals for odds ratios are based on the asymptotic standard error of the logarithm of the odds ratio, unless the odds ratio was infinite, in which case a lower bound for the confidence interval was obtained by the Cornfield method.

The sensitivity of the Goff index was estimated as the number of case patients who are positive on the index divided by the total number of case patients on whom the index could be evaluated. The specificity was estimated as the number of control subjects negative on the index divided by the total number of control subjects on whom the index could be evaluated. The results are sometimes shown as 1 − specificity to estimate the proportion of women without cancer who would test positive for a symptom index. Confidence intervals were based on the binomial distribution. For Indices 1 and 2, these statistics were based on the cross-validated analysis as described under "Development of Novel Symptom Indices."

The positive diagnostic likelihood ratio was defined as the ratio of the sensitivity to one minus the specificity (ie, sensitivity/[1 − specificity]). It is a useful measure for comparing the positive predictive value (PPV) of different tests (or indices) when the prevalence of disease is unknown because for a rare disease (and provided the specificity is not very close to 100%), the PPV is closely approximated by the prevalence times the positive diagnostic likelihood ratio. Confidence intervals for the positive diagnostic likelihood ratio were calculated using standard methods for the risk ratio from a 2 × 2 table. Comparisons of the sensitivity using 0–11 months compared with 3–14 months were on the basis of the difference in sensitivities and takes into account case patients who only had an index symptom 0–2 months before diagnosis and those who only had one 12–14 months before diagnosis. Confidence intervals were calculated as the difference plus or minus 1.96

standard errors, and the standard error was based on multinomial distribution.

The sensitivity with 95% confidence intervals of the Goff index for the questionnaire, telephone interview, and GP notes was calculated and stratified by early- and late-stage cancers (stage I–II and III–IV, respectively, by the International Federation of Gynecology and Obstetrics criteria). The null hypothesis that the sensitivities are the same for early- and late-stage cancers was tested for each data source using the Pearson χ^2 statistic.

All statistical analyses were performed using STATA for Windows (version 10.0, StataCorp LP, College Station, TX). A *P* value of less than .05 was considered statistically significant. All statistical tests were two-sided.

Results

Study Population Characteristics

There were 194 women with newly diagnosed ovarian cancer, and 268 healthy control subjects who met the study inclusion criteria. Exclusion of participants in the UK Ovarian Cancer Population Study who did not meet our additional inclusion criteria (incident ovarian cancer and aged between 50 and 79 years) resulted in fewer case patients than control subjects for our analysis. Questionnaires were completed by 191 (98%) case patients and 268 (100%) control subjects (Table 1). Among the participants, 76% of all study participants (case patients and control subjects) agreed to be contacted for a telephone interview. Of these, 111 (76%) of 147 case patients and 125 (61%) of 205 control subjects had a telephone interview. GP notes were obtained for 171 (88%) case patients and 227 (85%) control subjects.

The mean age of case patients was 65 years (range = 50–79 years); the mean age of control subjects was also 65 years (range = 52–78 years) (Table 1). Case patients (189 of 194 patients) and control subjects (263 of 268 control subjects) were predominantly of European descent. Among the case patients, 22 (11%) had cancers that were primary borderline epithelial. Seventy-three (38%) case patients were diagnosed with early-stage tumors (stage I–II), 108 (56%) were diagnosed with late-stage tumors (stage III–IV), and the stage was not available for 13 (7%).

Symptom Reporting and Positive Predictive Value of Data Sources

The most common symptoms reported 3–14 months before ovarian cancer diagnosis were abdominal bloating (reported by 61 of 191 case patients on the questionnaire), fatigue (reported by 48 of 111 case patients by telephone interview), and pelvic/abdominal pain or discomfort (recorded in GP notes for 44 of 171 case patients) (Table 2). For control subjects, the most common symptoms included fatigue (reported by 14 of 268 control subjects on the questionnaire and 14 of 125 control subjects by telephone interview), urinary frequency or urgency (reported by 29 of 191 control subjects on the questionnaire), pelvic/abdominal pain or discomfort (recorded in GP notes for 22 of 227 control subjects), and other urinary symptoms (recorded in GP notes for 22 of 227 control subjects).

Stepwise regression identified six symptoms from the questionnaire (pelvic/abdominal pain or discomfort, loss of appetite or

Table 1. Characteristics of case patients (n = 194) and control subjects (n = 268)

Characteristic	Case patients	Control subjects	All study participants
Mean age, y (range)	65 (50–79)	65 (52–78)	65 (50–79)
Race, No. (%)			
European descent	189 (97.4)	263 (98.1)	452 (97.8)
Black Caribbean	1 (0.5)	4 (1.5)	5 (1.1)
Jewish Ashkenazi	1 (0.5)	1 (0.4)	2 (0.4)
Jewish Sephardi	1 (0.5)	0	1 (0.2)
Mixed race	2 (1.0)	0	2 (0.4)
Type of tumor, No. (%)			
Invasive epithelial	166 (85.6)	0	166 (85.6)
Borderline epithelial	22 (11.3)	0	22 (11.3)
Non-epithelial	6 (3.1)	0	6 (3.1)
Stage*, No. (%)			
I–II	73 (37.6)	0	73 (37.6)
III–IV	108 (55.7)	0	108 (55.7)
Unknown	13 (6.7)	0	13 (6.7)
Symptom assessments, No. (%)			
Questionnaires	191 (98.5)	268 (100.0)	459 (99.4)
General practitioner notes			
Telephone interviews	171 (88.1)	227 (84.7)	398 (86.1)
Agreed to be interviewed	147 (75.8)	205 (76.5)	352 (76.2)
Completed interview	111 (57.2)	125 (46.6)	236 (51.1)

* Tumor stage was determined by applying the International Federation of Gynecology and Obstetrics staging criteria.

feeling full quickly, weight loss, increase in abdominal size, abdomen feels bloated, and able to feel a lump in the abdomen) and five symptoms from GP notes (pelvic abdominal pain or discomfort, loss of appetite, increase in abdominal size, able to feel a lump in the abdomen, and vaginal discharge) that were independently associated with ovarian cancer within 3–14 months before diagnosis (Table 2) forming Index 1 and Index 2, respectively. Four symptoms appeared in both indices (pelvic/abdominal pain or discomfort, loss of appetite, increase in abdominal size, and able to feel a lump in the abdomen).

Symptom reporting among control subjects, and consequently PPV, varied substantially depending on how symptoms were ascertained. In general, symptom reporting was higher in self-reported data (particularly by telephone interview) than in physician records (GP) (Table 3 and Figure 1). For the Goff index, the positive diagnostic likelihood ratio for 3–14 months before diagnosis was 24.2 (95% CI = 9.0 to 65.2) for the questionnaire, whereas for the telephone interview, it was 7.1 (95% CI = 3.7 to 13.7) (Table 3). Thus, the estimated PPV would be 3–4 times greater using questionnaire data compared with telephone interview data.

Sensitivity and Specificity of Symptom Indices for Ovarian Cancer

The cross-validated estimates of sensitivity and specificity for the 0- to 11-month and 3- to 14-month periods for Index 1, Index 2, and the Goff index were calculated (Table 4). Because Index 1 included all of the symptoms that are featured in the Goff index (in addition to weight loss and lump in abdomen), its sensitivity must be at least as high, whereas its specificity can be no higher than those of the Goff index. In practice, for the telephone interview data 0–11 months before diagnosis, the sensitivity of Index 1 was higher (91.0% vs 75.7% for Goff, difference = 15.3%, 95% CI = 5.6% to 25.1%), but the specificity was lower (76.0% vs 89.6%,

difference = 13.6%, 95% CI = 8.1% to 20.9%). The performance of Index 2 within this same period (sensitivity = 91.0%, 95% CI = 84.1% to 95.6%; specificity = 82.4%, 95% CI = 74.6% to 88.6%) was similar to that of Index 1 for the telephone interview (Table 4). The relative differences in symptom ascertainment between data sources were greatest for the control subjects. For example, the sensitivity of the Goff index 0–11 months before diagnosis was 61.8% (95% CI = 54.5% to 68.7%) for the questionnaire compared with 75.7% (95% CI = 66.6% to 83.3%) for the telephone interview, whereas the specificity was 98.9% (95% CI = 96.8% to 99.8%) vs 89.6% (95% CI = 82.9% to 94.3%), respectively (Table 4). For all three indices, sensitivity was lowest for data recorded in the GP notes and highest for the telephone interview, whereas specificity was lower for the telephone interview and higher for the questionnaire (eg, 1 − specificity for the time period of 3–14 months before diagnosis for telephone interviews vs questionnaires: for Index 1 = 19.2% vs 10.4%, difference = 8.8%, 95% CI = 1.0% to 16.6%; for Index 2 = 14.4% vs 6.7%, difference = 7.7%, 95% CI = 0.9% to 14.5%; and for the Goff index = 7.2% vs 1.5%, difference = 5.7%, 95% CI = 0.9% to 10.5%). The results for nonoverlapping intervals (0–2, 3–5, and 6–11 months) before diagnosis were as one would expect given those observed for the overlapping intervals (Supplementary Table 2, available online).

When the indices were applied to symptoms present 3–14 months before diagnosis, the sensitivity was reduced markedly, although specificity was largely unchanged (Table 4). The greatest sensitivity observed was 69.4% (for Index 1 on telephone interview) suggesting that at best, a symptom index might advance diagnosis of ovarian cancer by 3 months or more in about two-thirds of women. Depending on the data source and index, there was between 21.6% (for Index 1 telephone interview data) and 32.2% (for Index 1 GP note data) reduction in sensitivity when the period examined shifted from 0–11 to 3–14 months (eg, for telephone

Table 2. Checklist and other symptoms together with numbers reported by case patients and control subjects on questionnaire (Q), telephone interview (TI), and general practitioner (GP) notes, univariate and multivariable odds ratios and inclusion in the three indices*

Symptom	No. of participants reporting symptoms, case patients (control subjects)			Univariate OR (95% CI)†			Multivariable OR (95% CI)†		Index‡		
	Q, n = 191 (268)	TI, n = 111 (125)	GP notes, n = 171 (227)	Q	TI	GP notes	Q	TI	1	2	Goff
Checklist symptoms, reported 3–14 mo before diagnosis											
Pelvic/abdominal pain or discomfort	54 (8)	41 (11)	44 (22)	12.8 (5.9 to 27.7)	6.1 (2.9 to 12.6)	3.2 (1.8 to 5.6)	7.1 (3.1 to 16.4)	2.9 (1.6 to 5.2)	+	+	+
Back pain	61 (11)	35 (8)	9 (5)	2.9 (1.4 to 6.1)	4.0 (1.3 to 12.7)	1.9 (1.0 to 3.8)	NA	NA	−	−	−
Indigestion	55 (6)	45 (6)	12 (1)	3.7 (1.8 to 7.4)	2.3 (1.0 to 5.5)	1.3 (0.6 to 2.6)	NA	NA	−	−	−
Loss of appetite or feeling full quickly	32 (1)	26 (1)	8 (1)	53.7 (7.3 to 397.1)	37.9 (5.1 to 284.9)	11.1 (1.4 to 89.6)	18.4 (2.2 to 154.2)	7.6 (0.9 to 64.7)	+	+	+
Nausea or vomiting	17 (3)	11 (2)	12 (11)	8.6 (2.5 to 29.9)	6.8 (1.5 to 31.2)	1.5 (0.6 to 3.4)	NA	NA	−	−	−
Weight loss (unplanned)§	30 (5)	18 (3)	5 (2)	9.8 (3.7 to 25.8)	7.9 (2.3 to 27.5)	3.4 (0.6 to 17.7)	2.9 (0.9 to 9.9)	NA	+	−	−
Increase in abdominal size	55 (6)	45 (6)	12 (1)	17.7 (7.4 to 42.1)	13.5 (5.5 to 33.4)	17.1 (2.2 to 132.5)	5.4 (1.9 to 15.8)	14.9 (1.9 to 118.9)	+	+	+
Abdomen feels bloated	61 (11)	35 (8)	9 (5)	11.0 (5.6 to 21.5)	6.7 (3.0 to 15.3)	2.5 (0.8 to 7.5)	2.7 (1.1 to 6.7)	NA	+	+	+
Able to feel a lump in abdomen	8 (0)	6 (0)	6 (0)	∞ (3.2 to ∞)	∞ (1.8 to ∞)	∞ (2.1 to ∞)	NA	NA	+	−	−
Urinary frequency or urgency	29 (13)	29 (11)	19 (13)	3.5 (1.8 to 7.0)	3.7 (1.7 to 7.8)	2.1 (1.0 to 4.3)	NA	NA	−	−	−
Constipation	23 (6)	16 (1)	18 (10)	6.0 (2.4, 15.0)	20.9 (2.7 to 160.3)	2.6 (1.1 to 5.7)	NA	NA	−	−	−
Diarrhea	1 (0)	6 (1)	14 (8)	3.2 (1.2 to 8.5)	1.9 (0.4 to 8.2)	1.0 (0.4 to 2.5)	NA	NA	−	−	−
Fatigue	48 (14)	48 (14)	13 (14)	6.1 (3.2 to 11.4)	6.0 (3.1 to 11.8)	1.3 (0.6 to 2.7)	NA	NA	−	−	−
Irregular vaginal bleeding‖	13 (6)	6 (3)	6 (6)	3.2 (1.2 to 8.5)	2.3 (0.6 to 9.5)	1.3 (0.4 to 4.2)	NA	NA	−	−	−
Additional symptoms, reported 0–14 mo before diagnosis in GP notes											
Urinary other	2 (0)	10 (0)	27 (22)	NA	NA	2.4 (1.4 to 4.1)	NA	NA	−	−	−
Vaginal discharge	0 (0)	5 (2)	9 (3)	NA	NA	5.1 (1.4 to 4.1)	NA	4.6 (1.2 to 17.6)	−	+	−
Change in bowel habit	1 (0)	6 (1)	14 (8)	NA	NA	4.8 (2.3 to 10.1)	NA	NA	−	−	−
Leg swelling	5 (0)	3 (1)	4 (1)	NA	NA	12.6 (1.6 to 100.1)	NA	NA	−	−	−

* CI = confidence interval; NA = not applicable; OR = odds ratio.

† Odds ratios are estimated by unconditional logistic regression and refer to the odds of ovarian cancer in those with a symptom relative to the odds in those without a particular symptom (all symptoms are treated as binary). Confidence intervals are based on the asymptotic standard error of the log odds ratio unless the odds ratio is infinite in which case the lower limit is based on the Cornfield method. The multivariable odds ratios are adjusted for all other symptoms with odds ratio in the column of the table. Univariate odds ratios are unadjusted.

‡ Data from questionnaires were used to derive Index 1 and general practitioner notes were used to derive Index 2. For Index 1 and Index 2, a woman was considered to have a positive index if she reported at least one index symptom in the indicated period before diagnosis (3–14 or 0–11 months). For a woman to have a positive Goff index (3), she had to report any one of the following symptoms as having occurred 16–31 d/mo with a date of onset within the time period: pelvic/abdominal pain or discomfort, bloating/increased abdominal size, or feeling full quickly.

§ This symptom includes the appearance of weight loss.

‖ This symptom includes primarily postmenopausal bleeding but also postcoital and intermenstrual bleeding in premenopausal women.

Table 3. Positive diagnostic likelihood ratio for Index 1, Index 2, and the Goff index using data from questionnaires, telephone interviews, and general practitioner (GP) notes*

Index and time before diagnosis†	Positive diagnostic likelihood ratio (95% CI)		
	Questionnaire (n = 191 case patients and 268 control subjects)	Telephone interview (n = 111 case patients and 125 control subjects)	GP notes (n = 171 case patients and 227 control subjects)
Index 1			
0–11 mo	9.2 (6.3 to 13.6)	3.8 (2.8 to 5.2)	5.6 (3.9 to 8.1)
3–14 mo	5.1 (3.5 to 7.4)	3.6 (2.5 to 5.3)	2.4 (1.7 to 3.4)
Index 2			
0–11 mo	13.1 (8.1 to 21.1)	5.2 (3.5 to 7.6)	4.8 (3.4 to 6.7)
3–14 mo	6.9 (4.3 to 11.1)	4.2 (2.7 to 6.6)	2.3 (1.6 to 3.2)
Goff Index			
0–11 mo	55.2 (17.8 to 171.0)	7.3 (4.3 to 12.3)	7.0 (4.5 to 10.8)
3–14 mo	24.2 (9.0 to 65.2)	7.1 (3.7 to 13.7)	2.9 (1.9 to 4.5)

* The positive diagnostic likelihood ratio was calculated by dividing the sensitivity by 1 − specificity. CI = confidence interval.

† Data from questionnaires were used to derive Index 1 and GP notes were used to derive Index 2. For Index 1 and Index 2, a woman was considered to have a positive index if she reported at least one symptom in the indicated period before diagnosis (0–11 or 3–14 months for Index 1 and Index 2). For a woman to have a positive Goff index (3), she had to report any one of the following symptoms as having occurred 16–31 d/mo with a date of onset within the time period: pelvic/abdominal pain or discomfort, bloating/increased abdominal size, or feeling full quickly.

interviews, sensitivity for the time period 0–11 vs 3–14 months before diagnosis: for Index 1 was 91.0 % vs 69.4%, difference = 21.6%, 95% CI = 13.6% to 29.7%; for Index 2 was 91.0 % vs 60.4%, difference = 30.6%, 95% CI = 21.7% to 39.6%; and for the Goff index was 75.7% vs 51.4%, difference = 24.3%, 95% CI = 16.0% to 32.7%). On analysis of the questionnaire data, we found that 49 (25.7%) case patients who reported Goff index symptoms in the year before diagnosis did not report any of these symptoms within the period of 3–14 months before diagnosis (data not shown). The corresponding attrition on telephone interview was 27 (24.3%) case patients and 50 (29.2%) case patients for GP notes. The relative loss of sensitivity over time from diagnosis was similar for all three data sources, with between 22.5% (telephone

interview, Index 1) and 35.1% (GP notes, Index 1) of women with symptoms only having symptoms within 3 months of diagnosis and a further 17.2% (questionnaire, Goff index) to 29.8% (telephone interview, Index 1) only having symptoms 3–5 months before diagnosis (Supplementary Table 2, available online).

The cumulative incidence of symptoms by data source for each windex was also investigated by plotting the cumulative proportion of individuals with at least one index symptom over time (Figure 1). This was done separately for case patients and control subjects to provide a visual representation of how long before diagnosis case patients develop index symptoms and how this compares with control subjects. As expected, the proportion of case patients and control subjects with index symptoms start off similarly and then

Table 4. Sensitivity and 1 − specificity of Index 1, Index 2, and the Goff index for 0–11 and 3–14 months before diagnosis for each data source*

Participants by index and time before diagnosis	Questionnaire		Telephone interview		GP notes†	
	Sensitivity, % (95% CI)	1 − specificity, % (95% CI)	Sensitivity, % (95% CI)	1 − specificity, % (95% CI)	Sensitivity, % (95% CI)	1 − specificity, % (95% CI)
No. of participants	191	268	111	125	171	227
Index 1‡						
0–11 mo	82.7 (76.6 to 87.8)	9.0 (5.8 to 13.0)	91.0 (84.1 to 95.6)	24.0 (16.8 to 32.5)	69.6 (62.1 to 76.4)	12.3 (8.4 to 17.3)
3–14 mo	53.4 (46.1 to 60.6)	10.4 (7.1 to 14.7)	69.4 (59.9 to 77.8)	19.2 (12.7 to 27.2)	37.4 (30.2 to 45.1)	15.9 (11.4 to 21.3)
Index 2§						
0–11 mo	78.0 (71.5 to 83.7)	6.0 (3.5 to 9.5)	91.0 (84.1 to 95.6)	17.6 (11.4 to 25.4)	67.3 (59.7 to 74.2)	14.1 (9.8 to 19.3)
3–14 mo	46.6 (39.4 to 53.9)	6.7 (4.0 to 10.4)	60.4 (50.6 to 69.5)	14.4 (8.8 to 21.8)	38.0 (30.7 to 45.7)	16.7 (12.1 to 22.2)
Goff index‖						
0–11 mo	61.8 (54.5 to 68.7)	1.1 (0.2 to 3.2)	75.7 (66.6 to 83.3)	10.4 (5.7 to 17.1)	61.4 (53.7 to 68.7)	8.8 (5.5 to 13.3)
3–14 mo	36.1 (29.3 to 43.4)	1.5 (0.4 to 3.8)	51.4 (41.7 to 61.0)	7.2 (3.3 to 13.2)	32.2 (25.2 to 39.7)	11.0 (7.3 to 15.8)

* Sensitivity is the percentage of case patients with a positive index; 1 − specificity is the percentage of control subjects with a positive index. The estimates for Index 1 and Index 2 use a method (cross-validation), which corrects for the fact that the indices were derived from the same data. CI = confidence interval; GP = general practitioner.

† GP notes had no information available on the duration of symptoms, so all symptoms present were considered to have occurred 16–31 d/mo.

‡ Index 1 was on the basis of stepwise regression on questionnaire data for symptoms present in period before diagnosis.

§ Index 2 was on the basis of stepwise regression on GP note data for symptoms present in period before diagnosis.

‖ The Goff index was approximated as any one of pelvic/abdominal pain or discomfort, bloating/increased abdominal size, or feeling full quickly and occurred within 16–31 d/mo and the onset was within the indicated period.

Table 5. Sensitivity of the Goff index for early- and late-stage ovarian cancer*

	Sensitivity, % (95% CI)								
	Questionnaire			Interview			GP notes		
Time before diagnosis	Early stage (n = 72)	Late stage (n = 106)	P†	Early stage (n = 45)	Late stage (n = 60)	P†	Early stage (n = 69)	Late stage (n = 92)	P†
0–11 mo	61.1 (48.9 to 72.4)	67.0 (57.2 to 75.8)	.42	73.3 (58.1 to 85.4)	78.3 (65.8 to 87.9)	.55	55.1 (42.6 to 67.1)	68.5 (58.0 to 77.8)	.08
3–14 mo	31.9 (21.4 to 44.0)	39.6 (30.3 to 49.6)	.30	51.1 (35.8 to 66.3)	53.3 (40.0 to 66.3)	.82	36.2 (25.0 to 48.7)	31.5 (22.2 to 42.0)	.53

* Early-stage cancers include stages I–II and late-stage cancers include stages III–IV. Sensitivity is the percentage of case patients with a positive index. CI = confidence interval; GP = general practitioner.

† All P values were calculated by Pearson χ^2 test.

there is a marked acceleration in the proportion of case patients with symptoms closer to diagnosis. For all three indices, sharp increases in the cumulative symptom incidence were observed within 3 months before diagnosis for case patients. A symptom-based intervention was unlikely to be useful for advancing diagnosis in patients who only develop symptoms within 3 months of diagnosis because by this point in time, it was likely that the patient was already being investigated for suspected ovarian cancer (possibly precisely as a result of these same symptoms). Thus, the plots are useful for gauging the magnitude of the lead time that might be achieved through a symptom-based intervention.

In all situations, the sensitivities for the Goff index in early- and late-stage disease were similar (Table 5). In most situations, the sensitivity was greater in late-stage compared with early-stage disease. These findings indicate that there is little difference in the symptoms experienced by women with early- vs late-stage disease.

The Goff index has been validated by several groups in various countries with comparable results (4,6,7) (Table 6). The minor differences between our questionnaire and that used by Goff et al. (3) did not have a major impact on results, as demonstrated by the similarity of our results compared with those of Goff et al. (Goff study: sensitivity = 66.7%, specificity = 90.0%; our study: sensitivity = 61.8%, specificity = 98.9%). In particular, we did not find an increase in specificity despite using a shorter and more focused (eg, "unplanned weight loss" instead of "weight loss") symptom list.

Discussion

Although a high proportion of women with ovarian cancer experience symptoms before diagnosis, approximately one-third of the reported sensitivity of symptoms is because of symptoms that start within 3 months of diagnosis. Many of these symptoms will have been the initiator of the diagnostic process and even if they were not, there is little opportunity for advancing the date of diagnosis by more than a few weeks in these women. Symptom-based testing involving direct questioning of women is likely to lead to investigations in a higher number of women without ovarian cancer. We have shown that both the timing of symptoms and the mode of elicitation are important considerations when developing and evaluating a symptom index. The small differences in performance between all three indices (Index 1, Index 2, and the Goff index) indicate that there is little to gain from further research to derive a new ovarian cancer symptom index.

We report that between 21.2% (Index 1 telephone interview) and 32.2% (Index 1 GP notes) of women with ovarian cancer only have symptoms within 3 months of diagnosis, which is supported by Rossing et al. (6) who found that symptoms developing within 3 months of diagnosis accounted for roughly one-third of the sensitivity over 0–11 months, and also by Hamilton et al. (32) who found few symptoms other than bloating more than 100 days before diagnosis. Inclusion of such symptoms leads to artificially favorable estimates of the clinical value of any symptom index. Using three different indexes, we have shown that a symptom index is unlikely to detect more than

Table 6. Comparison of sensitivity and specificity of the Goff index from our study and previously published reports for questionnaire and telephone interview data*

Data source and study	Time before diagnosis, mo	No. of case patients	No. of control subjects	Age, y	Sensitivity, % (95% CI)	Specificity, % (95% CI)	Reference
Questionnaire							
Goff et al.	0–11	≤245	≤75	≥50	66.7 (NA)	90.0 (NA)	(3)
Kim et al.	0–11	116	209	18–77	65.5 (56.1 to 74.1)	84.7 (79.1 to 89.3)	(4)
Our study	0–11	191	268	50–79	61.8 (54.5 to 68.7)	98.9 (96.8 to 99.8)	NA
	3–11	191	268	50–79	29.8 (23.5 to 36.9)	98.9 (96.8 to 99.8)	NA
Telephone interview							
Jordan et al.	0–11	NA	NA	55–79	68.4 (65.2 to 71.6)	81.9 (79.2 to 84.5)	(7)
Rossing et al.†	0–11	2592	1300	55–74	66.2 (63.4 to 69.0)	95.7 (94.6 to 96.9)	(6)
	3–11	2592	1300	55–74	42.6 (38.6 to 46.7)	95.5 (94.3 to 96.6)	(6)
Our study	0–11	111	125	50–79	75.7 (66.6 to 83.3)	89.6 (82.9 to 94.3)	NA
	3–11	111	125	50–79	50.5 (40.8 to 60.1)	93.6 (87.8 to 97.2)	NA

* CI = confidence interval (if available or computable from data in the article); NA = not available.

† This study excluded women with missing index status (positive or negative) or unknown onset date of first symptom.
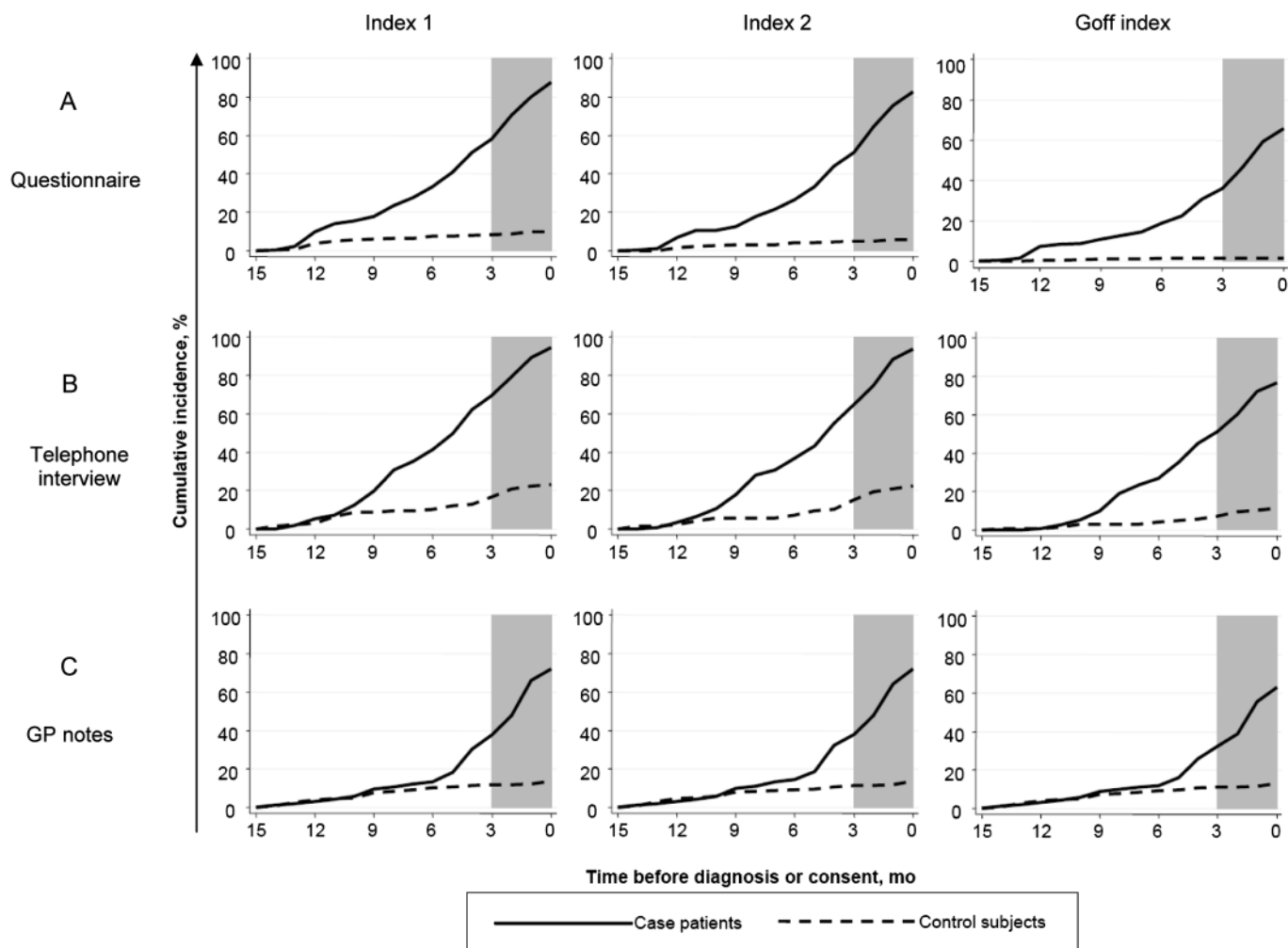
Figure 1. The cumulative incidence of positive indices among case patients and control subjects throughout 0–14 months before diagnosis for each data source. The 3 months before diagnosis are **shaded** to demarcate the period during which women are already in, or about to enter, the referral system. The **shaded areas** indicate a peri-diagnostic period when investigations specific to the diagnostic work-up for ovarian cancer were likely to already be underway. **A)** Data from 191 case patients and 268 control subjects were analyzed for the questionnaire; **B)** data from 111 case patients and 125 control subjects were analyzed for the telephone interview; and **C)** data from 171 case patients and 227 control subjects were analyzed for the general practitioner (GP) notes.

one- to two-thirds (depending on its specificity) of women with ovarian cancer more than 3 months before diagnosis.

It is often implicitly assumed that when symptoms are present, the cancer would be detected by the screening and diagnostic work-up that would be triggered. However, the relationship between symptoms and CA125 or ultrasound is poorly understood. We know of only one prospective study that investigated this relationship, which was small and did not include a single cancer (34). It is likely that some women investigated for a symptom index will be declared cancer-free only to be diagnosed with ovarian cancer some months later. For example, in a study with 75 case patients and 254 control subjects, Andersen et al. (2) found that of the 48 case patients with symptoms, 40 were also positive for CA125 (the threshold was chosen so that 5% of the screening population was positive). Using the same threshold, none of the 30 control subjects with symptoms were also positive for CA125.

A key strength of this study is the comparison of different data sources, which previously has not been done. However, an assessment of concordance between the three data sources would require

a complex analysis and is the subject of a separate article being prepared by this group. The inclusion of women from 10 centers in this study adds to the external validity of our findings.

This study also has several limitations. The indices were derived and assessed using the same data. However, the cross-validation makes better use of the data (compared with splitting the sample into training and verification samples) and provides unbiased estimates because no woman is used for both training and validation of the same model. Also, one of the symptoms in the Goff index, "difficulty eating/feeling full quickly," was not on our questionnaire checklist and therefore may have been underreported as it was not directly elicited. However, this would underestimate rather than overestimate sensitivity. We cannot rule out recall bias for questionnaire and telephone interview data and case patients may have reported more symptoms than control subjects. This was minimized by telephone interviewing within 3 months of diagnosis and often before definitive diagnosis. The potential bias from differential lack of date of symptom onset between case patients and control subjects was minimized by treating all such symptoms as being "longstanding," as supported by

data from other sources. Even if all women with missing symptom onset dates had symptoms starting within 3–14 months before diagnosis, this would only have increased the sensitivity of the Goff index on questionnaire by 1.6% (three case patients) and decreased its specificity by 0.4% (one control subject). Other potential drawbacks include recruitment bias of case patients and a possible "healthy volunteer" effect leading to lower symptom prevalence in the control subjects who were recruited from screening clinics. Excluding symptoms that started more than 15 months before diagnosis may have underestimated symptom lead time. However, this is justifiable because increasing evidence suggests that symptoms manifest at most about a year before ovarian cancer diagnosis (19,32,34–36). Inclusion of older symptoms tended to reduce the specificity by the same amount as it increased the sensitivity.

We have derived two new symptom indices, which were both qualitatively and quantitatively similar to the Goff index. The small differences between the three indices indicate that there is little to gain from deriving new symptom indices. The data source (ie, mode of symptom elicitation) has substantial impact on results, in particular on the symptoms reported by control subjects. The loss of sensitivity if the 3 months before diagnosis are discounted is substantial, indicating that simply reporting the PPV of a symptom index overestimates the potential added benefit that it could have in advancing diagnosis. Consequently, the potential for any index to affect time to diagnosis in ovarian cancer is much smaller than previously suggested. At best, a symptom index might advance diagnosis of ovarian cancer by 3 months or more in two-thirds of women. For a more specific index, the sensitivity would be approximately one-third. The mortality impact of such a lead time is unknown. The design of future studies needs to carefully consider the approach to symptom ascertainment. There is a need for prospective studies using an existing index to better ascertain the likely number needed to test to diagnose one ovarian cancer and the associated lead time distribution.

References

1. Andersen MR, Goff BA, Lowe KA, et al. Use of a symptom index, CA125, and HE4 to predict ovarian cancer. *Gynecol Oncol.* 2009;116(3): 378–383.
2. Andersen MR, Goff BA, Lowe KA, et al. Combining a symptoms index with CA 125 to improve detection of ovarian cancer. *Cancer.* 2008;113(3): 484–489.
3. Goff BA, Mandel LS, Drescher CW, et al. Development of an ovarian cancer symptom index: possibilities for earlier detection. *Cancer.* 2007; 109(2):221–227.
4. Kim MK, Kim K, Kim SM, et al. A hospital-based case-control study of identifying ovarian cancer using symptom index. *J Gynecol Oncol.* 2009; 20(4):238–242.
5. Lurie G, Thompson PJ, McDuffie KE, Carney ME, Goodman MT. Prediagnostic symptoms of ovarian carcinoma: a case-control study. *Gynecol Oncol.* 2009;114(2):231–236.
6. Rossing MA, Wicklund KG, Cushing-Haugen KL, Weiss NS. Predictive value of symptoms for early detection of ovarian cancer. *J Natl Cancer Inst.* 2010;102(4):222–229.
7. Jordan SJ, Coory MD, Webb PM. Re: predictive value of symptoms for early detection of ovarian cancer. *J Natl Cancer Inst.* 2010;102(20): 1599–1601.
8. Gynecologic Cancer Foundation SoGO, American Cancer Society. Ovarian Cancer Symptoms Consensus Statement. http://www.wcn.org/ov_cancer_cons.html. Accessed May 11, 2011.
9. Eve Appeal, Ovacome. Ovarian Cancer UK Consensus Statement. http://www.eveappeal.org.uk/media/5381/ovarian%20cancer%20uk%20consensus%20statement.pdf. Accessed February 1, 2011.
10. Department of Health. Ovarian Cancer: Key Messages for Health Professionals. http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_110534. Accessed February 4, 2011.
11. Kroenke K. Studying symptoms: sampling and measurement issues. *Ann Intern Med.* 2001;134(9, pt 2):844–853.
12. Balogun N, Gentry-Maharaj A, Wozniak EL, et al. Recruitment of newly diagnosed ovarian cancer patients proved challenging in a multicentre biobanking study. *J Clin Epidemiol.* 2010;64(5):525–530.
13. Menon U, Gentry-Maharaj A, Hallett R, et al. Sensitivity and specificity of multimodal and ultrasound screening for ovarian cancer, and stage distribution of detected cancers: results of the prevalence screen of the UK Collaborative Trial of Ovarian Cancer Screening (UKCTOCS). *Lancet Oncol.* 2009;10(4):327–340.
14. Attanucci CA, Ball HG, Zweizig SL, Chen AH. Differences in symptoms between patients with benign and malignant ovarian neoplasms. *Am J Obstet Gynecol.* 2004;190(5):1435–1437.
15. Chan YM, Ng TY, Lee PW, Ngan HY, Wong LC. Symptoms, coping strategies, and timing of presentations in patients with newly diagnosed ovarian cancer. *Gynecol Oncol.* 2003;90(3):651–656.
16. Eltabbakh GH, Yadav PR, Morgan A. Clinical picture of women with early stage ovarian cancer. *Gynecol Oncol.* 1999;75(3):476–479.
17. Flam F, Einhorn N, Sjovall K. Symptomatology of ovarian cancer. *Eur J Obstet Gynecol Reprod Biol.* 1988;27(1):53–57.
18. Goff BA, Mandel L, Muntz HG, Melancon CH. Ovarian carcinoma diagnosis. *Cancer.* 2000;89(10):2068–2075.
19. Goff BA, Mandel LS, Melancon CH, Muntz HG. Frequency of symptoms of ovarian cancer in women presenting to primary care clinics. *JAMA.* 2004;291(22):2705–2712.
20. Kirwan JM, Tincello DG, Herod JJ, Frost O, Kingston RE. Effect of delays in primary care referral on survival of women with epithelial ovarian cancer: retrospective audit. *BMJ.* 2002;324(7330):148–151.
21. Olson SH, Mignone L, Nakraseive C, et al. Symptoms of ovarian cancer. *Obstet Gynecol.* 2001;98(2):212–217.
22. Smith EM, Anderson B. The effects of symptoms and delay in seeking diagnosis on stage of disease at diagnosis among women with cancers of the ovary. *Cancer.* 1985;56(11):2727–2732.
23. Vine MF, Calingaert B, Berchuck A, Schildkraut JM. Characterization of prediagnostic symptoms among primary epithelial ovarian cancer cases and controls. *Gynecol Oncol.* 2003;90(1):75–82.
24. Vine MF, Ness RB, Calingaert B, Schildkraut JM, Berchuck A. Types and duration of symptoms prior to diagnosis of invasive or borderline ovarian tumor. *Gynecol Oncol.* 2001;83(3):466–471.
25. Webb PM, Purdie DM, Grover S, et al. Symptoms and diagnosis of borderline, early and advanced epithelial ovarian cancer. *Gynecol Oncol.* 2004;92(1):232–239.
26. Wikborn C, Pettersson F, Silfversward C, Moberg PJ. Symptoms and diagnostic difficulties in ovarian epithelial cancer. *Int J Gynaecol Obstet.* 1993;42(3):261–264.
27. Yawn BP, Barrette BA, Wollan PC. Ovarian cancer: the neglected diagnosis. *Mayo Clin Proc.* 2004;79(10):1277–1282.
28. Andersen RS, Vedsted P, Olesen F, Bro F, Sondergaard J. Patient delay in cancer studies: a discussion of methods and measures. *BMC Health Serv Res.* 2009;9(189). doi:10.1186/1472-6963-9-189.
29. Bankhead CR, Collins C, Stokes-Lampard H, et al. Identifying symptoms of ovarian cancer: a qualitative and quantitative study. *BJOG.* 2008;115(8): 1008–1014.
30. Funch DP. Predictors and consequences of symptom reporting behaviors in colorectal cancer patients. *Med Care.* 1988;26(10):1000–1008.
31. Wallander MA, Dimenas E, Svardsudd K, Wiklund I. Evaluation of three methods of symptom reporting in a clinical trial of felodipine. *Eur J Clin Pharmacol.* 1991;41(3):187–196.
32. 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. doi:10.1136/bmj.b2998.

33. Rufford BD, Jacobs I, Menon U. Feasibility of screening for ovarian cancer using symptoms as selection criteria. *BJOG*. 2007;114(1):59–64.

34. Friedman GD, Skilling JS, Udaltsova NV, Smith LH. Early symptoms of ovarian cancer: a case-control study without recall bias. *Fam Pract*. 2005; 22(5):548–553.

35. Smith LH, Morris CR, Yasmeen S, et al. Ovarian cancer: can we make the clinical diagnosis earlier? *Cancer*. 2005;104(7):1398–1407.

36. Wynn ML, Chang S, Peipins LA. Temporal patterns of conditions and symptoms potentially associated with ovarian cancer. *J Womens Health (Larchmt)*. 2007;16(7):971–986.

Affiliations of authors: Centre for Cancer Prevention, Wolfson Institute of Preventive Medicine, Barts and The London School of Medicine and Dentistry, Queen Mary University of London, London, UK (AWWL, DM, PS); Department of Women's Cancer, Institute of Women's Health, University College London, London, UK (AG-M, NB, UM); Faculty of Medical and Human Sciences, University of Manchester, Manchester, UK (IJ).