INVESTIGATING THE POTENTIAL FOR EXPEDITING DIAGNOSIS OF OVARIAN CANCER VIA PROMPT SYMPTOM RECOGNITION & 'TARGETED SCREENING'

Thesis submitted for the degree of Doctor of Philosophy by

Anita Wey Wey Lim

Wolfson Institute of Preventive Medicine

Queen Mary University, London

April 2009

Abstract

This thesis explores the potential for using symptoms as a tool to bring forward the diagnosis of ovarian cancer. Current evidence supports the existence of prediagnostic symptoms, however symptom lead time has yet to be adequately quantified. 'Targeted screening' is one possible approach to expediting diagnosis. This would involve offering a blood test (e.g. CA125 or a future biomarker) to postmenopausal women presenting to primary care with symptoms possibly related to ovarian cancer. Key barriers include the non-specific nature of ovarian cancer symptoms and potential impact on GP workload. The main aims of this PhD research project were:

- To quantify the lead time of symptoms in ovarian cancer.
- To estimate the GP workload associated with offering a blood test to postmenopausal women with ovarian cancer symptoms.

Following a background to ovarian cancer symptoms research; a brief overview of the epidemiology of ovarian cancer, a case-control study to quantify symptom lead time, and a cross-sectional pilot study to estimate GP workload and symptom specificity in women aged 45-74 in the general population is presented. This is complemented by a systematic review focussed on the evidence for symptom lead time in ovarian cancer since 1980, with an update on ovarian cancer symptoms and a discussion of some of the methodological issues. The main findings showed that the diagnostic process could be initiated at least 3 months prior to the current date of diagnosis, in 45% to 74% of cases. However, pilot data suggest that 13%-35% of women in the general population aged ≥45 would be offered 'targeted screening' in 1 year. Delays in ovarian cancer diagnosis of concern were identified but require further examination. Timing of symptoms is also an important consideration.

The concluding chapter summarises the main findings of this thesis and discusses possibilities for future research.

Table of Contents

List of Figures

List of Tables

ACKNOWLEDGEMENTS

I would like to express my gratitude to Prof. Peter Sasieni for his supervision, continued support and for giving me the opportunity to undertake this thesis project.

I am also extremely grateful to my second supervisors Dr Usha Menon and Prof. Stephen Duffy for their invaluable input and support, in particular to Usha for her much needed clinical expertise and knowledge of ovarian cancer.

I am indebted to David Mesher for teaching me how to use Stata, and for providing excellent statistical support and guidance, all with admirable patience. I would also like to thank Dr Clare Bankhead, Prof. Joan Austoker and Megan Goldsmith for their generosity, expertise, collaboration and support, Dr Aarti Sharma for her help with clinical aspects, Keith Prescott, Gladys Foudjour and Mark Cronin for their guidance and help with navigating the EMIS software, Richard Ostler for his work on developing the EMIS anonymiser, Anne Lancely for her advice and debriefing for the telephone interviews, Rob Edwards for helping with the development of the access database, Prof. Donna Lamping for her advice regarding symptom measurement issues, Prof. Jane Wardle for her advice on the piloting and developing the questionnaire, Willie Hamilton for being so forthcoming in sharing his knowledge, Professor Ian Jacobs for allowing me to work within his group, and to Prof. Vaskar Saha and Prof. Stephan Schug for encouraging me to undertake this thesis.

I am also grateful to the UKOPS team for allowing me to collaborate with them; in particular I would like to thank Nyala Balogun and Aleksandra Gentry-Maharaj for their invaluable support and dedication to the UKOPS study. A special thanks to all the UKOPS nurses, and to all the women who participated in the UKOPS study. Thank you to Chrisp Street practice, Wokingham, Northcroft and Woosehill surgeries for allowing me to access their data.

I would also like to thank Penny Allen and my colleagues Alex, Iva, Laurence and Benoit for their support and advice, and Prof. Jack Cuizick and everyone at the Wolfson Institute of Preventive Medicine for providing such a stimulating academic environment to work in.

A very special thank you to Cancer Research UK, the Oak Foundation and the Eve Appeal for providing funding for the project.

Finally, I would like to thank my parents, Bernadine and Andreas for their constant support and encouragement.

ABBREVIATIONS

AJCC	American Joint Committee on Cancer
CA125	Cancer Antigen 125
BLOT	Borderline Ovarian Tumour
EMIS	Egton Medical Information Systems
EOC	Epithelial Ovarian Cancer
EPR	Electronic Patient Record
FIGO	International Federation of Gynaecology and Obstetrics
HRT	Hormone replacement therapy
IBS	Irritable Bowel Syndrome
NICE	National Institute for Health and Clinical Excellence
NHS	National Health Service
ONS	Office of National Statistics
PMB	Postmenopausal bleeding
PV	Pelvic Vaginal (examination)
RMI	Risk of Malignancy Index
ROC	Risk of Ovarian Cancer
TMN	Tumour, Metastasis, Node involvement
UCL	University College London
UKCTOCS	United Kingdom Collaborative Trial of Ovarian Cancer
UKOPS	United Kingdom Ovarian Cancer Population Study
UTI	Urinary Tract Infection

PART I: BACKGROUND

Part I will contextualise the setting for this thesis project by providing a summary of:

- Ovarian cancer symptoms
- Epidemiology of ovarian cancer
- Current screening possibilities
- Treatment
- Why this research was needed

Also included is a systematic review of the recent literature on ovarian cancer symptoms with a focus on the timing of ovarian cancer symptoms and methodological issues.

1 CHAPTER 1: Ovarian Cancer Overview

1.1 Introduction

Ovarian cancer is the leading cause of mortality for women with gynaecological malignancies in the United Kingdom (UK) and many other Western countries.[1] The majority of women are diagnosed at advanced stages when prognosis is poor and 5-year survival rate is as low as 15%.[2] Early stage disease is associated with a more favourable 5-year survival rate of 70%, but only 20% of women present at this stage.[2] Recently, there have been improvements in 5-year survival, and a substantial increase in 10-year survival is expected in the near future.[3] Stage at diagnosis is one of the strongest predictors of survival[4] hence, there is great incentive to diagnose women 'early' while disease is localised, however this has proven to be challenging. Although ovarian cancer is an ideal candidate for screening, there is currently no proven screening mechanism, even for high-risk groups.[5] Given that an acceptable screening program has yet to be identified, 'early' diagnosis relies heavily on timely presentation to health care and physician ability to make the diagnosis. The situation is compounded by the absence of an identifiable premalignant lesion or specific biomarker, and a tenuous symptom profile.

Historically, ovarian cancer has been perceived as a 'silent' or 'whispering' disease that rarely produces symptoms until disease has spread beyond the ovaries. Yet, symptoms are the most frequently cited reason for the medical consultation leading to diagnosis in ovarian cancer.[6, 7] However, this is not surprising given that there is no screening programme. A plethora of studies have served to disprove the 'silent disease' myth, and there is great interest in using symptoms as a diagnostic tool for detecting ovarian cancer. Hence, research efforts in symptoms have intensified. One of the key papers to trigger this surge in research was a survey by Goff *et al.* which helped to identify and establish the existence of prediagnostic symptoms in ovarian cancer.[8] Importantly, the study also linked delays in diagnosis and women ignoring symptoms to late stage disease. Although this study helped to raise awareness of prediagnostic symptoms in ovarian cancer; it relied solely on subjective data from women, with no objective data to support the findings. Nevertheless, the main findings have been confirmed by several subsequent medical record studies.[9-16]

Overall, symptoms appear to be relatively common in the immediate period leading up to diagnosis, but it remains unclear what role, if any, they could have in promoting diagnosis at earlier stages of disease. Unfortunately, most of the symptoms are subtle

and non-specific, and none are exclusively associated with ovarian cancer. In general, symptoms tend to be abdominal or gastrointestinal rather than gynaecological,[8, 12, 15, 17-19] and most are relatively common complaints in postmenopausal women. Furthermore, the fact that the symptoms are likely to be common in the general population raises the critical issue of low specificity. The specificity of any 'screening' test for ovarian cancer is crucial given that false positives can lead to serious psychological and physiological morbidity. As such, while patient support groups moot that symptoms hold the key to early diagnosis, most of the medical community have remained largely reticent; unconvinced until there is stronger evidence. This disagreement has sparked a 'help versus harm' debate (discussed in further detail later in this Chapter).

Regrettably, only sparse information exists regarding the actual timing of symptoms prior to ovarian cancer diagnosis. Much of the current data represents symptoms *at* diagnosis, and the data relating to timing have been presented in a convoluted manner. Presumably, symptoms fluctuate with time in terms of type, severity and frequency. Therefore, it is sensible to examine what symptoms are present at different periods before diagnosis, and in what proportion of women. This would give a clearer picture of symptom development and any potential opportunities for earlier diagnosis. So far, only three studies have used this approach.[10, 20, 21]

There is growing pressure to act upon symptoms in ovarian cancer. In January 2007 'Joanna's Law' was passed in the United States (US), which committed $16.5 million of federal funding to educate women and their healthcare providers on ovarian cancer signs and symptoms (*http://www.johannaslaw.org*). This was followed by an ovarian cancer consensus symptoms statement (June 2007) which stated that any women with pelvic/abdominal pain, bloating, difficulty eating/feeling full quickly or urinary frequency/urgency almost daily for more than a few weeks, should visit a physician for pelvic examination (*http://www.wcn.org/materials/gcam.html*). The UK followed suit in October 2008, with a consensus statement published as a collaborative effort between the University College London (UCL) Institute for Women's Health, the Eve Appeal and Ovacome (*http://www.ovacome.org.uk/Resources/OvarianCancerUKConsensusStatement*). Although the UK consensus statement was more conservative than its US equivalent, it still encourages women to seek healthcare for symptoms which are both highly subjective and vague. Thus, there is ongoing promotion of using symptoms as a tool for earlier ovarian cancer diagnosis, however the evidence to support any benefit is non-existent. Many researchers and clinicians believe that such efforts are premature

especially given that prevalence of the same symptoms in the general population is unknown.

1.1.1 Symptoms Research

Symptoms are ubiquitous and represent the primary reason for most outpatient clinic visits.[22] Presenting symptoms are those that cause a person to seek healthcare, however they may not always be the first symptom(s) to appear. For research purposes, the subjective nature of symptoms is a major complicating factor. The presence of an objective manifestation such as cough makes symptom assessment much more straightforward, but is a rarity. Furthermore, the background prevalence of common symptoms is substantial in outpatient clinic populations, therefore many coincidental associations may be incorrectly interpreted as causal.[23, 24] Symptoms are often self-limiting, and there is likely to be a large proportion of women with vague symptoms in the non-consulting general population.[25]

Other issues in symptom research include the variable periodicity of symptoms (e.g. vomiting may occur 5 times in 1 day but constipation could be all day for one week), and interference from medication and other exogenous compounds (due to side effects and/or efficacy).[24, 26] Perhaps an even greater hurdle is the significant impact of elicitation method on symptom reporting. The number and type of symptoms detected can vary dramatically depending on whether they are elicited by checklist or spontaneous reporting.[27, 28]

All of these factors need to be taken into consideration when designing studies investigating ovarian cancer symptoms. Yet, many of these issues have not been acknowledged in the ovarian cancer symptoms literature.

1.1.2 'Help Versus Harm' Issue

The rationale for increasing symptom awareness is based on the assumption that earlier detection of symptoms will lead to earlier stage disease, and therefore better survival. An online survey carried out in the US found that only 15% of women aged ≥40 knew about ovarian cancer symptoms (*http://www.mnovarian.org/article-090705.htm*). In the UK, a nationwide survey performed by Dr Foster Intelligence found that 66% of women were unable to name a single symptom associated with ovarian cancer (*http://www.ovarian.org.uk/news/awareness_study.asp*). While these findings certainly identify a gap in knowledge, they do not validate any symptoms-based testing. Prompt diagnosis after symptom recognition is likely to have

a psychological benefit, but there is currently no evidence for any favourable effects on survival or overall mortality. Furthermore, encouraging women to go to their general practitioner (GP) for vague symptoms could impact on GP workload by increasing the number of presentations from the 'worried well'. In addition, there are other possible deleterious effects such as unnecessary anxiety and potentially, invasive surgery. There are minimal data to support the effectiveness of increased cancer symptom awareness except for well defined symptoms such as breast lumps and breast cancer. Equally, there is little evidence to suggest increased awareness affects GP workload or patient anxiety.[29]

Raising awareness of ovarian cancer symptoms in GPs is complicated by the fact that the average GP will see just one case of ovarian cancer every five years.[30] However, many still hope that encouraging GPs to at least consider an ovarian malignancy may hasten diagnosis.

1.2 Rationale for Research Project

Research focussed on prediagnostic symptoms in ovarian cancer has gained momentum over recent years, however, scarce data exist on the UK population. It is clear that symptoms in ovarian cancer exist, but the current knowledge-base relies heavily on symptoms *at* diagnosis. The next logical step is to identify whether these symptoms are detectable early enough to expedite diagnosis (via earlier referral), and in turn prevent or at least postpone death from ovarian cancer. Essentially, this relates to the quantification of symptom lead time and the proportion of women with ovarian cancer that may have this lead time. A crucial element of this is the proportion of women in the general population who also have the symptoms.

Clearly, it would be inappropriate to offer invasive and costly procedures to every woman that presents to primary care with vague symptoms. However, it may be feasible to identify a cluster of symptoms that would increase the index of suspicion for ovarian cancer sufficiently to justify 'screening' in a subpopulation. In this scenario, women in whom there is a moderate to high suspicion of disease would still be managed via the current routes (i.e. urgent ultrasound or rapid referral).

'Screening' is most conventionally applied to an asymptomatic population. It may be defined as "the systematic application of a test or enquiry, to identify individuals at sufficient risk of a specific disorder to benefit from further investigation or direct preventative action, among persons who have not sought medical attention on account of symptoms of that disorder".[31] The term 'targeted screening' will be employed in this

thesis, even though testing based on symptoms is not strictly compatible with the aforementioned definition of screening. However, this nonclamenture may be justified by the fact that testing would only be offered to women in whom there is only a very low suspicion of ovarian cancer, and the vast majority of women tested are expected to be disease-free.

'Targeted screening' is most likely to be offered to postmenopausal women only (or based on an age restriction). Such a restriction would significantly improve the positive predictive value of symptoms by taking advantage of the age-specific incidence of ovarian cancer (>85% of women with ovarian cancer are aged 45 and over[32, 33]) and the fact that most women with disease are postmenopausal. In addition, because prognosis tends to be poorer in these women, there is a higher chance of impacting on survival and mortality.

Ovarian cancer symptom data are extremely complex and most studies have not addressed this in their design. Previous studies have been inconsistent in terms of study population, methodology and symptom definitions. There is a clear need to better establish symptom incidence (cases and controls) and any associated diagnostic delays in a UK population.

This thesis will explore the potential for 'targeted ovarian cancer screening' by:

1) Establishing the distribution of symptom duration prior to ovarian cancer diagnosis

2) Estimating the proportion of apparently healthy women that would be offered 'targeted screening' over the course of 1 year

3) Taking into account some of the practical issues of 'targeted screening', such as impact on GP workload.

1.3 Studies Included in this Thesis

This thesis was based on a Cancer Research UK project grant awarded to Professor Peter Sasieni to investigate the scope for 'targeted' ovarian cancer screening. A series of studies were performed under the umbrella of this project grant, of which this thesis will include three:

1) A systematic review of the timing of prediagnostic ovarian cancer symptoms.

2) A case-control study to investigate symptoms and events preceding ovarian cancer diagnosis.

3) A pilot primary care study to estimate the GP workload associated with offering 'targeted' ovarian cancer screening.

All of the participants of the symptoms case-control study were identified within the framework of a larger case-control study called UK Ovarian Cancer Population Study (UKOPS), funded by the Oak Foundation via the Eve Appeal (lead investigators Dr Usha Menon, Dr Simon Gayther, Professor Ian Jacobs). Data in the symptoms study were collected as an addition to the data already planned for collection in the main UKOPS study, with the exception of the questionnaire and the basic case details (histology, date of diagnosis etc.).

1.4 Epidemiology of Ovarian Cancer

Ovarian cancer is a disease of predominantly older postmenopausal women. In general, the natural history of ovarian cancer is poorly understood. There are 3 major categories of ovarian cancer including epithelial, germ cell and stromal cell. Epithelial ovarian cancer (EOC) is by far the most common comprising 90% of all ovarian malignancies.[34] As such, research efforts are largely directed at this subtype. EOC is thought to arise from the surface epithelium covering the ovary and may differentiate into serous, endometrioid, clear cell, mucinous and poorly differentiated subtypes.

Increased parity, combined oral contraceptive use, tubal ligation and hysterectomy are all established as being protective for ovarian cancer. In contrast, increasing age, and mutations in BRCA1 and BRCA2 are the strongest risk factors for disease.

1.4.1 Incidence

Incidence in the UK has been rising, a total of 6806 ovarian cancer cases were diagnosed in 2005.[32, 35, 36] Ovarian cancer incidence increases with increasing age, almost 84% of affected women are aged 50 or older (5745 of 6806). Sharp increases in incidence coincide around the typical age of menopause (incidence in postmenopausal women is about 40 per 100,000 per year).[37] In the future, incidence may start to decrease as a more women who have used oral contraceptives (which are protective for disease) will become postmenopausal,[38] however this could be countered by increases in the number of women who are living longer.

1.4.2 Survival

The UK age-standardised five-year survival rates have doubled between 1971 and 2001, hence long-term survival is also expected to increase.[39] Stage at diagnosis

remains the strongest predictor of survival, and women with early disease have five-year survival rates in excess of 70%.[37]

1.4.3 Mortality

Ovarian cancer is the fifth leading cause of death for women with 4407 dying from disease per year.[35, 40] This is equivalent to 6% of all female cancer deaths.

1.4.4 Risk Factors

Age

Increasing age is a major risk factor, and the highest incidence rates are for women aged 65 and above.[32] The highest incidence rates are observed for women aged 65 and over.[32, 35, 36] Only a minority of cases are aged under 50, and these tend to be non-epithelial subtypes with favourable prognoses.[41]

Geography

The highest rates of ovarian cancer are observed in Northern Europe and the United States, while the lowest are in Asia and Africa.[37] Population migrant studies have shown that risk increases in women who migrate from countries of lower to higher risk, providing evidence for a role of environmental factors.[42]

Reproductive Factors

A long menstrual life has been postulated to be a risk factor for ovarian cancer. Overall, late menopause and age at menarche have not been shown to have any effect on the risk of disease. Only weak associations between increasing age at natural menopause and risk of disease have been identified.[42]

Parity

The protective effect of parity is well established in ovarian cancer. One study showed a risk reduction of 40% for the first birth, with further risk reduction for each subsequent birth.[43] Conversely, nulliparity predisposes women to an increased risk of ovarian cancer. The effect of incomplete pregnancies is less clear, although these may be mildly protective.[42]

Oral Contraceptive Use

The association between combined oral contraceptive pill use and reduced ovarian cancer risk is another well demonstrated protective factor.[37] Risk reduction increases with duration of use and the trend continues for up to 20 years after termination of

use.[44] The mechanism is thought to be via cessation of ovulation and/or reduction in gonadotrophin levels.

Breast feeding

Overall, breastfeeding is thought to be protective. This may be specific to histological subtype.[37]

Infertility & Infertility Drugs

Since nulliparity can increase the risk of ovarian cancer, it is difficult to study the effect of infertility alone. Women who are nulliparous are not necessarily subfertile. Studies examining the effects of infertility and fertility treatment have produced conflicting results. The situation is complicated by the difficulties with defining infertility type and controlling for the various fertility treatments. A pooled analysis of eight case-control studies showed that risk of ovarian cancer was increased in nulligravid women who had at least 5 years of unsuccessful attempts at pregnancy.[45]

So far, no clear relationship between fertility drugs and risk of ovarian cancer has been identified.[37, 42]

Hormone Replacement Therapy (HRT)

Oestrogen replacement therapy (ERT) alone has been shown to be associated with greater risk increases than oestrogen-progestin therapy.[46] Ever-users versus never-users of ERT have an increased risk between 19-24%.[46]

Genetic

Only 5-10% of ovarian cancer cases are attributable to familial links, the remainder are sporadic.[37] While hereditary predisposition accounts for a very small percentage of cases, it is the strongest risk factor for disease. A meta-analysis showed that relative risk with an affected first-degree relative is 3.1 (95% CI 2.6, 3.7) and increases proportionately with each additional first or second degree relative affected.[47] The majority of relevant mutations are BRCA1- and BRCA2-mediated. Lifetime risk for BRCA1 carriers is 16-44% and 27% for BRCA2 carriers.[42] Hereditary non-polyposis colorectal cancer syndrome (HNPCC) also increases the risk of developing disease. However, this is to a lesser extent with only 10% lifetime risk for HNPCC carriers.[48]

Other

Evidence from most studies suggest a protective effect of tubal ligation and hysterectomy. Risk reductions between 10% and 80% have been observed for tubal

ligation and a modest protective effect for hysterectomy.[37] Notably, protection diminishes within 20 years after surgery implying that the observed effect may be due to inspection of ovaries at the time of surgery.

An increased incidence of endometriosis has been found in women with endometrioid and clear cell ovarian cancers.[42] Many of these were found to be in continuum with the carcinomas suggesting that endometriosis may play a precursory role in the oncogenesis of these histological subtypes.

The relationship between diet and ovarian cancer is unclear.[37, 42] High consumption of olive oil and green vegetables is thought to decrease risk while a high intake of animal fats and meat have been associated with increased risk.

A link between body mass index (BMI) and ovarian cancer has been purported, particularly in association with the upper percentiles of body mass index.[37] However, the precise details of this relationship is yet to be established.

Talcum powder use on the perineum is a contentious risk factor for ovarian cancer. A meta-analysis of 16 studies demonstrated a relative risk of 1.33 (95%CI 1.16 to 1.45) however no dose-response relationship could be identified.[49] More recently, further studies have found weak associations between talc use and ovarian cancer.[50]

Polycystic ovary syndrome (PCOS) is thought to increase the risk of disease, however, there is insufficient evidence in this area.[37]

1.5 Early Detection & Screening in Ovarian Cancer

In ovarian cancer, a single screening test would require particularly high specificity since a positive test usually mandates an invasive surgical procedure. An arbitrary goal of at least 10% Positive Predictive Value (PPV) is generally accepted, which would mean there would be ten unnecessary operations performed for every case found.[51] However, it is becoming increasingly unlikely that a single test will be employed in ovarian cancer screening.

The main strategies for screening involve both biochemical and morphological markers. The most thoroughly investigated of these include antigen serum marker CA125 and transvaginal ultrasound. Neither of these techniques alone has sufficient specificity and sensitivity to utilise for screening.

Screening is unlikely to be beneficial if most screen-detected cancers are borderline or tumours that would have in any case been confined to the ovaries at presentation. A

reduction in mortality is more likely to be associated with screen-detection of early-stage serous or undifferentiated carcinomas since they are usually advanced at presentation. There are currently two large randomised controlled studies investigating the efficacy of screening in the general population using CA125 testing and transvaginal ultrasound. In the UK, there is the United Kingdom Collaborative Trial of Ovarian Cancer Screening (UKCTOCS) and in the US there is the Prostate, Lung, Colorectal and Ovarian cancer screening trial (PLCO). Both are due to report mortality data around 2015, but until then the possible mortality benefit from screening remains unknown.

1.5.1 Transvaginal Sonography (TVS)

Transvaginal sonography (TVS) is used to assess ovarian volume and visualise ovarian morphology. It has relatively high sensitivity but insufficient specificity in asymptomatic women.[52] In addition, the high costs are restrictive, and specialist skills are required for accurate scan interpretation. As a first-line test TVS has demonstrated higher sensitivity for early stage disease.[53]

1.5.2 Tumour Markers

Cancer antigen 125 (CA125) is a high molecular weight glycoprotein, and it is the most extensively investigated tumour marker for ovarian malignancies. Serum levels are elevated in 50%-60% of stage I cancers (not screen-detected)[54] and in more than 90% of women with more advanced stages.[55] Elevation may be observed as long as 10 to 60 months before diagnosis.[54] However, CA125 expression is minimal or absent in 20% of ovarian cancers, particularly in mucinous subtypes.[55] Furthermore, CA125 is non-specific and may be raised during menstruation and in various other conditions including fibroids, diverticulitis and endometriosis.[56] A useful distinguishing feature of malignant versus benign suprathreshold CA125 values is rising levels from serial testing. Combination of CA125 with other markers to improve sensitivity has been evaluated, however, to date this has only been achieved at the expense of specificity.[54, 57] Recent advancements include the use of an algorithm that accounts for the rate of change of CA125 which increases sensitivity.[58] The risk of ovarian cancer (ROC) algorithm accounts for a woman's age-specific incidence of ovarian cancer and serial CA125 profile to give a more accurate estimate of risk.[58]

There are a host of other biomarkers being evaluated but to date, none have met the stringent sensitivity and specificity requirements for screening in ovarian cancer.

1.5.3 Multimodal Screening (MMS)

So far, the best screening results have been achieved with multimodal screening (MMS) using both CA125 and transvaginal ultrasound in a postmenopausal population.[59, 60] This approach was evaluated in 22,000 postmenopausal women, specificity was 99.9%, sensitivity was 78.6% and PPV was 26.8% at 1-year follow-up.[59] A randomised controlled trial using the same population found a significant difference in the median survival rate for screened versus control women (72.9 months vs. 41.8 months, p=0.011).[61]

As mentioned previously, there are two ongoing studies designed to assess the efficacy of multimodal screening (PLCO and UKCTOCS). Prevalence screening results have been published for both studies.[62, 63] In the PLCO study, 28,816 women had at least one of CA125 and TVS.[62] 1706 (5.9%) women had at least one abnormal test, resulting in 570 surgical investigations. A total of 29 neoplasms were identified (26 ovarian, 2 fallopian and 1 primary peritoneal), of which 19 were epithelial invasive cancers. Of the 29 neoplasms, 9 women were positive on both tests (31%). PPV of either test for invasive epithelial cancer was just 1.1%, but 23.5% if both tests were positive.

In UKCTOCS, 50,075 women received multimodal screening, of whom 4555 had repeat testing, leading to 97 surgical evaluations.[63] A total of 42 ovarian or tubal cancers were detected. Of these 34 were primary invasive epithelial ovarian and tubal cancers, and a further 4 were detected as interval cancers. Corresponding sensitivity for primary invasive epithelial ovarian and tubal malignancies was 89.5% and specificity was 99.8%. PPV was relatively high at 43.3%. For TVS alone, sensitivity, specificity and PPV were 84.9%, 98.2% and 5.3%, respectively. The results for MMS are encouraging but no conclusions may be drawn until mortality data are processed in 2015.

1.5.4 Proteomics

Proteomics aims to identify protein patterns (signatures) or peptides in serum that may be specific to the disease in question. Advances in bioinformatics have contributed immensely to proteomics and a huge amount of research is ongoing. Although proteomics represents a promising area of research, extensive validation is required for any putative markers discovered. Reproducibility of some encouraging results has already proven to be problematic.[64, 65] Consequently, it may be some time before any proteomic biomarker for ovarian cancer reaches clinical practice.

1.6 Challenges to Early Detection

1.6.1 Dual Pathways of Oncogenesis

Potentially, there is an additional challenge in the detection of early stage ovarian cancer. Our understanding of the natural history of ovarian cancer is poor, and to date, no premalignant lesion has been identified (although this is a contentious issue). It is possible that disease progression in epithelial ovarian cancer is heterogeneous with some early stage tumours becoming large without regional or distal spread while some advanced stage malignancies may develop rapidly with little opportunity for 'early detection'.[66] Thus, some tumours that are commonly detected at early stages could be biologically distinct from their late stage counterparts with slower growth rates and a lower propensity to metastasize.[66] Observations that support this include a tendency for early stage carcinomas to have larger tumour size, better differentiation, non-serous histology, and to be more common in younger women.[9, 12] Specifically, only 3% of women with late stage disease were found to have a grade I tumour versus 54% of women with early stage disease.[12] Seventy-eight percent of early stage disease patients presented with a tumour >10 cm in diameter compared with only 32% of late stage patients. Attanucci *et al.* also found that tumour size was significantly larger in early stage disease (13.5 ± 6.8 cm versus 9.0 ± 6.3 cm, p=0.004), although the sample size was small (early n=33, advanced n=81). Additionally, several studies have reported longer symptom duration in borderline versus invasive tumours,[6, 67, 68] and early versus late stage disease.[69] One would expect the opposite if there was temporal relationship between symptom duration and stage of disease.

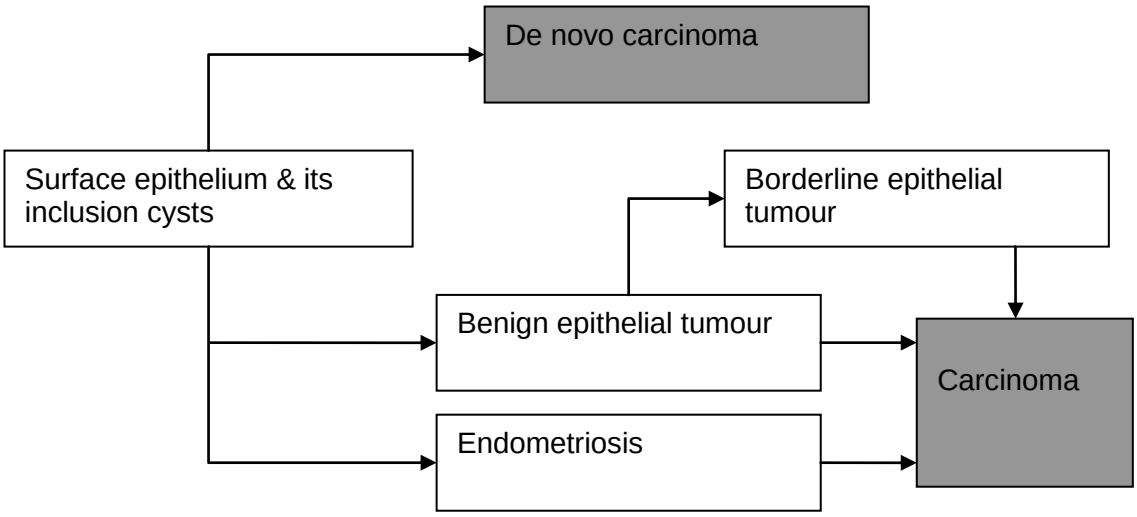
There are morphologic and molecular genetic data to support this theory. Each histological subtype in epithelial ovarian cancer (EOC) appears to undergo distinct changes and develop via different pathways.[70] Certainly, the most common subtypes of EOC are serous and undifferentiated and these are much more likely to be stage III-IV at diagnosis.[71] On the other hand, borderline, mucinous, endometrioid cancers are more likely to be confined to the ovaries at presentation are usually clinically detected at early stages and have a more favourable prognosis.[71]

Based on this evidence, it has been proposed that a significant proportion of high grade serous cancers arise *de novo* from the ovary surface epithelium or its inclusion cysts enabling rapid spread when only microscopic in size.[72] Conversely, low grade serous and mucinous tumours probably arise from within, or contiguous to, their benign or borderline counterparts and follow the classical stepwise adenoma-borderline-carcinoma sequence. Clear cell and endometrioid cancers appear to originate from

endometriosis[73] (see Figure 1-1). This theory of divergent pathways for oncogenesis challenges the presumption that temporally 'early' diagnosis will invariably translate into detection of disease while still confined to the ovaries.

Nonetheless, it is unlikely that *all* advanced stage cancers develop at such a rate that attempts to bring forward diagnosis are futile or impossible, and early detection via symptoms remains a critical area of investigation. It remains entirely plausible that some women are detected at advanced stages as a result of prolonged intervals between symptom onset and diagnosis.

Figure 1-1 Dual Pathways of Epithelial Ovarian Cancer Oncogenesis (adapted from Scully *et al.* 2000)[74]



1.6.2 Misdiagnoses

Due to the non-specific nature of the symptoms, ovarian cancer may be initially misdiagnosed as other diseases or conditions such as irritable bowel syndrome (IBS), diverticulitis, cholecystitis, and urinary tract infection (UTI) or cystitis.[8, 11, 16, 17, 75-78] A study performed in the UK found that 60% of women were initially sent to non-gynaecological hospital departments.[11] In the same study, half of the 16 women with a delayed referral of three months or longer were initially misdiagnosed with IBS. Survey results from 1725 self-selected women with ovarian cancer found that 30% believed that misdiagnosis was a barrier to receiving a prompt diagnosis.[8] Importantly, the same study found that initial misdiagnosis was associated with later stage disease.

Although misdiagnosis is regarded as physician failure, it should also be recognised that the average GP will see only one case every five years in the UK.[30] Hence, the rationale for initially suspecting other more common diseases is substantiated.

In the literature, IBS and diverticulitis have often been perceived as being associated with misdiagnosis of ovarian cancer.[10, 11, 16, 77] Hence, the proportion of women that had each disease (documented in the medical records) was examined for both of the studies in this thesis. Misuse of terms pertaining to diverticular conditions is common in both lay persons and medical professionals.[79] Diverticulum (diverticula plural) is a pouch or sac that bulges out from the colon that is thought to be caused by increasing intraluminal pressure required to eliminate faeces, particularly in low fibre diets. The condition of having diverticula is referred to as diverticulosis, which can be completely asymptomatic.[80] Diverticular disease is simply diverticulosis with symptoms which can include abdominal cramping, bloating, diarrhoea and constipation.[81] Diverticulitis is when diverticula become inflamed and in complicated cases these may become infected and perforate. For the purposes of this thesis, all of the terms for diverticular conditions were included as diverticulitis since we expected them to be used interchangeably in the medical records.

1.7 Current UK Referral Guidelines for Ovarian Cancer

Under current UK policy, anyone with suspected cancer should be seen by the appropriate specialist within two weeks of urgent GP referral.[82] An indication that this policy is being adhered to is that 70% of British women with ovarian cancer were seen in hospital within 2 weeks of GP referral (including urgent and non-urgent referral).[11] Although these results seem encouraging, the enduring perception amongst patients is that delays in ovarian cancer diagnosis are commonplace. A more recent UK study had less supportive findings, only 24% of women diagnosed with ovarian cancer had an urgent referral.[83]

According to the NICE (National Institute for Clinical Excellence) guidelines, abdominal palpation should be carried out in women presenting with 'any unexplained abdominal or urinary symptoms'.[84] Pelvic examination is recommended if there is 'significant concern' and if 'appropriate and acceptable to the patient'. The guidelines also stipulate that:

- Women with a palpable abdominal or pelvic mass should be sent for urgent ultrasound (if the mass is not obviously uterine fibroids or of gastrointestinal or urological origin) and urgent referral if the scan is suggestive of malignancy.
- Women with postmenopausal bleeding (if not on HRT) should be referred urgently.

The degree to which these guidelines are having a positive impact on the speed of ovarian cancer detection is unclear. One possible barrier is GP uncertainty with regard to urgent ultrasound and the decision to use rapid referral.[85]

1.8 Ovarian Cancer Treatment

Women with suspected ovarian cancer have a risk of malignancy index (RMI) calculated based on ultrasound results, menopausal status and CA125 levels. Those who have a RMI >200 are considered to be at high risk of disease and receive care under a gynaecologist oncologist. The standard treatment for all stages of ovarian cancer includes cytoreductive ('debulking') surgery, chemotherapy and/or radiotherapy. The specific combination of these treatments is dependent on prognostic factors such as stage, tumour grade and clear cell histology. FIGO (International Federation of Gynaecology and Obstetrics) staging is performed by surgery and this largely determines the need for further treatment (see Table 1-1).

Table 1-1 FIGO Staging[86]

Stage I	Tumour confined to the ovaries
IA	Tumour limited to one ovary; no tumour on external surface; capsule intact. No malignant cells in ascites or peritoneal washings
IB	As above, but tumour limited to both ovaries
IC	Tumour limited to one or both ovaries with any of the following: tumour on external surface; ruptured capsule; malignant cells in ascites or peritoneal washings
Stage II	Tumour involving one or both ovaries with pelvic extension
IIA	Extension and/or implants in uterus and/or fallopian tubes. No malignant cells in ascites or peritoneal washings
IIB	Extension to other pelvic organs. No malignant cells in ascites or peritoneal washings
IIC	Tumour staged either IIA or IIB with malignant cells in ascites or peritoneal washings
Stage III	Tumour involving one or both ovaries with microscopically confirmed peritoneal metastases outside the pelvis and/or regional lymph node metastases. Liver capsule metastasis equals Stage III.
IIIA	Microscopic peritoneal metastases beyond the pelvis
IIIB	Macroscopic peritoneal metastasis beyond the pelvis, none exceeding 2cm in diameter
IIIC	Peritoneal metastases beyond the pelvis greater than 2cm in diameter and/or regional lymph node metastasis
Stage IV	Distant metastasis

1.8.1 Treatment for Early Stage Disease

Primary surgery is first-line treatment for early stage disease. This typically involves total hysterectomy, bilateral salpingo-oophorectomy (removal of both ovaries and the fallopian tubes), omentectomy and removing as much of the cancerous tissue as possible (debulking or cytoreduction). Tumour grade refers to the degree of differentiation of the tumour cells, that is, the extent to which the cells resemble the tissue of origin. Grade 1 (well-differentiated) tumours have cells that closely resemble normal ovarian tissue, grade 2 (moderately differentiated) tumours have some cells that resemble the tissue of origin but display signs of frank malignancy, and grade 3 (poorly differentiated) tumours have cells that appear abnormal. In general, the better differentiated a tumour is, the less aggressive it behaves. Hence, patients with well-

differentiated stage IA and IB tumours are usually managed without adjuvant chemotherapy. Clear cell carcinomas are the exception to this rule since they are much more likely to recur (in comparison to other histological types), tend to have a poor response to chemotherapy and generally have a less favourable prognosis.[87] As such, usually all women with clear cell histology receive chemotherapy regardless of stage. Accurate surgical staging is critical for early stage disease since it determines whether or not the patient will require chemotherapy. If a woman is under-staged she may not receive potentially life-saving chemotherapy.

1.8.2 Treatment for Advanced Stage Disease

For ovarian malignancies of stage IC and higher, primary cytoreductive surgery is performed to minimise tumour burden before chemotherapy. In advanced stage disease, optimal debulking (i.e. leaving <1 cm of residual disease) is critical since it translates into a survival benefit.[88] Chemotherapy is usually administered after cytoreductive surgery however there is ongoing research to find out if treatment given in the reverse order (i.e. neoadjuvant chemotherapy) would be advantageous (EORTC 55971 and MRC-CHORUS study). Women with very advanced disease may also receive first-line chemotherapy as their disease may not be safely operable. Adjuvant chemotherapy for ovarian carcinoma typically involves six cycles of a platinum compound such as cisplatin or carboplatin with or without paclitaxel (Taxol®).

Nowadays, radiotherapy is rarely used, but it may be given to women with advanced disease to help control symptoms or to women with stage IC or II cancers who have no signs of malignancy in the abdomen and minimal deposits in the pelvis.

2 CHAPTER 2: Systematic Review of the Timing of Prediagnostic Ovarian Cancer Symptoms

2.1 Introduction

Over the past few decades, numerous studies on ovarian cancer symptoms have been carried out. Most of these set out to prove that women with ovarian cancer experience symptoms before diagnosis. Overall, the study designs and methods used for symptom studies have been inconsistent, but despite this heterogeneity, certain themes have emerged with regard to what types of symptoms are common prior to diagnosis. Specifically, the vast majority of women have at least one symptom, regardless of the data collection method (direct from women or medical records), including women with early stage disease.[6-10, 12-17, 19, 67, 68, 75, 78, 89-92] In 2005, a systematic review on ovarian cancer symptoms was published by the Primary Care group in Oxford.[93] This review provided a much needed summary of the evidence for symptoms associated with ovarian cancer diagnosis, however the timing of prediagnostic symptoms was not evaluated. Symptoms have been reported to be present for months or even years, prior to diagnosis, but the evidence to support these claims is extremely variable and needs clarification.

Although many studies have purported to examine prediagnostic symptoms, in reality, most have presented data that include a mixture of symptoms before *and* at diagnosis. This is partly due to the fact that 'symptoms before diagnosis' is likely to have been interpreted in various ways (e.g. symptoms at presentation, before referral, before biopsy results, or before final diagnosis). As such, there is a high probability that many of the symptoms reported, started just weeks or days before diagnosis, when most women would already be in the secondary or tertiary care system. The underlying goal of research on prediagnostic symptoms in ovarian cancer is to find out whether diagnosis can be bought forward by acting on symptoms. Therefore, the crucial period to quantify is the interval between symptom onset and referral, when there is potential to intervene by expediting referral, and thereby, diagnosis. In addition, symptom data are complex to collect and analyse. Many of the issues relating to these complexities have not been taken into account in the literature.

Using symptoms to distinguish between benign and malignant tumours or early and late stage disease has obvious appeal. Unfortunately, so far researchers have failed to demonstrate clear separation in symptom profile according to histology, tumour type or stage of disease. Borderline, benign and malignant masses appear to generate a

similar spectrum of symptoms.[12, 68, 75] Certainly, some symptoms appear more prominently within particular groups, however, there are no features which would accurately discriminate between groups.

This review will provide an updated overview of the studies performed to date, and it will suggest that symptoms are indeed present prior to diagnosis, however, the timing of these is still unclear. The available evidence for the duration of symptoms before diagnosis will be summarised, and the most common prediagnostic symptoms will be highlighted. Potential symptom discriminatory factors for stage, tumour type and presence or absence of ovarian cancer will be discussed in light of their relevance to this project. Emphasis will be on data from the population of women who are most likely to benefit from a symptoms-based diagnostic tool, i.e. women aged ≥45, with invasive epithelial ovarian cancer. Finally, some of the key methodological issues will be discussed.

2.2 Aims

2.2.1 Primary Aim

- To evaluate and summarise the existing evidence on the timing of symptoms before diagnosis of ovarian cancer.

2.2.2 Secondary Aims

- To provide an updated overview of symptoms prior to ovarian cancer diagnosis, with a focus on the most common symptoms.
- To summarise key methodological issues in ovarian cancer symptom studies.
- To summarise the existing evidence on symptoms associated with tumour status (benign, borderline, invasive) and stage of disease (early, late).

2.3 Methods

2.3.1 Search Strategy

A detailed electronic search using OVID MedLine® (1980 to September 2007) and CINAHL was carried out (see Table 2-1). It was difficult to obtain a specific search strategy, since the term ‘symptom’ commonly appears in journal articles of all types. The original search included the MeSH (medical subject heading) term SIGNS & SYMPTOMS however, this resulted in over 1000 additional papers so this was excluded in the final search.

Table 2-1 Search Strategy

1	Ovar$3
2	OVARIAN-NEOPLASMS#.DE.
3	(cancer$3 OR malignan$4 OR tumour$3 OR tumor$3 OR neoplasm$3 OR carcinoma$3 OR adenocarcinom$3).TI, AB.
4	Symptom$7
5	1 NEAR 3
6	3 OR 5
7	4 WITH 5

A total of 251 articles were retrieved using the search strategy (166 from Medline and 85 from CINAHL). Certain restrictions were imposed on the methods due to limited resources. Only papers written in English were considered. The SIGLE (System for Information on Grey Literature in Europe) database was not searched. However, the PhD thesis of Clare Bankhead[94] constituted one source of grey literature. Several abstracts relating to ovarian cancer symptoms from conference proceedings were assessed, but none were identified as relevant. No journals were hand-searched, however the bibliographies of retrieved papers in the final set were scoured for additional papers, as were several review articles.[68, 92, 93] This resulted in 5 additional papers[11, 95-98] that were not detected in the electronic database search. EMBASE was not searched since it is a more pharmaceutically-inclined database. The final update was performed in December 2008 which identified 7 additional papers.[21, 99-104] This resulted in a total of 264 articles to be assessed for inclusion.

2.3.2 Selection

The following criteria were used to select eligible studies for the review:

- Study design – Original research articles with quantitative symptom data. Case studies were excluded, as were papers focussed on a single symptom.
- Publication year between 1980 to present - Only recently published studies were of interest for several reasons; (1) a systematic review in this area was performed in 2005 and many of the same papers were included; (2) symptom duration/delays may have changed (3) the symptom spectrum in primary care may have changed; and (4) interest in symptoms as a tool for early diagnosis is a relatively new concept.
- Participants – women newly diagnosed with primary ovarian cancer (not recurrent disease). Studies focussed on women aged below 45 were excluded since younger women tend to have earlier stage disease with more favourable prognosis, non-serous histology, different symptoms and comprise less than 15% of women with ovarian cancer.[32, 33] Likewise, articles which focussed on non-epithelial

subtypes were excluded since these are rare and symptoms often different from those produced by epithelial ovarian cancer.

- Outcomes – Duration of symptoms before diagnosis of ovarian cancer (including data on delays in diagnosis), proportion of women reporting symptoms before diagnosis (individual and grouped symptoms), the most common symptoms reported by women prior to ovarian cancer diagnosis.
- Other – Exclude if not in English.

Two reviewers (Anita Lim; AL and Alejandra Castanon; AC) independently evaluated the eligibility of the studies and selected the relevant papers using information from the title and abstract. Any disagreements were discussed, which resulted in 43 papers to be assessed in greater detail. Fourteen studies[15, 18, 76, 77, 95, 99, 103, 105-111] were further excluded from this short-list during secondary review using the entire paper. Questions over study validity or any other aspect of the data were discussed on an ad-hoc basis with a third reviewer (Peter Sasieni, PS). Reasons for all exclusions are listed in Table 2-2.

Table 2-2 Reasons for Excluding Studies

First Review (based on title and abstract)
Case study (n=14)
Not original research (review, news article, letters etc.) (n=67)
Focus on recurrent disease (n=3)
Not about ovarian cancer (n=45)
Focus on Non-EOC disease (n=5)
Not about pre-diagnostic symptoms (n=84)
Study population younger than 45 (n=3)
Second Review (based on entire article)
Focus on a single symptom (n=2)
Symptoms data split by histological subtype (n=2)
Same data already reviewed in another source by the same author (n=1)
No data on symptoms from women with ovarian cancer (n=1)
Not a study (n=1)
Aim to assess or develop symptom index not evaluate prediagnostic symptoms (n=2)
Qualitative study (no quantitative data) (n=2)
Not clear where data were from (n=1)
Symptom data inadequate (n=1)
Insufficient case details (n=1)

2.3.3 Data Extraction & Validity

A total of 29 studies were eligible for final review.[6-14, 16, 17, 19-21, 67-69, 75, 78, 89, 90, 94, 96-98, 100, 101, 104, 112] Data extraction was performed unblinded by only one reviewer (AL) due to limited resources.

The studies were grouped according to data collection method (questionnaire/interview, medical record review, or coded insurance billing records) and design.

2.4 Results

2.4.1 Timing of Prediagnostic Symptoms

Establishing the time between symptom onset and diagnosis (or referral) is a critical step in any paradigm for promoting earlier diagnosis via symptoms. From this, it is possible to ascertain the potential lead time of symptoms. In the literature, the interval between symptom onset and diagnosis has been labelled as either symptom duration, or (total) delay. The term 'delay' implies that there has been unnecessary postponement of an event, whereas 'duration' merely suggests that the time the event spans over has been quantified. Although these terms are distinct, they are derived using the same dates (symptom onset date, date of diagnosis) and therefore, refer to the same interval.

Twenty-three[6-8, 10-14, 16-21, 67-69, 75, 78, 89, 94, 104, 112] of the studies identified in the original search contained information relating to the timing of symptoms. However, the measurement of symptom duration or delay was rarely a primary objective. Hence, the data are scant, variable and have been presented in ways which have made it difficult to perform any formal quantitative pooling or meta-analysis. Two of the studies will not be discussed since the data on duration were not useful (i.e. combined with data on benign tumours,[104] or split by histological type[18]).

Typically, duration data have been presented as mean or median time from symptom onset to diagnosis, or as the proportion of women with any symptom at predefined periods before diagnosis (e.g. <1 month, <3 months, >6months). The distribution of duration data is often skewed, thus, presenting a mean may be inappropriate. Often, it is not clear if calculations relating to duration and delay were restricted to symptomatic women, or if asymptomatic women were also included. Inclusion of asymptomatic women would unduly skew the data towards shorter duration or delay, since they would be designated a duration or delay of zero.

Another consideration is the possibility of contamination by symptoms that are coincidental rather than due to ovarian cancer. This is a particular problem in ovarian cancer due to the non-specific nature of the symptoms (e.g. abdominal pain, fatigue). Indeed, some of the data available on symptom duration and delay suggest suspiciously long periods of time.[6, 13, 112] Although it is impossible to know for certain which symptoms are truly attributable to ovarian cancer and which are not, it is important to acknowledge, and try to control for, this issue. Moreover, this has not been discussed in the ovarian cancer symptom literature.

Finally, duration and delays will differ depending on whether data are self-reported by women or extracted from medical records. In self-reported data, the accuracy of symptom onset dates is dependent on the women's ability to recall and is naturally subject to recall bias (if collected retrospectively). Further inaccuracies may stem from the women's perception of what constitutes a 'symptom'. For medical record data, recording error and recording bias can be a problem, also the visit date may be used rather than the true symptom onset date. Data obtained using these two methods can be vastly different and should be considered separately. The issues relating to this are discussed in more detail later.

As a corollary of the issues and inconsistencies described above, the tables relating to duration or delay may appear to be incomplete; however, the data provided are as comprehensive as possible. Table 2-3 and Table 2-4 provide a summary of the available data relating to time from symptom onset to ovarian cancer diagnosis (data presented in weeks were converted to months to allow comparability across studies). In the tables, duration for borderline and invasive disease were separated (whenever possible), since borderline symptom duration appeared to be longer.[6, 13, 67, 68] Also, both the total number of women and the number (or proportion) of symptomatic women were presented because symptom duration was often calculated for symptomatic women only. Finally, median duration was used to complete the proportion of women with >3 or >6 months duration, whenever possible (e.g. from a median of 4 months it can be deduced that more than 50% of women had a duration of >3 months).

Overall, the data were fairly heterogeneous, particularly with regard to the range of symptom duration. The largest range reported for any symptom was 0 to 8 years (medical record data).[13] Symptomatic ovarian cancer that is undiagnosed for 8 years is probably extremely rare and potentially, this range may have been exaggerated by the inclusion of symptoms that were present due to other diseases. The majority of studies have found that the mean or median symptom duration is less than 6 months.[6, 13, 67, 78, 89, 112] However, a median of 12 months has been reported by a US and a UK study.[7, 94] In general, the mean and median symptom duration for medical record studies were shorter than those derived from survey studies. Most medical record studies have shown that 50%-70% of women are diagnosed within 3 months of symptom onset. Questionnaire and interview study results are much more mixed. Collectively, the data on symptom duration are extremely heterogeneous and no strong conclusions can be drawn. It is unclear whether this heterogeneity stems from differences in methodology or differences in study population or country-specific differences.

Table 2-5 shows symptom duration for individual symptoms. No obvious patterns could be observed with regard to symptom duration and symptom type, or data collection method. Duration was very similar within each study, values were typically between 2 to 4 months.

Three studies have examined the timing of prediagnostic symptoms by dividing the months before diagnosis into periods and measuring symptom reporting in each period for cases versus controls.[10, 20, 21] According to these studies, symptoms of ovarian cancer were detectable within 12 months of diagnosis. Smith and colleagues found significantly higher case symptom frequencies on health insurance claims 10-12 months before diagnosis for abdominal swelling and pelvic pain.[20] Medical records from a prepaid health care system showed that abdominal pain was recorded significantly more in late stage cases (stage IC-IV) than controls (20% vs. 6%, respectively) 6-12 months prior to diagnosis.[10] In the same study, very few symptoms were documented in excess for early stage disease (IA-IB), although as with late stage disease, abdominal pain was the most common symptom 6 months prior to diagnosis. Wynn *et al.* found that case-control symptoms diverged 2-3 months (60-90 days) before diagnosis on health insurance claims.[21] Unfortunately, symptom reporting is likely to be underestimated by these studies as only indirect methods of data collection were used.

Symptom duration was a main study outcome in a study based in Israel.[112] The authors examined the effect of symptom duration on prognostic factors. No relationship between stage, grade, residual tumour and symptom duration was found. Unfortunately, the quality of the data was questionable; only 371 of 1005 (37%) cases had symptoms recorded in medical notes, and just 187 (50%) of these had symptom duration data. Therefore, it is unclear how reliable the data were.

Table 2-3 Symptom Duration for Any Symptom (i.e. Symptom Onset to Diagnosis) – Questionnaire or Interview Studies

Study	Data Source	Total No.	Symptomatic No.	Symptom Duration for Any Symptom (months)				
				Mean	Median	Range	>3m*	>6m*
Goff *et al.* [8] 2000 (USA)	Q	1725 (T) 385 (E), 1209 (L)	95% (T) 89% (E),97% (L)	4 (E) 5 (L)	-	-	70% (T)	35% (T)
Vine *et al.*[6] 2001 (USA)[a]	I	616 (IOC) 151 (BL)	556 (IOC) 127 (BL)	-	4 (IOC) 6 (BL)	1-99 (IOC) 1-99 (BL)	*>50% (IOC)* *>50% (BL)*	*<50% (IOC)* *50% (BL)*
Koldjeski *et al.*[78] 2003 (USA)	Q, I	19	19	3[b]	-	0-12[b]	-	-
Chan *et al.*[89] 2003 (HK)	Q	80 (T) 43 (E), 37 (L)	72 (T)	1[b] (T) 1[b] (E), 0.4[b] (L)	-	-	-	-
Vine *et al.*[7] 2003 (USA)	Q, I	200 (IOC) 67 (BL)	95% (T)	-	12 (T)	-	*>50%*	*>50%*
Webb *et al.*[68] 2004 (Australia) [a]	I	218 (E) 447 (L)	185 (E) 389 (L)	-	-	-	29% (E) 27% (L)	-
Bankhead C.[94] 2005 (UK)	Q, I	44	44	-	12	2-48	*>50%*	*>50%*

*Proportion of women diagnosed more than 3 or 6 months after (first) symptom onset
[a] Values calculated based on number of symptomatic women with invasive disease
[b] Values converted from weeks to months by multiplying by 7, dividing by 30.4375 and rounding to the nearest month (rounded to 1 decimal place if less than 1 month)
Q=Questionnaire, I=Interview, IOC=Invasive Ovarian Cancer, BL=Borderline, T=Total, E=Early Stage, L=Late Stage
Note: Italicised values are deduced from the median

Table 2-4 Symptom Duration for Any Symptom (i.e. Symptom Onset to Diagnosis) – Medical Record Studies

Study	Data Source	Total No.	Symptomatic No.	Symptom Duration for Any Symptom (months)				
				Mean	Median	Range	>3m*	>6m*
Eltabbakh *et al.*[67] 1999 (USA)[a]	MR	50 (IOC) 22 (BL)	41 (IOC) 15 (BL)	3 (IOC) 8 (BL)	-	0-14 (IOC) 1-24 (BL)	-	24% (IOC+BL)
Yawn *et al.*[16] 2004 (USA)	MR	107	91	-	-	0-18	-	13%[a]
Thulesius *et al.*[14] 2004 (Sweden)[b]	MR	99	74	-	2	-	-	-
Lataifeh *et al.*[12] 2005 (Australia)[c]	MR	200 (T) 100 (E), 100 (L)	190 (T) 90 (E), 100 (L)	-	-	-	30% (T) 30% (E), 31% (L)	-
Paulsen *et al.*[13] 2005 (Norway)[d]	MR	486	349	-	2	0-105	-	-
Menzcer *et al.*[112] 2008 (Israel)	MR	1005	187	-	2	1-60	-	-

*Proportion of women diagnosed more than 3 or 6 months after (first) symptom onset
[a] Proportion of all women (i.e. symptomatic and asymptomatic).
[b] Values converted from weeks to months by multiplying by 7, dividing by 30.4375 and rounding to the nearest month
[c] Unclear if proportions are calculated based on all women or symptomatic women only
[d] Symptom onset to date of first hospitalisation
MR=Medical Records, IOC=Invasive Ovarian Cancer, BL=Borderline, T=Total, E=Early Stage, L=Late Stage

Table 2-5 Symptom Duration for Individual Symptoms

Study	Data Source	n	Mean or Median Symptom Duration (months)					
			Abdominal Pain	Abdominal Distension	Bloating	Fatigue	Bowel Irregular	Urinary
Vine *et al.*[6] 2001 (USA)[a]	I	616	2[f]	-	-	-	3	3
Olson *et al.*[69] 2001 (USA)[b]	Q, I	168 (T) 37 (E) 118 (L)	6[d] (T) 8[d] (E) 6[d] (L)	-	5[e] (T) 6[e] (E) 5[e] (L)	6 (T) 7 (E) 5 (L)	6/5[c] (T) 5/8[c] (E) 6/4[c] (L)	6 (T) 7 (E) 5 (L)
Vine *et al.*[7] 2003 (USA)[a, g]	Q, I	200	2-4	2-4	-	5-7	5-7	5-7
Goff *et al.*[75] 2004 (USA)[a]	Q	44	4	3	-	3	3/3[c]	4
Paulsen *et al*[13]. 2005 (Norway)[a, h]	MR	486	2	2	-	3[i]	3	2

[a] Median
[b] Mean
[c] Constipation/Diarrhoea
[d] Unusual abdominal or lower back pain
[e] Unusual bloating, fullness and pressure in the abdomen or pelvis
[f] Pelvic discomfort
[g] Values were only given as a range in the paper
[h] First symptom experienced to first hospitalisation
[i] Persistent fatigue or weight loss
Q=Questionnaire, I=Interview, MR=Medical Records, T=Total, E=Early Stage, L=Late Stage

Diagnostic Delays

As aforementioned, delays are inherently linked with symptom duration and lead time since they are both derived from the same information; and it is only the way in which the data are interpreted that differs. Although total delay is the same as symptom duration, it is pertinent to discuss the possible causes of delay and to closer examine the different components of delay (i.e. patient versus provider delay). The term 'delay' carries negative connotations which imply that there is an avoidable prolongation of events. Indeed, the investigation of diagnostic delays seeks to identify the length of time that diagnosis can be feasibly brought forward by the eradication of *unnecessary* delays. However, there are certain processes between symptom onset and diagnosis that are valid and/or mandatory steps in the diagnostic pathway. These may include the perfectly reasonable (patient) decision to wait several days or weeks before consulting for vague symptoms, and the equally sensible GP approach of 'watching and waiting' after initial presentation for vague symptoms (unless symptoms are specific e.g. abdominal distension). Also, the processing of CA125 tests, performance of ultrasound, and time taken to obtain a specialist appointment and surgery date, all require a minimum passing of time. Hence, the labelling of the interval between symptom onset and diagnosis as 'delay' is inaccurate and misleading. All of these factors should be considered when interpreting delay data. Moreover, it would be prudent to devise and standardise a better term to describe the interval in question. For the purposes of this thesis, the term 'delay' will be employed to keep consistency with the literature.

Diagnostic delays can be divided into patient delay (ignoring or misinterpreting symptoms, putting off GP visits) and provider delay (misdiagnosis, referral waiting times, omission of appropriate tests). Crucially, delays and women ignoring symptoms have been associated with late stage disease.[8] The calculation of delays is subject to the same limitations as the derivation of symptom duration data. As such, existing delay data are as sporadic as duration data.

Patient Delay

Patient delay is usually defined as time from symptom onset (i.e. date that the woman becomes aware of the symptom) to first clinic visit for symptoms. Only limited data on patient delay are available. Several studies have found that most women (~80%) presented to clinic within 3 months of symptom onset (see Table 2-6).[11, 12, 17, 19, 75] Mean and median patient delay was 1 month or less for most studies according to both survey and medical record data. The largest study reported a longer delay however, only a range was presented for the median (2-3 months) which is unusual, and

explanatory data were not provided.[8] Two papers have reported a very short median/mean time from symptom onset to presentation.[11, 89] The first, was a UK GP record study which found that 78% of women presented to the GP within four weeks of symptom onset.[11] Interestingly, this study found that 18-month survival was not affected by delays however the authors did not adjust for stage in the analysis.[113] The second, was a Hong Kong-based study which found that according to self-reported data, most women sought medical help within two weeks from symptom onset.[89]

Overall, the literature suggests that most women with ovarian cancer present to a doctor within 1 month of symptom onset (see Table 2-6). However, approximately a third of women have a delay of 3 months or more, which may be an area for improvement. The distribution of delay in this 30% of women requires further elucidation.

Provider Delay

Provider delay is usually defined as time from first presentation to a doctor for a symptom to date of diagnosis. This period covers both health system delays (waiting times for test results, imaging, referral appointments) and time taken for physicians (primary care and specialists) to make the diagnosis. Identifying which was the first critical consultation for ovarian cancer symptoms is not straightforward and can vary depending on the definitions and methods used to derive this date. In addition, deciding which symptoms were present and due to ovarian cancer can be highly subjective. Provider delays may be underestimated in the literature by studies that use medical records as the sole data source, since many symptoms reported by patients go unrecorded. On the other hand, one might reason that provider delay is more accurate than patient delay since all the required dates are recorded in the medical notes and do not depend on retrospectively recalled data from women. Date of diagnosis was not clearly defined in most studies. If date of diagnosis was self-reported from women, it is likely to differ from those taken from medical notes since this could be interpreted as the date they were first told they might have cancer as opposed to date of surgery or positive biopsy.

Table 2-6 shows that approximately 60%-70% of women are diagnosed within 3 months of presentation. As with patient delay, mean and median provider delay appears to be 1 month or less for most studies, regardless of data collection method. Once mandatory diagnostic procedures are accounted for, these already short provider delays would become even more trivial. However, again it is important to note that there are still around 30%-40% of women with symptoms that are diagnosed more than 3 months after first presentation to a doctor.

The diverse methodology and forms of data presentation make it difficult to assess whether provider and patient delays are comparable across studies. However, within studies it seems that patient and provider delay delays are similar (see Table 2-6, note that delays in weeks were converted to months).[14, 75, 89]

Collectively, the data imply that most women with ovarian cancer have a relatively short patient or provider delay. However, roughly a third of women have a patient or provider delay of >3 months. In addition, there is a tendency for delays in early versus late stage, and borderline versus invasive disease, to be longer (Table 2-6).

Table 2-6 Patient and Provider Delay

Study	Data Source	n	Patient Delay (months)			Provider Delay (months)		
			Mean	Median	>3m*	Mean	Median	>3m*
Smith *et al.*[17] 1985 (USA)	Q, I	82	-	1	23%	-	-	-
Flam *et al.*[19] 1988 (Sweden)	I	172 (E) 190 (L)	-		14% (E) 23% (L)	-	-	12% (E) 20% (L)
Goff *et al.*[8] 2000 (USA)	Q	1725	-	2-3	-	-	-	45%
Vine *et al.*[6] 2001 (USA)	I	616	-	-	-	-	1	25%
Vine *et al.*[7] 2003 (USA)	Q, I	65 (E) 135 (L)	-	4 (E) 2 (L)	-	-	3 (E)	-
Chan *et al.*[89] 2003 (HK) [a]	Q	80 (T) 43 (E), 37 (L)	0.4 (T) 0.3 (E), 0.4 (L)	-	-	0.2 (T) 0.2 (E), 0.1 (L)	-	-
Webb *et al.*[68] 2004 (Australia) [a]	Q, I	665 (IOC) 146 (BL)	-	0.9 (IOC+BL)	-	-	0.5 (IOC+BL)	-
Goff *et al.*[75] 2004 (USA)	Q	44	-	-	40%	-	-	-
Kirwan *et al.*[11] 2002[b] (UK)	MR	102	-	-	22%	-	-	27%
Thulesius *et al.*[14] 2004 (Sweden)[a]	MR	99	-	1	-	-	1	-
Paulsen *et al.*[13] 2005 (Norway)	MR	486	-	-	55%	-	-	39%

*Proportion of women diagnosed more than 3 months patient or provider delay

[a] Values converted from weeks to months by multiplying by 7, dividing by 30.4375 and rounding to the nearest month (rounded to 1 decimal place if less than 1 month)

[b] Time from first presentation to referral

Q=Questionnaire, I=Interview, MR=Medical Records, IOC=Invasive Ovarian Cancer, BL=Borderline, T=Total, E=Early Stage, L=Late Stage

2.4.2 Updated Overview of Ovarian Cancer Symptoms

In the current review, a total of 29 papers were evaluated. Eleven studies were based on medical record data,[9-14, 16, 67, 96, 97, 100] 14 used self-reported data,[6-8, 17, 19, 68, 69, 75, 78, 89, 90, 94, 104] and 3 utilised coded insurance billing records.[21, 92, 101] Two further papers derived data from a mixture of medical records and interview.[98, 112] Not all of the studies that were evaluated in the previous systematic review were included in the current review (and vice-versa), as the aims were not the same.

A meta-analysis from the previous systematic review showed that the majority of women with ovarian cancer have symptoms before (or at) diagnosis (93% and 77% according to self-reported and medical record data, respectively).[93] Women who are reported as 'asymptomatic' are presumably either diagnosed incidentally while undergoing examinations or procedures for other morbidities, or detected via some sort of screening procedure. Women who present acutely as emergency cases may also have been classified as 'asymptomatic', however this is not quite correct since they must have had symptoms to prompt the emergency visit. Unfortunately, most authors have not provided details of the classification, thus the true proportion of asymptomatic women is not clear.

Table 2-7 to Table 2-10 summarise the studies performed to date in terms of the participants, selection of cases and controls, data collection method, symptom elicitation method, period over which symptoms were asked about, and the three most common (sensitive) symptoms reported by ovarian cancer cases. All of the studies used retrospective data (i.e. the data collected pertained to events that had occurred in the past). Some of the studies included borderline and benign tumours in their definition of 'cases', however for simplicity, some of the data in the tables were limited to invasive disease.

Abdominal/pelvic pain or discomfort, abdominal distension/swelling and bloating are the most commonly reported symptoms regardless of the data collection methods (see Table 2-7 to Table 2-10).[7, 20, 69, 75, 101] Other symptoms of note include fatigue, indigestion, change in bowel habit, constipation, urinary symptoms, loss of appetite, nausea, palpable mass, weight loss, and irregular vaginal bleeding.[6-8, 15-19, 67-69, 75, 89, 91] In general, the comparability across studies is limited by grouping of symptoms (discussed further in methodological issues section). Several studies[9, 19, 20, 68, 112] have presented gastrointestinal (GI) symptoms as a group, which makes it impossible to assess sensitivity of individual GI symptoms. Also, the classification of GI symptoms may have differed between studies, and it is not clear in some studies what was

included under this heading. Nevertheless, it is clear that abdominal and gastrointestinal symptoms predominate in ovarian cancer. Gynaecological symptoms are among the least frequently reported symptoms in most studies.[8, 12, 15, 17-19] Presumably, this is due to the fact that most women develop disease after menopause when the ovaries lack physiological function. Other symptoms that have been associated with ovarian cancer include back pain,[8, 10, 11, 19, 75, 90, 98, 104] leg swelling,[75, 90, 98] pain with intercourse,[7, 8, 104] difficulty breathing,[7, 13, 14, 17, 19, 90, 104] and diarrhoea.[17, 69, 75] However, these are reported less frequently (less than <50% of women).

As expected, symptom prevalence is much lower in medical record and coded insurance bill studies in comparison to questionnaire/interview studies. However, the overall symptom spectrum remained similar for all of the data sources. Similarly, no patterns were observed for studies that examined the 12 months before diagnosis versus studies that evaluated no specific period.

Table 2-7 Questionnaire & Interview Studies

STUDY	STUDY DESIGN	PARTICIPANTS	PERIOD COLLECTED	MOST COMMON CASE SYMPTOMS
Smith & Anderson[17] 1985 (USA)[a]	RCS, Q, I, SR	82 OC cases (36 local, 46 distant) from population-based cancer registry Age 20-54	Symptoms that prompted healthcare-seeking. Mean time from diagnosis: 10 weeks, range 1-3 months.	Abdominal swelling (70%) Fatigue (56%) Abdominal pain (48%)
Flam *et al.*[19] 1988 (Sweden)[b]	RCS, I, unclear if CL or SR	362 OC cases (172 stage IA-IIA,190 stage IIB-IV) referred to a specific gynaecology department. Age not given	Women questioned prior to admission about symptoms before treatment.	Abdominal swelling (27%/24%) GI symptoms (15%/24%) Abdominal pain (17%/11%)
Igoe *et al.*[90] 1997 (USA)[c]	RCS, Q, CL, SR	50 invasive & BL OC identified via the internet. Age 35-75	Women asked about symptoms preceding diagnosis. Median time from diagnosis: 20 months	Fatigue (82%) Abdominal swelling (78%) Indigestion (72%)
Goff *et al.*[8] 2000 (USA & Canada)	RCS, CL	1725 OC cases (385 stage I-II, 1209 stage III-IV, 131 unknown) identified from mailed survey Age 18-84	Symptoms before diagnosis. Median time from diagnosis: 24 months	Increased abdominal size (61%) Abdominal bloating (57%) Fatigue (47%)

[a] Percentages shown are for proportion of symptomatic women
[b] Percentages shown are split by early stage/late stage
[c] Percentages shown are for invasive and borderline
RCS=Retrospective Case Series, CC=Case-control, Q=Questionnaire, I=Interview, CL=Checklist, SR=Spontaneous reporting, OC=Ovarian Cancer, BL=Borderline

STUDY	STUDY DESIGN	PARTICIPANTS	PERIOD COLLECTED	MOST COMMON CASE SYMPTOMS
Olson *et al.*[69] 2001 (USA)	CC, Q, I, CL, SR	138 OC cases (37 stage I-II, 118 stage III-IV) identified at two hospitals, 251 population controls. Age ≥18	Symptoms present in the 6-12 months before diagnosis. Mean time from diagnosis: 4.7 months	Unusual bloating, fullness & pressure in abdomen/pelvis (71%) Unusual abdominal or lower back pain (52%) Unusual lack of energy (43%)
Vine *et al*[6]. 2001 (USA)[a]	RCS, Q, I, CL, SR	767 OC cases (616 invasive, 151 BL) from existing CC study. Age 20-69	Symptoms before diagnosis. Diagnosed within last 6 months.	Pelvic discomfort (71%) Bowel irregularity (47%) Urinary frequency/urgency (37%)
Vine *et al*[7]. 2003 (USA)[b]	CC, I, CL, SR	267 EOC cases (65 stage I-II, 135 stage III-IV, 67 BL), 317 controls from existing population-based CC study. Age 20-74	Time from diagnosis median: 4.6 months. Symptoms present for at least 2 weeks during year before diagnosis.	Pelvic/abdominal discomfort/pain (64%) Bloating or feeling of fullness (62%) Distended/hard abdomen (59%)
Chan *et al*[89]. 2003 (HK)	RCS, I, SR	80 EOC cases (43 stage I-II, 37 stage III-IV) from a single gynaecology department. Age 18-70	Interviewed before treatment. Symptoms before diagnosis.	Abdominal pain/discomfort (26%) Abdominal distension (25%) Menstrual symptoms (15%)

[a] Percentages shown are for invasive cases only

[b] Cases could also report symptoms on the checklist of any duration

RCS=Retrospective Case Series, CC=Case-Control, I=Interview; CL=Checklist; SR=Spontaneously Reported, EOC=Epithelial Ovarian Cancer, OC=Ovarian Cancer, BL=Borderline

STUDY	STUDY DESIGN	PARTICIPANTS	PERIOD COLLECTED	MOST COMMON CASE SYMPTOMS
Koldjeski *et al.*[78] 2003 (USA)	RCS, I, Q, CL, SR	19 OC cases referred from regional cancer clinics (6 stage I-II, 13 stage III-IV). Age 28-73	Time from diagnosis: 2-3 weeks. Prediagnostic symptoms.	Bloating (84%) Vague abdominal pain (68%) Indigestion (63%)
Webb *et al.*[68] 2004 (Australia)[a]	CC, Q, I, SR	811 EOC cases (218 stage I-II, 447 stage III-IV, 146 BL) from an existing CC study. Age 18-79	Symptoms that first prompted doctor visit. Maximum of 4 different symptoms could be reported.	Abdominal pain/pressure (44%) Abdominal swelling/tightening (39%) GI (15%)
Goff *et al.* [75] 2004 (USA)[b]	CC, CL	128 ovarian tumours (44 invasive EOC cases [11 stage I-II, 33 stage III-IV], 74 benign, 10 BL), approached preoperatively at 2 cancer centres. 1011 controls visiting 2 primary care clinics. Age 15-90	Cases completed survey before definitive diagnosis. Symptoms in 12 months before diagnosis.	Increased abdominal size (64%) Bloating (70%) Fatigue (61%)

[a] Percentages shown are for proportion of symptomatic women
[b] Percentages shown are for invasive cases only
RCS=Retrospective Case Series, CC=Case-Control, I=Interview, CL=Checklist, SR=Spontaneously Reported, EOC=Epithelial Ovarian Cancer, BL=Borderline,

STUDY	STUDY DESIGN	PARTICIPANTS	PERIOD COLLECTED	MOST COMMON CASE SYMPTOMS
Olsen *et al.*[104] 2007 (Australia)[a, b]	RCS, I, SR, CL	244 invasive OC cases (89 stage I-II, 155 stage III-IV) 62 BL, 151 benign, identified via gynaecology cancer centre. Age 18-79	Time from diagnosis: median 12 months. Symptoms before diagnosis and at presentation.	Abdominal pain/pressure (56%/58%) Abdominal swelling (43%/60%) Gas/nausea/indigestion (34%/44%)
Bankhead C.[94] 2005 (UK)[c]	O, I, SR	44 cases (40 OC, 2 PP, 2 other gynaecological cancers), 80 non-cancer cases (59 benign, 21 normal findings), all referred to hospital for suspected ovarian malignancy. Mean age 59 (cases), 48 (non-cancer cases)	Interviewed before definitive diagnosis or soon after diagnosis. Symptoms prior to interview	Abdominal distension (±Bloating) (86%) Malaise (73%) Abdominal pain (59%)

[a] Percentages shown are for proportion of symptomatic women
[b] Percentages shown are split by early stage/late stage
[c] Percentages shown are for the 44 cases only
RCS=Retrospective Case Series, CC=Case-Control, O=Observational, I=Interview, CL=Checklist, SR=Spontaneously Reported, OC=Ovarian Cancer, BL=Borderline, PP=Primary Peritoneal

Table 2-8 Medical Record Studies

STUDY	STUDY DESIGN	PARTICIPANTS	PERIOD COLLECTED	MOST COMMON CASE SYMPTOMS
Kennedy & Gordon[97] 1981 (UK)	RCS, HR	97 OC cases from a single hospital. Age 11-83	Presenting symptoms	Abdominal pain (34%) Abdominal distension (21%) Abdominal mass (10%)
Piura *et al.*[96] 1998 (Israel)[a]	RCS, HR	52 OPSC Stage III-IV, 15 PPSC from a single hospital Mean age OPSC 56 Mean age PPSC 62	Presenting signs and symptoms	Abdominal mass (79%) Ascites (39%) Pleural effusion (12%)
Eltabbakh *et al.*[67] 1999 (USA)[a]	RCS, HR	50 stage I-II IOC cases, 22 BL, identified from cancer registry. Age 16-89	Presenting symptoms	Abdominal/pelvic pain (38%) Increased abdominal girth/bloatedness (28%) Vaginal bleeding (22%)
Kirwin *et al.*[11] 2002 (UK)	RCS, GP	135 EOC cases (38 stage I-II, 78 stage III-IV, 19 unstaged) from audit of Mersey area. Age Not given	12 months before diagnosis	Abdominal pain (48%) Change in bowel habit (25%) Abdominal swelling (19%)

[a] Percentages are for invasive ovarian cancer cases only
RCS=Retrospective Case Series, HR=Hospital Records, GP=GP Records, EOC=Epithelial Ovarian Cancer, IOC=Invasive Ovarian Cancer, BL=Borderline, OPSC=Ovarian Papillary Serous Carcinoma, PPSC=Peritoneal Papillary Serous Carcinoma

STUDY	STUDY DESIGN	PARTICIPANTS	PERIOD COLLECTED	MOST COMMON CASE SYMPTOMS
Attanucci *et al.*[9] 2004 (USA)[a, b]	CC, HR	114 IOC cases, 33 BL, 76 benign identified via tumour registry board. Age 15-86	Symptoms recorded at initial consultation before surgery.	Pain (57%) GI (54%) Constitutional [c] (43%)
Thulesius *et al.*[14] 2004 (Sweden)	RCS, GP, HR	99 OC (41 stage I-II, 58 stage III-IV) identified via regional tumour registry. Mean age 63	12 months before diagnosis	Urinary & GI (compression) (38%) Abdominal pain (38%) Abdominal swelling (33%)
Yawn *et al.*[16] 2004 (USA)[d]	RCS, GP, HR	107 IOC (42 stage I-II, 62 stage III-IV, 3 unstaged) identified via diagnostic database for Olmsted County Age 30-98	Two years before diagnosis	Abdominal pain (38%) Bloating/increased abdominal girth (13%) Urinary symptoms (13%)
Lataifeh *et al.*[12] 2005 (Australia)	RCS, HR	200 invasive EOC cases (100 stage IA-IB, 100) stage IIIC) from a tertiary referral centre for gynaecological cancer. Age not given	Presenting symptoms	Abdominal pain (48%) Abdominal swelling (47%) Abdominal bloating (12%)

[a] Percentages are for invasive ovarian cancer cases only
[b] Percentages calculated by Anita Lim
[c] Fatigue, fever, weight loss, weight gain
[d] Primary peritoneal cancers were included
RCS=Retrospective Case Series, HR=Hospital Records, GP=GP Records, OC=Ovarian Cancer, IOC=Invasive Ovarian Cancer, EOC=Epithelial Ovarian Cancer, BL=Borderline

STUDY	STUDY DESIGN	PARTICIPANTS	PERIOD COLLECTED	MOST COMMON CASE SYMPTOMS	
Friedman *et al.*[10] 2005 (USA)[a]	RCS, GP, HR	102 OC cases, (33 stage IA-IB, 69 stage IC-IV), 102 age-matched controls. All women enrolled on Kaiser Permanente Medical Care Program. Age 29-87	Two years prior to diagnosis divided into 0-6, 6-12, 12-24 months	Abdominal pain (55%) Fatigue/weakness/lack of energy (39%) Headache (38%)	Anytime (2 years)
				Abdominal Pain (46%) Abdominal bloating/ fullness/pressure (23%) Nausea (20%)	6 months before diagnosis
				Abdominal pain (20%) Nausea (15%)	1 year-6 months before diagnosis
Paulsen *et al.*[13] 2005 (Norway)[b]	RCS, HR	623 EOC (486 invasive, 137 BL) identified via cancer registry. Age 25-94	Data are from notifications to cancer registry and medical records	Abdominal pain/discomfort (53%) Distended/tense abdomen (44%) Bowel irregularity (26%) Persisting fatigue or weight loss (26%)	
Deligdisch *et al.*[100] 2007 (USA & France)	RCS, HR	76 stage I IOC cases. Age not given	Symptoms at presentation	Symptomatic pelvic mass (46%) Vaginal bleeding (28%) Asymptomatic pelvic mass (21%)	

[a] Stages IC-IV only
[b] Percentages are for invasive cases only
RCS=Retrospective Case Series, HR=Hospital Records, GP=GP Records, IOC=Invasive Ovarian Cancer, EOC=Epithelial Ovarian Cancer, OC=Ovarian Cancer, BL=Borderline

Table 2-9 Interview and Medical Record Data Studies

STUDY	STUDY DESIGN	PARTICIPANTS	PERIOD COLLECTED	MOST COMMON CASE SYMPTOMS
Beck *et al.*[98] 2001 (USA)[a]	RCS, I, HR	52 early stage (I-II) OC cases operated on by the lead author. Age 30-85.	Time from diagnosis: median 62 months. Symptoms at presentation from medical records (n=52), and interview with those still alive (43/46).	Pain (abdominal or pelvic) (35%) Increased abdominal girth (33%) Vaginal bleeding (15%)
Menczer *et al.*[112] 2008 (Israel)[b]	I, unclear if CL or SR, HR	1005 EOC cases identified from existing CC study.	Symptoms at presentation	Abdominal pain (65%) Abdominal swelling (35%) GI (27%)

[a] Unclear if percentages relate to combination of interview and medical record data or medical record data alone
[b] Only 37% had symptoms recorded
RCS=Retrospective Case Series, CC=Case-Control, I=Interview; HR=Hospital Records, CL=Checklist; SR=Spontaneously Reported, OC=Ovarian Cancer, EOC=Epithelial Ovarian Cancer

Table 2-10 Coded Insurance/Billing Claim Records

STUDY	STUDY DESIGN	PARTICIPANTS	PERIOD COLLECTED	MOST COMMON CASE SYMPTOMS	
Smith *et al.*[114] 2005 (USA)[a]	CC. Predetermined symptom list, ICD-9 codes	1895 OC cases (287 stage IC-II, 1453 stage III-IV, 245 unstaged), 10941 breast cancer controls, 6024 cancer-free controls. All women enrolled in Medicare in California. Age 68-101	36 months before diagnosis, divided into quarters starting at 1-3 months prior to diagnosis	Abdominal pain (31%) Abdominal swelling (17%) GI (8%)	1-3 Months
				Abdominal pain (9%) Abdominal swelling (2%) GI (2%)	4-6 Months
Ryerson *et al.*[101] 2007 (USA)[b]	RCS, Predetermined symptom list, ICD-9 codes	3250 OC cases (601 stage I-IB, 196 stage IC-II, 2453 stage III-IV) from the SEER-Medicare database. Age 65-98	Claims with symptoms within 12 months of diagnosis	Abdominal pain and tenderness (49%) Abdominal or pelvic swelling (43%) Constipation, diarrhoea, other digestive disorders (18%)	
Wynn *et al.*[21] 2007 (USA)[c]	CC. Predetermined symptom list, ICD-9 codes	920 OC cases, 2760 matched controls from a national claims and encounters database (MarketScan). Median age 59	National claims and encounters database. Symptoms within 9 months of diagnosis	Abdominal symptoms (36%) Chest/respiratory symptoms (17%) Urethra/urinary tract disorders (13%)	270-31 days before diagnosis

[a] Medicare provider claims linked to California SEER database
[b] Medicare claims linked to SEER cancer registries
[c] Thomson Healthcare's Medstat MarketScan Commercial Claims & Encounters & Medicare Supplemental Databases.
RCS=Retrospective Case Series, CC=Case-Control, OC=Ovarian Cancer, SEER=Surveillance, Epidemiology, and End Results

2.4.2.1 Case-Control Studies

Five case-control studies have been published comparing women with ovarian cancer to women without ovarian cancer.[7, 10, 20, 69, 75] Notably, the Vine and Olson studies used healthy volunteers which can overinflate case-control differences via a 'healthy volunteer effect'.[7, 69] The remaining 3 studies used women seeking healthcare.[10, 20, 75] Overall, the main findings include that women with ovarian cancer have symptoms of shorter duration (i.e. recent onset), greater severity and higher frequency in comparison to cancer-free women.[7, 10, 20, 69, 75] Women with ovarian cancer also experienced a higher number of symptoms.[7, 69, 75]

As expected, larger odds ratios (OR) were observed in the studies using controls who were not actively seeking healthcare (see Table 2-11).[7, 20, 69, 75] The highest odds ratios obtained for three of the case-control studies were for abdominal distension.[7, 20, 75]

The proportion of controls that reported symptoms was relatively high in some of the studies that collected data directly from women.[69, 75] In one study, 38% reported bloating and 30% had abdominal pain.[75] Conversely, symptom reporting was much lower in studies using medical records or coded insurance claims.[10, 20]

Table 2-11 Frequency & Odds Ratios for Common Symptoms in Case-Control Studies

	Olson *et al.*[69] 2001 (USA)		Vine *et al.*[7] 2003 (USA)		Goff *et al.*[75] 2004 (USA)		Smith *et al.*[20] 2005[*] (USA)[e]		Friedman *et al.*[10] 2005 (USA)	
Cases (Ca)	168		267[b]		44		1985		102	
Controls (Co)	251		317		1011		6024		102	
Symptom	% Ca/Co	OR (95%CI)	% Ca/Co	OR (95%CI)	% Ca/Co	OR (95%CI)	% Ca/Co	OR (95%CI)	% Ca/Co	OR (95%CI)
Abdominal Pain	52/15[a]	6.2 (4.0, 9.6)	64/10[c]	16.2 (10.3, 25.3)	50/30	2.3 (1.2, 4.4)	31/4	6.2 (5.2, 7.4)	55/19	8.9 (4.0, 20.3)
Abdominal Distension	-	-	59/5[d]	29.2 (16.5, 51.8)	64/19	7.4 (3.8, 14.2)	17/0[f]	39.2 (22.5, 68.1)	-	-
Bloating	71/9	25.3 (15.6, 40.9)	62/10	14.6 (9.4, 22.8)	70/38	3.6 (1.8, 7.0)	-	-	30/6[g]	6.4 (2.6, 26.3)

[a] Unusual abdominal or lower back pain
[b] Includes borderline and invasive cases
[c] Pelvic/abdominal discomfort/pain
[d] Distended or hard abdomen
[e] 1-3 months before diagnosis
[f] Abdominal or pelvic swelling or mass
[g] Abdominal bloating, fullness, pressure

Severity & Frequency

It has been suggested that subtleties in symptom frequency and severity may have a critical role in facilitating earlier diagnosis via symptoms.[69, 75, 94] Indeed, ovarian carcinomas have been shown to give rise to symptoms that are persistent and frequent.[69, 75] Only one study has asked women about symptom severity directly. They found that patients with ovarian malignancy had significantly more severe symptoms in comparison to women with benign masses and other clinic patients.[75]

However, severity and frequency of symptoms are liable to change over time and with disease progression. Also, the two parameters may not be independent, and perceived severity probably intensifies with increased frequency. The definition of what constitutes a single episode of symptoms can vary with different symptoms, and the frequency of symptoms can be irrelevant for some symptoms. For example, a single episode of postmenopausal bleeding can be of greater concern than four episodes of nausea or vomiting, and both could occur over a single day. As such, comparability of frequency data for different symptoms can be limited. Similarly, asking about the severity of an abdominal lump is vague and potentially irrelevant. Thus, the collection and interpretation of severity and frequency data requires careful consideration. Ultimately, to be a useful discriminatory factor, differences in symptom severity and frequency would need to be apparent relatively early on in symptom onset to provide a sufficient window of opportunity to intervene. Although, potentially worsening of severity and/or frequency could be useful.

Number of Symptoms

The number of symptoms experienced was significantly greater in ovarian cancer compared to controls in two studies performed in the US.[69, 75] In the Olson study, cases reported a mean of 3.0 (± 1.8) symptoms while controls had a mean of 0.8 (± 1.3) ($p<0.001$).[69] Ninety-three per cent of cases reported at least one symptom, compared with only 43% of controls. Goff's group showed that the median number of symptoms was 8 for cases versus 2 for controls.[75] Vine *et al.* found that cases were more likely to report having multiple (≥3) symptoms than population-based controls (81% versus 18%, respectively).[7]

Symptom Duration

In a questionnaire study, Goff *et al.* reported median symptom durations between 3-6 months for women with ovarian cancer compared to 11-12 months for clinic controls.[75] However, there were age differences between the two groups, women in the clinic

population included a vast age range (15-90 years) with a median of 45, whereas women with ovarian carcinoma had a median age of 55 (range not reported).

2.4.2.2 Early versus Advanced Disease

A symptom profile specific to early stage disease has obvious appeal, given the significantly more favourable early stage survival rates. Unfortunately, once patients have been divided into early versus late stage, the numbers have often been too small for any meaningful comparisons. Ideally, one would hope to be able to identify symptoms in women with advanced disease, which were also present when the tumour was still localised (i.e. early stage). This would provide evidence that a stage shift was possible via 'early' symptom detection and a 'targeted screening' approach. Although women with early stage disease have symptoms, to date, a clear temporal relationship between symptoms of early and late disease has not been established.

The limited data available indicate that women with early and late stage disease share the same symptom spectrum. There are however, some disparities. Abdominal swelling, [7, 68, 104] fatigue,[7, 68, 69] and gastrointestinal symptoms,[8, 9, 19, 68, 89, 101] are cited more often by women with advanced stage disease. In contrast, urinary symptoms[9, 68] and gynaecological symptoms[9, 12, 68, 101] and an abdominal mass[9, 68] are more frequently associated with early stage disease.

Most studies have reported that women with early disease are more likely to be asymptomatic before diagnosis in comparison to patients with advanced disease.[7, 9, 12, 16, 19, 68, 69, 89, 104] Only one study reported to the contrary.[8] Women with advanced disease also tend to have a higher number of symptoms.[8, 12, 75, 104]

Some groups have attempted to identify symptoms specific to early stage disease by focussing on those that are reported first chronologically.[7, 68] Overall, the first symptoms reported were similar in early and late stage disease, and only non-significant trends have been identified. Vine *et al.*[7] found that pelvic/abdominal discomfort or pain was the most commonly reported first or second symptom for early and advanced stages. Webb and colleagues found that abdominal pain/pressure and abdominal swelling/tightening were most commonly reported as the first symptoms in early (I-II) and late (III-IV) stage disease.[68]

Two questionnaire studies have reported longer symptom duration associated with early stage disease (Table 2-5) however, one had very few early stage cases (n=37 stage I-II),[69] and the other defined early stage narrowly as IA-IIA.[19] Longer symptom duration with early stage disease could be indicative of a less aggressive cancer (as is

the case in breast cancer).[115] In contrast, several other studies found no difference in symptom duration for early versus late stage disease.[8, 12, 68]

In general, there is very little data to suggest that there is a direct relationship between earlier presentation and diagnosis of early stage disease. Several researchers have postulated that differences in tumour biology may explain this phenomenon.[9, 12, 66, 112] Potentially, women detected at advanced stages could have disease that is more aggressive and rapidly metastasizing such that symptoms only appear when disease has already progressed. Conversely, women found with early stage malignancies may have disease that is less aggressive, allowing tumours to grow to a large size while disease is still localised.

2.4.2.3 Benign & Borderline Tumours versus Malignant Tumours

The focus of this project was at the primary care level, and how much earlier the referral to gynaecological-oncology could be made by acting on symptoms. The process of discriminating between benign, borderline and malignant tumours would then be performed at the secondary or tertiary care level. Hence, using symptoms for discriminating between these tumour types was not of interest in this thesis. Furthermore, the existing evidence suggested that it would not be possible (given that each tumour type seems to produce similar symptoms).[6, 7, 9, 13, 67, 68, 75] Therefore, benign tumours were excluded from the case-control study (Part II). Conversely, borderline tumours were included because of their inclusion in the survival and mortality national statistics for ovarian cancer.

Two studies have compared the symptoms associated with benign versus malignant tumours.[9, 75] Both showed that women with benign and malignant masses had comparable symptom profiles. Also, the proportion of women that present with symptoms is comparable for benign and malignant ovarian neoplasms.[9, 75]

Borderline and invasive tumours display no distinctive symptom qualities that would help separate the two groups with certainty, although some minor differences may exist.[6, 7, 13, 67, 68] A Norwegian study reported significantly more abdominal pain, bowel irregularity, fatigue or weight loss and respiratory difficulties with invasive versus borderline disease.[13] Likewise, a study in North America found that pelvic discomfort and bowel irregularity were significantly more common in women with invasive versus borderline tumours.[6] Studies have also shown that women with borderline disease are more likely to be asymptomatic at diagnosis and tend to have longer symptom duration than women with invasive cancer.[6, 13, 67, 68]

2.4.3 Key Methodological Issues

On a research level, examination of prediagnostic symptoms is hampered by the relatively low incidence of disease and the undifferentiated nature of the symptoms. A case-control study design can overcome the problems with investigating rare diseases, but the accurate assessment of highly subjective symptoms requires careful consideration of methodology and study design. The methodological flaws in some studies have limited the interpretation. The main issues will be discussed in this review, including:

- Data Collection Issues - Methods of symptom elicitation (spontaneous reporting versus checklist, lack of validated questionnaires)
- Recall and Recording Errors
- Retrospective versus Prospective Study Design
- Selection Bias
- Semantic Issues
- Poor Definition of Study Populations

Data Collection Issues

Symptom data can be derived either directly (from women) or indirectly (from medical notes); both methods have their inherent drawbacks and each may detect symptoms that differ in quality and threshold. Data collected from women (for research) are considered to be flawed by subjectivity whereas data from medical notes are regarded as objective but limited (for reasons discussed later). In general, a research setting is likely to capture symptoms that patients do not report in clinic. Obtaining retrospective symptom data directly from women has typically involved either the application of various checklists within a questionnaire (most of which are unvalidated), or the use of open-ended questions (spontaneous symptom reporting). These may be in the format of an interview or a self-administered questionnaire. Spontaneous reporting of symptoms is known to elicit lower response rates than specific questioning.[27, 28] Also, the threshold for spontaneous symptom reporting may be altered by variables such as patient personality and skills of the interviewer or physician.[24] In contrast, data that are obtained directly from the subject may over inflate symptom prevalence due to the tendency for subjects to answer positively to symptoms asked about on a checklist.[24] However, this method can also detect symptoms that physicians fail to recognise or record.

Studies with data extracted from medical records have been classified as retrospective however the symptom data collected are really contemporaneous in that they have

been recorded at the time of reporting by the women. For self-administered questionnaires, the women's comprehension of the terms (or jargon) used to describe the symptoms is crucial. Similarly, the words used by women to describe vague symptoms are subject to interpretation and may be paraphrased by the interviewer or GP when recorded. For instance, the phrase "I'm really blown up" could be recorded as abdominal bloating, abdominal swelling or gas. Such scenarios introduce some degree of variation, but this can be reduced by applying strict predetermined symptom definitions. Another important factor is that medical records tend to be diagnosis-driven;[24] often a diagnosis will be recorded rather than symptoms. In case-control studies this is less important, since the effect will be uniform for both groups, however symptom sensitivity (i.e. the full spectrum of symptoms experienced) will be underrepresented.

Symptom reporting is also affected by whether or not women associate their symptoms with cancer.[17, 107] If women perceive their symptoms to be due to other causes, symptoms may go unreported.

A second but different issue is with data that are recorded for non-research purposes (e.g. medical records). Symptom recording in this setting is prone to incompleteness and recording errors.

Recall and Recording Errors

Recall bias is a particular problem in the ovarian cancer population as the diagnosis is traumatic and patient anxiety and mood at the time of recall can affect perception of symptoms.[116] If symptom data are collected after surgery or chemotherapy, one might expect additional recall bias. Furthermore, symptoms may abate or change post-treatment, and chemotherapy could induce new symptoms or worsen pre-existing symptoms. Recruiting women who are about to undergo surgery for pelvic masses of (as yet) unknown aetiology may circumvent relative recall bias (i.e. recall bias between women with malignant versus benign tumours). However, this approach does not address the problem of absolute recall bias (i.e. comparisons with women from the general 'healthy' population), since all women with suspected malignancy will have some degree of anxiety over their pending diagnosis and major surgical procedure. In addition, some women may be warned preoperatively that a cancer diagnosis is almost certain if there are strong clinical indications of malignancy. For these reasons, it is unlikely that study design will completely overcome the issue of recall bias when retrospective data are collected directly from women. A very large prospective study is required to achieve this.

Retrospective data collection is also subject to general inaccuracy from errors in recall. Various factors such as age, gender and severity may contribute to differences in the ability to recall events. However, a more important determining factor in accuracy of recall is time-elapsed since the event. The longer this period is, the greater the memory decay and recall error.[117] The Memorial Sloan-Kettering group asked women about symptoms that were present in the 6-12 months before diagnosis.[69] Although this is a useful interval to examine in terms of possible lead time from prediagnostic symptoms, it seems unlikely that women would be able to accurately recall such a specific time period.

Data extraction from medical notes eliminates recall bias and error but relies on clinicians to record symptoms completely and accurately. Recording error can occur in the form of recording inaccuracies (e.g. mislabelling of symptoms reported) or complete omission of symptoms, particularly for symptoms that are deemed insignificant. Indeed, prediagnostic symptoms are much more common when (retrospectively) self-reported than when ascertained from medical notes.[16, 118] This is exemplified by the meta-analysis results of Bankhead's systematic review.[93] The proportion of patients who were asymptomatic was 23% when data were extracted from medical record studies but only 7% when studies collected data directly from patients. While recall bias almost certainly plays a role in this, it is also likely that not all symptoms mentioned by women are recorded during medical consultation. Equally, women may fail to mention all of their symptoms and focus only on the most bothersome.[94]

A new consideration is the unknown impact of the move from paper to electronic patient records (EPR) on data quality. In general, the quality of disease documentation is probably enhanced but the recording of symptoms may be reduced.[119-122] The use of clinical codes could potentially discourage GPs from documenting all symptoms mentioned since accurate recording requires intimate knowledge of the coding system. There have been reports of large inter-practice variation in coding including multiple ways in which the same clinical concept may be represented.[123] In addition, clinicians who struggle to navigate around computer systems may record less symptom details, particularly if under time constraints in a busy practice. Most EPR systems also allow free text entry for consultation data, and this is the most likely place for symptom data to be recorded. Unfortunately, free text is much more difficult to search and extract data from for research purposes in comparison to codes. In the future, electronic patient records may prove to be of great utility, such as in the rapid identification and flagging of symptom complexes (or clusters) associated with ovarian malignancy, or simply to encourage consideration of ovarian cancer as a differential diagnosis.

Lastly, there may be disparities between symptoms that are extracted from hospital notes versus GP medical records given the difference in focus between primary, secondary and tertiary care. For example, hospital records may only contain symptoms present at admission or those deemed relevant to the disease being investigated. Recording error probably also varies between countries with diversities in medical record keeping practices.[124]

Retrospective versus Prospective Study Design

A prospective study would be ideal, but difficult given the relatively low incidence of ovarian cancer. The key issues pertaining to retrospective versus prospective design in ovarian cancer symptom studies relate to the whether or not women are required to report past events. Recall bias not only affects symptom data that have been reported by women after they have been diagnosed, but also impacts on data once women become aware that they may have a serious morbidity (i.e. before definitive diagnosis). Recall error refers to the inaccuracy that stems from asking for data after events have occurred. Thus, recall error will affect any retrospectively reported symptom data. A survey study carried out by Goff *et al.* purported to be prospective because data were collected data pre-surgery.[75] A truly prospective study of ovarian cancer symptoms would involve recording symptoms as they occur in a cohort of 'healthy' women and following the women for a number of years or until ovarian malignancy is diagnosed.

Selection Bias

There are two main types of bias that affect ovarian cancer studies - survivor bias and self-selection bias. Ovarian cancer has both high morbidity and mortality, hence recruitment of an unbiased patient population to research studies can be challenging. Five-year survival rates for ovarian cancer are poor, hence studies using women who are more than 5 years post-diagnosis can create a survivor bias (over 40% of ovarian cancer patients will die within one year of diagnosis in England[39]).

Self-selection bias can be an issue, especially with studies that recruit via ovarian cancer support groups or their websites.[8, 90] Certain women may be overrepresented in these studies (e.g. proactive women, women who are internet-savvy), while others may choose not to participate. Non-participation may occur to a greater extent amongst women with advanced stage cancers who have a poor prognosis or women who are generally in poor health at the time of approach for recruitment. Potentially, physicians or family members may not even allow researchers to approach women for recruitment. Also, cultural attitudes may have an effect on the decision to participate in research studies.[125]

The ideal control group for a case-control study depends on what the aims of the study are. If a study aims to identify symptom differences that would aid in making a differential diagnosis, then women presenting to primary care with symptoms would comprise a suitable control group. However, a considerable number of symptoms do not prompt clinic visits for various reasons (e.g. self-treatment, non-debilitating, spontaneous resolution, alternative medicine).[24, 126] Therefore, if the desired outcome is to identify symptoms that would help discriminate disease based on those experienced by women in the general population, women who are not consulting need to be included in the control group. Selecting women who are actively seeking healthcare creates a bias towards women who are more likely to be symptomatic and are less likely to have ignored symptoms. However, symptoms that are ignored may be qualitatively different from those that prompt healthcare-seeking behaviour (e.g. mild and transient symptoms are less likely to be reported).

Semantic Issues

Symptoms have been described and defined inconsistently in the ovarian cancer literature. This has limited the comparability of studies and made it difficult to decipher what is truly being measured by the terms used. Several different words and phrases haven been used to confirm the presence of abdominal bloating and abdominal distension.[16, 67, 75] The terms 'abdominal bloating', 'abdominal swelling', 'abdominal distension' and 'increased abdominal size/girth' have been used interchangeably in different studies. For example, Eltabbakh and colleagues have used the terms 'bloatedness' and 'increased abdominal girth' indiscriminately.[67] Similarly, an American study used 'bloating or increased abdominal girth' as a single symptom category in a review of medical records.[16] In reality, symptoms such as abdominal bloating and swelling may not be independent. If the abdomen is distended or swollen, a sensation of bloating is likely to be present in tandem. These are the sorts of issues that a validated questionnaire would help with.

Of equal relevance is how women interpret these terms in symptom checklists or when asked to spontaneously report symptoms. Research performed by Dr Clare Bankhead (Oxford University) as part of a PhD thesis, found that women used symptom terms interchangeably.[94] Specifically, it was noted that women were referring to two distinct events when using the term 'bloating'. One was a persistent distension of the abdomen and the other was a transient distension or a fluctuating sensation of discomfort.[94] Also, use of the word 'symptom' itself may lead to underreporting of symptoms in ovarian cancer studies. This is because many women (and possibly their doctors) do not associate the health changes experienced with ovarian cancer even

after the diagnosis has been made. Constipation is another term which can be confusing. Prevalence of this symptom has been shown to differ when self-reported (yes/no) compared to when specific symptom criteria are gathered (i.e. information on straining, number of defecations per week, incomplete evacuation).[127]

Mislabelling and misinterpretation of symptoms is an important consideration when designing studies with direct questioning of women, as non-medical subjects are unlikely to understand the meaning of anatomical and medical terms. This is a likely source of confusion for women completing questionnaires, and may be a barrier for symptom communication in the clinical setting.

Inconsistent grouping of symptoms is also an issue. For example, a study may report on the presence of 'urinary symptoms' which can include any combination of (unspecified) urinary symptoms,[16, 68, 69, 89] whereas others may report on individual urinary symptoms separately (such as dysuria, frequency, urgency, stress incontinence).[12] Prevalence for individual symptoms is often presented alongside prevalence of grouped symptoms, thus care must be taken when interpreting these values. Also, symptoms have been presented in overly broad categories such as 'constitutional', 'mass effect', 'gastrointestinal', and 'pain'.[9] Consequently, only broad conclusions can be drawn on symptoms. Another issue is the unusual grouping of symptoms, for example dyspnoea has been grouped with back pain.[19]

The issues pertaining to the use of the term 'delays in diagnosis' are also important, but have already been discussed.

Poor Definition of Study Populations

Use of diverse groupings has also made comparisons across studies awkward. For example, some studies have defined 'early stage' disease as stage I-II while others have used stage IA-IB. [10, 19] Furthermore, some studies have included cases in which staging has not been confirmed. Since there is great interest in identifying symptoms associated with early stage disease, complete FIGO staging should be available for all women, unless staging was not possible (e.g. inoperable tumours). It should be noted that staging in women who have received neoadjuvant chemotherapy can be misleading since due to downstaging of metastases after tumour shrinkage. Likewise, despite the heterogeneity associated with prognosis and histology, histological details are often omitted or incomplete in the literature.[17, 19, 69, 75, 90, 110]

Tumour status is another important variable that has been inconsistently reported. Some studies have not provided any details of what tumour types were included. This

is particularly important for tumours that may behave differently from invasive ovarian cancer such as borderline tumours. Borderline tumours have such different prognosis and behaviour from invasive disease that they mandate consideration as a separate entity.[128] Groups that provide national cancer statistics (such as the Office for National Statistics) have acknowledged the inaccuracies in survival data stemming from the inclusion of borderline cancers in the tumour registries.[1] Also, primary peritoneal cancers have been included as ovarian cancer cases in some studies.[16, 94] These malignancies can be indistinguishable at presentation and treatment is also similar, nevertheless primary peritoneal is still regarded as a separate disease and should probably be regarded as such.

Finally, it is crucial to discriminate between consulting and non-consulting women in the control population. Symptoms in the non-consulting population are often managed outside of the formal healthcare system, and only a minority ever reach primary care.[25] If symptom prevalence in the non-consulting population is used, case-control differences will be underestimated and numbers needed to investigate (NNI) will be overestimated.

2.5 Discussion

Table 2-12 Key Ovarian Cancer Symptom Attributes

Most women with ovarian cancer have symptoms before or at diagnosis
Abdominal pain, bloating and abdominal distension are the most common symptoms regardless of data collection method
Symptom profiles of benign/borderline/invasive disease are similar

It has already been established that ovarian cancer elicits symptoms before diagnosis (key attributes of ovarian cancer symptoms are listed in Table 2-12). In this respect, research into ovarian cancer symptoms has turned the first corner. To take this area of research on to the next level, the goal should be to quantify the potential lead time of symptoms. This review has demonstrated that the timing of symptoms prior to ovarian cancer diagnosis requires further clarification.

It is paramount to measure the period between symptom onset and diagnosis, since this interval needs to be sizeable enough to justify any further research into using symptoms as a diagnostic tool for ovarian cancer. Evidence so far has been extremely heterogeneous (see Table 2-3 and Table 2-4). Crucially, interpretation of symptom lead time needs to allow time for both the woman to make the GP appointment (expecting women to present on the same day as symptom onset is unreasonable), and mandatory diagnostic procedures to take place (CA125 testing, TVS, etc.). It is also important to note that delays in diagnosis may be changing due to factors such as

the implementation of directives to reduce cancer waiting times and greater consumer and physician awareness of the symptoms linked to ovarian cancer. Furthermore, one might expect delays to be shorter in countries where patients can make a specialist appointment without GP referral, and GPs have no 'gate-keeping' role. There is a need to establish the current length and sources of delay, and also to evaluate any potential impact on survival. However, it is also important to note that a shorter interval between oncogenesis and diagnosis may not influence survival.

Finally, any future studies should ensure that they address the various methodological issues raised in this review. The terminology used in ovarian cancer symptom studies needs to be clarified. This includes the words and phrases used to define and describe symptoms, and the terms used to define and refer to symptom duration and delays. Inconsistent and unclear definitions have hindered the interpretation of studies and the pooling of data. With regard to delays in diagnosis, it would be sensible to move away from using this pejorative terminology and to establish a more appropriate term. Future studies should ensure greater transparency in the methods used to derive symptom duration and delays. Standardisation of the methodology for quantifying the interval between symptom onset and diagnosis is also overdue. In this thesis, the symptoms checklist in the case-control study (Part II) was specifically designed to try and address some of the issues with terminology. Thus, any potentially ambiguous terms were accompanied by descriptions in parentheses. Additionally, use of the term 'symptom' was avoided in the questionnaire in order to encourage completion and full data capture.

Recently, a validated ovarian cancer symptoms questionnaire was developed in the UK as part of a PhD project. Although previous authors[78] have claimed to have used a validated ovarian cancer symptoms questionnaire, no details of the validation process were provided and the questionnaire does not appear to have been used in any other published studies. The symptoms questionnaire developed in the UK has been fully validated using the EORTC (European Organisation for Research and Treatment of Cancer) guidelines. As such, use of this questionnaire should help resolve some of the issues with poorly defined symptoms.

PART II: A CASE-CONTROL STUDY TO INVESTIGATE SYMPTOMS & EVENTS PRECEDING OVARIAN CANCER DIAGNOSIS

Part II is comprised of a case-control study to examine symptom lead time and events prior to ovarian cancer diagnosis.

3 CHAPTER 3: Background & Methods

3.1 Background

Most ovarian cancer symptoms studies have been small retrospective studies carried out in the United States or Scandinavia. Very few have been performed in a UK population.[11, 97, 99] Given the lack of symptom lead time data and the potential for country-specific differences, it seemed pertinent to undertake a study that would establish the situation in a UK population. The various methodological issues (discussed in the systematic review) further justified a new study. Ascertainment of the potential symptom lead time and refinement of ovarian cancer symptomology were of particular interest. An excellent opportunity arose to collaborate with a large UK-based case-control study (UKOPS) that had already been set-up and was about to start recruitment.

3.1.1 UK Ovarian Cancer Population Study (UKOPS)

UKOPS (UK Ovarian Cancer Population Study) was a multicentre case-control study focused on biological markers of ovarian cancer. Specifically, the study hoped to identify novel biomarkers for early diagnosis, prediction of prognosis and monitoring of recurrence of ovarian cancer, and to identify genes with moderate penetrance for ovarian cancer using Single Nucleotide Polymorphism (SNP). The lead investigators were Dr Usha Menon, Dr Simon Gather, and Professor Ian Jacobs (University College London, Gynaecological Cancer Research Centre). Subjects were recruited from 10 centres across the UK including London, North Wales, Middlesbrough, Southend, Gateshead, Bristol, Portsmouth, Belfast, East Kent and Manchester. The study planned to recruit 2000 cases (1000 newly diagnosed women and 1000 previously diagnosed women), 1500 women with benign/borderline neoplasms and 5000 'apparently healthy' control women. Various sub-studies were nested within UKOPS, each with their individual objectives. The main UKOPS study completed recruitment in April 2009.

UKOPS was launched early on in this PhD and provided an ideal opportunity to collect ovarian cancer symptom data from a cohort of women with theoretically reduced recall bias. In addition, the UKOPS protocol was already finalised and submitted for Ethics approval before collaboration for the current project was agreed, which eliminated the usual study set-up delays. Another convenient factor was that the co-ordinating centre

was in London (UCLH) where all the data could be easily accessed. All of the subjects in the current chapter were collected within the framework of UKOPS.

Recruitment to UKOPS took place in three different settings; (1) women attending with suspected ovarian cancer (potential final diagnoses includes invasive ovarian or fallopian tube cancer, invasive non-ovarian cancer, borderline ovarian cancer, benign neoplasms, normal ovaries) (2) women undergoing treatment or follow-up of ovarian cancer; and (3) unselected women from the UK Collaborative Trial of Ovarian Cancer Screening (UKCTOCS) population-based trial of ovarian cancer screening (see 3.1.2).

At baseline, all study participants were required to give blood (for SNP and biomarker analysis) and to complete a questionnaire about putative risk factors for ovarian cancer. In addition, women undergoing surgery (i.e. potential cases) were also asked to donate tissue.

Cases in the UKOPS study were women aged 25 and above with histological confirmation of primary invasive or borderline ovarian or fallopian tube cancer (see below for full eligibility criteria). The Gynaecological Cancer Research Centre (GCRC) is also the lead centre for UKCTOCS, which is a population-based ovarian cancer screening study (see section 3.1.2 for details). UKOPS control subjects were postmenopausal women participating in the multimodal arm of the UKCTOCS study. They were approached for UKOPS study recruitment at their annual UKCTOCS screening visit. The UKOPS eligibility criteria below pertain to the main study. For the symptoms substudy, only women aged 45 and above who met case criteria (1) and (2) were included as cases. Similarly, only women who met control criterion (2) were included as controls. However, women who met control criterion (1) and were later confirmed as borderline cancers were included as cases in the symptoms substudy (if aged ≥45). More specific details are listed under 3.3.1 Symptom Substudy Population.

UKOPS Inclusion Criteria

Cases:

(1) Women with an adnexal mass suspicious of ovarian cancer who were about to undergo surgery. (Once final histology was confirmed, women with borderline or benign tumours were reclassified in line with the endpoints of each UKOPS sub-study. Women with normal findings were excluded).

(2) Women with a probable diagnosis of primary invasive ovarian cancer who were not undergoing surgery (i.e. women with unresectable tumours or who were too unwell for surgery) or who were scheduled for neoadjuvant chemotherapy.

(3) Women with a confirmed diagnosis of primary invasive or borderline ovarian/fallopian tube cancer (i.e. previously diagnosed women).

Controls:

(1) Women with a possible benign or borderline adnexal mass who were about to undergo surgery. (Women with normal findings were excluded).

(2) Apparently healthy women recruited from the multimodal arm of UKCTOCS when they attended for annual screening. (More detailed eligibility criteria are below).

UKOPS Exclusion Criteria

(1) Currently active non-ovarian malignancy

(2) Age below 25 years

(3) Unable to give informed consent

3.1.2 UK Collaborative Trial of Ovarian Cancer Screening (UKCTOCS)

UKCTOCS is a population-based study designed to assess the impact of screening on ovarian cancer mortality in the UK. It is the largest randomised controlled study in ovarian cancer screening to date, with over 200,000 women recruited from 12 regional centres (London, Bristol, North Wales, Portsmouth, Belfast, Middlesbrough, Manchester, Liverpool, Nottingham, Cardiff, Derby, Gateshead). Notably, some of the UKCTOCS and UKOPS centres overlap (i.e. London, Bristol, North Wales, Portsmouth, Belfast, Middlesbrough, Manchester, Gateshead). A total of 1,243,282 women were sent invitations of which 202,638 were randomised.

Participants were randomised to one of three arms:

- Multimodal Group (50,640) – annual screening with serum CA125 as the primary test and repeat CA125 and transvaginal ultrasound as the secondary test.
- Ultrasound Group (50,639) – annual screening with transvaginal ultrasound as the primary test and repeat detailed ultrasound in 6-8 weeks as the secondary test
- Control Group (101,359) – no screening

Women in the screening arms are to be screened until 31st December 2011, and the primary endpoint is ovarian cancer mortality on 31st December 2014. The UKCTOCS recruitment phase was completed in September 2005. Participation was by invitation only. Letters of invitation were sent to women aged 50-74 randomly selected from the age/sex registers of the health authorities geographically-related to the 13 study centre catchment areas. Women who responded could participate if they met the following eligibility criteria.

UKCTOCS Inclusion criteria

- Age 50-74 years inclusive
- Postmenopausal:
    - >12 months amenorrhoea following a natural menopause or hysterectomy or;
    - >12 months of hormone replacement therapy commenced for menopausal symptoms

UKCTOCS Exclusion criteria

- Bilateral oophorectomy
- Currently active non-ovarian malignancy (excluding skin cancer)
- Women who have had an ovarian malignancy in the past
- Women at high-risk of ovarian cancer due to a familial predisposition
- Women participating in other ovarian cancer screening trials

3.2 Aims & Objectives (Symptom Substudy)

3.2.1 Study Aim

To investigate the potential for diagnosis of ovarian cancer to be expedited via prompt symptom recognition.

3.2.2 Primary Objectives

- To identify the type, duration, severity and frequency of symptoms present at different time periods prior to diagnosis of ovarian cancer (cases) or study enrolment (controls).
- To calculate the proportion of cases that would be identified using symptoms at different time points before diagnosis.
- To identify the frequency of GP visits in the 2 year period prior to diagnosis of ovarian cancer (cases) or recruitment (controls).

3.2.3 Secondary Objectives

- To assess the time interval between symptom onset and first visit to a GP.
- To assess the time interval between first GP visit for symptoms related to ovarian cancer to referral, and from referral to diagnosis (cases only).
- To relate symptom history to the type and stage of cancer.
- To distinguish between new and chronic symptoms reported.
- To explore the potential for developing a crude symptoms index.

3.3 Methods

3.3.1 Symptom Substudy Population

3.3.1.1 Cases

Symptom Substudy Inclusion Criteria

- Histological or cytological confirmation of primary ovarian (ICD-10 C56) or fallopian tube (ICD-10 C57.0) cancer
- Age 45 and above
- Newly diagnosed (within 3 months of consent to main UKOPS study)
- Telephone interview – performed a maximum of 3 months after diagnosis

All histological subtypes were included but preference was given to women with invasive epithelial ovarian cancer. This is because these women contribute heavily to the low survival rates, and therefore were of greatest interest. Newly diagnosed women did not necessarily have definitive diagnosis at the time of study entry. However, cases who were recruited into UKOPS before definitive diagnosis were only included in the present study once histology was confirmed. Previously diagnosed cases were excluded to minimise recall bias and recall error. Inclusion was initially limited to women who had confirmed diagnosis no more than one month before recruitment. Unfortunately, case recruitment was slower than anticipated, so data were also collected on women who were consented within 3 months of diagnosis.

The age restriction was chosen to comply with the overall project aim of investigating the potential for 'targeted' ovarian cancer screening, which is unlikely to be offered to premenopausal women. The rationale for this is that premenopausal women tend to have tumours with a favourable prognosis, and more than 85% of women with ovarian cancer are aged 45 or older.[32] Thus, the positive predictive value of any symptoms index would be substantially lower in a premenopausal group. At first, cases aged over 74 were also excluded since screening is not routinely offered to this age group due to

the limited potential for gain in life years. However, part way through the study it was decided that inclusion of women over 74 would be worthwhile to increase study numbers since recruitment was poor.

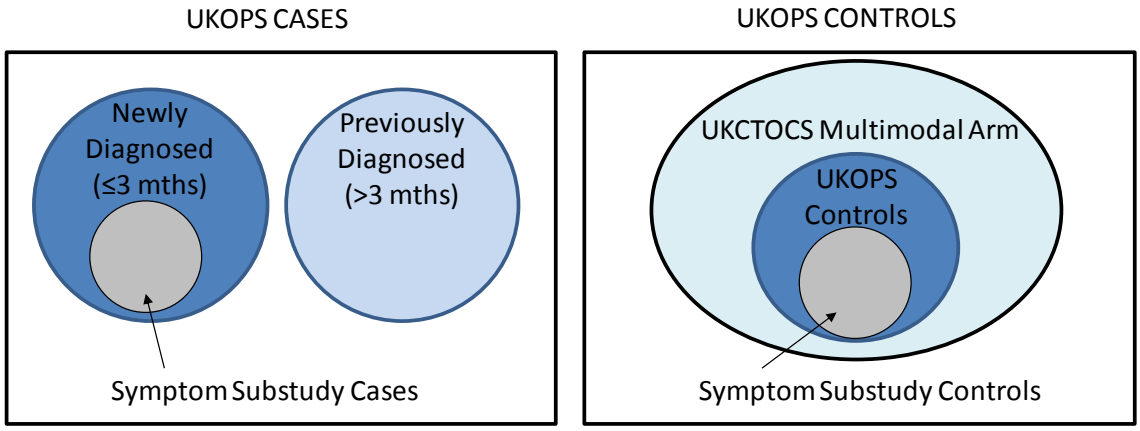
3.3.1.2 Controls

Symptom Substudy Inclusion Criteria

- Postmenopausal women from the multimodal arm of UKCTOCS who agreed to participate in UKOPS main study (approached at their annual screening visit)
- Aged 50-74 (at UKCTOCS entry)
- Telephone interview – performed a maximum of 3 months after consent to main UKOPS study.

Since there were significantly more controls than cases recruited to UKOPS, only a subset of UKOPS control women were included in the current study (see Figure 3-1). At study outset, control women were selected sporadically according to no specific criteria (n=56). After June 2007, controls were randomly selected using a frequency matching approach to balance for year of birth and agreement to telephone interview. The majority of controls were collected in this randomised manner (212/268).

Figure 3-1 Study Schematic



3.3.2 Study Design

This was a multicentre, retrospective case-control study in the UK. Recruitment took place at the following ten study centres: London, Bristol, North Wales, Portsmouth, Belfast, Middlesbrough, Gateshead, Manchester, Southend and East Kent. Cases were collected from all ten centres, and controls were collected from all but East Kent and Southend.

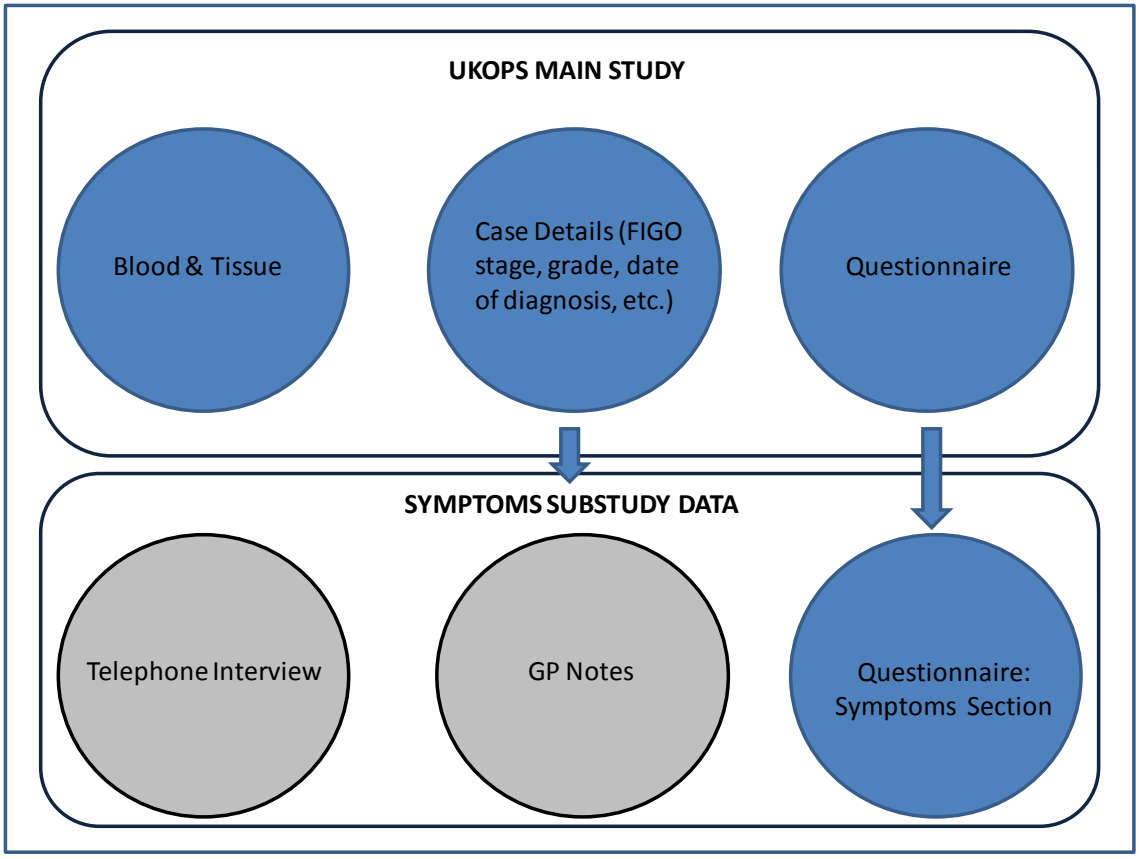
Symptom data were obtained both directly (questionnaire and telephone interview) and indirectly (GP notes) from the women. In order to gain the best possible symptom data, three different collection methods were used (in different proportions):

- Paper questionnaire
- Telephone interview
- GP notes

All women recruited to the main UKOPS study completed a symptoms questionnaire (see Appendix I) and gave permission for the research team to access their medical notes on the main study consent form. The telephone interview was optional and consent for this was indicated on a tick box on the symptoms questionnaire. For questionnaire and interview, symptoms of any duration were included as long as they were present (anytime) during the 12 months prior to diagnosis or consent. For GP notes, any symptoms that were recorded during the 24 months prior to diagnosis or consent were included.

Recruitment of women and questionnaire administration were carried out by the designated research nurses for UKOPS. The data extraction and analysis for all 3 data sources were performed by the author of this thesis (Anita Lim), as were the telephone interviews and the GP note collection and review (see Figure 3-2). Case details were taken from the main UKOPS study (i.e. FIGO staging, tumour grade etc.).

Figure 3-2 Schematic Diagram Showing Data Collected for Symptoms Substudy & Data Taken from the Main UKOPS Study



Developing and Piloting of the Questionnaire

The UKOPS group had already designed and submitted a questionnaire for ethical approval before the addition of this thesis project was agreed. It contained a symptom section which was based largely on a questionnaire used in one of the Goff studies.[75] However, it did not ask for sufficient data to meet our study objectives. Furthermore, it was positioned at the end of a 20 page document which was intended for women to complete unassisted. Hence, a second version (amendment 1) was quickly developed in collaboration with core members of the UKOPS study team, to be implemented before recruitment started. Version 2 was the first questionnaire used in the symptoms substudy. Several factors were taken into account in this second version. Firstly, based on research performed by Dr Clare Bankhead (Oxford University) as part of a PhD thesis, it was clear that women used particular symptom terms interchangeably (for example 'abdominal bloating' and 'abdominal swelling').[94] Thus, any potentially ambiguous terms were accompanied by descriptions in parentheses. Secondly, nausea and vomiting were combined as they were presumed to be highly correlated.

The third version (amendment 3) was produced in June 2006 after there had been time to review the data quality of returned questionnaires which was relatively poor. The evolution between version 2 and 3 of the questionnaire was based on informal piloting of several different versions and brief consultation with a psychologist who was experienced with questionnaire development (Professor Jane Wardle). Initially, five non-medical female members of staff in the department completed the draft questionnaires and their feedback was incorporated. Next, the amended drafts were piloted in face-to-face interviews with five cases and a 'watch and listen' technique was applied. This technique involved watching subjects complete the questionnaire and noting any issues with completion that were observed or vocalised. A final version was produced based on the collective feedback. In addition, the title was changed to 'Health Changes' since there was debate over the ambiguity of the term 'physical changes' which was used in version 2.

These procedures were not ideal in terms of questionnaire development and validation, but timelines and limited resources did not permit formal validation. The desired goal of the piloting was to improve data quality. In brief, frequency was changed to an ordinal categorical scale, the wording of the questions was revised, and an extra page was added for additional information. Also, a sentence was added to specify that severity and frequency should be rated for the symptom at the time it first started since these are likely to change over time. In particular, one would expect worsening severity and frequency with disease progression.

3.3.3 Questionnaire

Baseline questionnaires were completed at the time of recruitment. For controls, recruitment was during attendance for annual UKCTOCS screening. For cases, recruitment was during pre-admission visits or on hospital wards before surgery or chemotherapy (i.e. pre-treatment).

There were two different versions of the questionnaire used in the symptom substudy (version 2 and 3, see Table 3-1 and Appendix I). Both versions asked women to indicate if they had experienced any symptoms from a checklist in the last 12 months.

Version 2 contained a checklist of 17 symptoms associated with ovarian cancer (as identified by previous studies) with an additional box for 'other changes'. Women were asked to provide the following information:

- Severity on an ordinal categorical scale (mild, moderate, severe)
- Frequency in number of days per month

- Duration in months indicated as a range of calendar months (how long the symptom was present prior to diagnosis/questionnaire completion)
- If the symptom was ongoing (tick box)
- If the symptom was reported to their GP and if so, when?

Version 3 consisted of a checklist of 14 symptoms with a box for ‘other symptoms’, with a second page for additional information. Subjects were asked to provide details of the following:

- Start date (month and year)
- Severity (at time symptom started) on an ordinal categorical scale (mild, moderate, severe)
- Frequency (at time symptom started) in number of days per month on an ordinal categorical scale (1-4, 5-15, 16-31)
- Duration in months (how long the symptom was present prior to diagnosis/questionnaire completion)
- If the symptom was ongoing (Y/N)
- If the symptom was reported to their GP and if so, the approximate date of consultation (month/year).

Table 3-1 Comparison of Questionnaire Checklist Symptoms for Version 2 & 3

Version 2	Version 3
Pelvic/Abdominal Pain	Pelvic/Abdominal Pain or Discomfort
Back Pain	Back Pain
Indigestion	Indigestion
-	Loss of Appetite
Nausea or Vomiting	Nausea or Vomiting
Weight Loss (unplanned) or appearance of weight loss	Weight Loss (unplanned) or appearance of weight loss
Abdominal Swelling (actual increase in size or girth)	Increase in Abdominal Size
Abdominal Bloating (feels full and tight)	Abdomen Feels Bloated
Able to Feel a Lump in Abdomen	Able to Feel a Lump in Abdomen
Urinating more often than usual or feeling an urgent need to go	Urinating more often than usual or feeling an urgent need to go
Constipation	Constipation
Diarrhoea	Diarrhoea
Fatigue	Fatigue
Irregular Periods	Irregular Vaginal Bleeding*
Bleeding after menopause	-
Bleeding with Intercourse	-
Pain during Intercourse	-
Leg swelling	-

*Due to limited space on the questionnaire irregular vaginal bleeding included other vaginal bleeding irregularities (e.g. post-coital bleeding, postmenopausal bleeding etc.).
Note: See Appendix I for copies of the questionnaires.

The questionnaire was designed to be self-administered however, early feedback indicated that support was needed. Hence, research nurses were instructed to explain how to complete the questionnaire and to be available to answer questions. If the nurse could not be present (e.g. due to time constraints); subjects completed the questionnaires unassisted. In order to encourage consistency, a list of symptom definitions (see Appendix II) was given to the nurses for reference. The list also included specific definitions for each severity rating (see Table 3-2). These definitions were only intended to be used if women were unclear about how to assign a severity rating.

Table 3-2 Severity Definitions

Mild	Symptom easily tolerable, did not interfere with usual activities
Moderate	Symptom hard to tolerate, interfered somewhat with usual activities
Severe	Symptom was so intolerable that usual activities could not be performed

3.3.4 Telephone Interview

Participation in the telephone interviews was optional and only those who gave specific consent were contacted. All interviews were performed by the same interviewer (Anita Lim). They were conducted unblinded since it was highly probable that cases would discuss events surrounding their cancer diagnosis. Also, cases were expected to be psychologically and physically frail (or possibly even deceased) by the time contact was attempted for interview, and the interviewer needed to be sensitive to this. Interviews for cases took place preferably before chemotherapy began and within 3 months of surgery or positive biopsy. No limits were placed on the number of contact attempts for cases if there was no answer or the women were not available for interview (provided the women were happy to be re-contacted at a later point).

As mentioned previously, agreement to telephone interview was used as one of the stratifiers for random control selection. Only controls who were recently recruited (within 3 months of consent) were contacted. A record of women who consented to telephone interview but were not contactable was maintained. A maximum of 3 contact attempts were made for each woman after which they were replaced with the next control on a ranked random number list. All controls were interviewed within 3 months of consent. If a subject was not contactable within the 3 month period no further attempt to interview was made on the grounds that recall would be less reliable.

Each telephone call began in the same manner based on a predetermined script, to ensure consideration for the women's state of well-being at the time of the call and consistency (see Appendix III and IV). Women were given the opportunity to withdraw their consent for interview or reschedule the interview for another time. All interviews were performed using a standardised questionnaire (see Appendix V). Any verbatim comments of interest were also recorded. These were used to help inform the interview answers but were not formally analysed. The interview consisted of the same symptom list as the paper questionnaire with some additional questions. One deviation from the questionnaire was that 'loss of appetite' was combined with 'feeling full quickly' (i.e. 'loss of appetite or feeling full quickly'). This symptom combination had appeared on one of the draft questionnaires but was later edited to 'loss of appetite'. Unfortunately, this was not updated on the telephone survey, and when the discrepancy was noticed, many interviews had already been performed, hence the

combination was kept to maintain consistency. As with the questionnaire, symptoms of any duration were included as long as they were present (anytime) during the 12 months prior to diagnosis or consent. Cases were also asked about their diagnostic route and were reminded to exclude any symptoms that started after diagnosis. This was particularly emphasized to cases who had already started chemotherapy at the time of interview. Again, the severity definitions listed in Table 3-2 were only referred to if women were confused about the categories.

Since the interview was intended to help gain the best possible symptom data, they were originally conducted with a copy of the subject's paper questionnaire present so that any ambiguous symptoms on paper could be clarified or probed. The interviewer did not inform the subjects of any inconsistencies between the two questionnaire answers. Referral to the subject's questionnaire was discontinued early on in the study since there seemed to be little additional information gained by probing based on questionnaire answers and the questionnaire was not always received at the coordinating centre in time.

As with the questionnaire, aspects of the interview were amended in the early phases of the study. The first few interviews included questions about medication or treatments used to treat symptoms however, on further consideration it was decided that these data would not contribute greatly to the study. Interviews conducted between September 2006 and July 2007 (cases n=72, controls n=59) included four additional symptoms based on Dr Clare Bankhead's thesis findings.[94] These were shoulder pain, dizziness, change in taste and night sweats. Case-control differences were not detected on an anecdotal level for any of the additional symptoms. The decision to stop asking about these symptoms was taken following discussions with the clinical investigators.

After several interviews, it became apparent that some symptoms were associated with activity (e.g. back pain and gardening) or were seasonal (e.g. leg swelling in hot weather). Also, some women were reluctant to report symptoms for which a cause was already attributed or that were mild and transient. Also, if a symptom occurred for a few days but only 2-3 times a year, this information would not be captured by the frequency question making it difficult to complete data for this parameter. In an attempt to deal with these issues, a note was made if the symptom was:

- Mild and transient
- Attributed to an illness or condition
- Activity-dependent

- Intermittent, seasonal or episodic

Subjects were also asked if the symptom worsened, improved or stayed the same over time.

In order to aid recall accuracy, the use of an event history calendar (EHC) was originally planned for telephone interviews. This was to consist of the subject's birthday, significant life events (e.g. anniversary, deaths, retirement, holidays, family celebrations) and major public holidays. These events were to act as cues to help the subjects report more accurately on the timing of their symptoms by taking advantage of the ways in which autobiographical memories are stored and structured.[129] It was to take place before the formal interview took place. This event calendar was to cover the 12 months leading up to the month of diagnosis and events were to act as timing cues. As there was uncertainty with regard to the benefits and feasibility of using such a tool in this setting, the EHC was treated as a pilot and several mock interviews with volunteers (not in the study) were carried out. These 'dry-runs' indicated that interview time was greatly prolonged by the EHC (to approximately 1 hour) and given the poor health status of the patient population it was deemed that the potential improvement in data quality at the cost of patient well-being was not justified. Thus, wherever possible anchor points (e.g. public and personal holidays, seasons, birthdays, Christmas etc) were used in place of the EHC.

3.3.5 Face-to-Face Interview

Five newly diagnosed cases were interviewed face-to-face (as inpatients) to allow informal piloting during symptom questionnaire development. Due to logistical and monetary issues, these were restricted to patients who were treated at UCLH. These interviews were analysed together with telephone interview data and are not presented separately due to small numbers.

3.3.6 GP Notes

GP notes were requested for all subjects in the symptoms substudy for the two year period prior to date of diagnosis or consent for cases and controls, respectively. Despite this request being made for all subjects, expectations were that records for only a proportion of women would be received. The information requested included medical history, consultation notes, test results and clinic letters. For GP practices with electronic and paper records, copies of both were requested. Subjects for whom only clinic letters were received were excluded from the GP note analysis. Data were extracted from the notes by the same researcher, and again, this was conducted

unblinded for reasons of practicality. Clinic letters included incoming and outgoing letters, and were only databased if any symptoms possibly related to ovarian cancer were recorded (see below).

Data extracted for all subjects included:

- Dates of any consultations in the two years prior to 'cut-off date' (cases) or consent (controls). See definition of 'cut-off' date below
- Any signs or symptoms (coded and free text)
- Symptom onset dates. If this was missing, the visit date was used instead
- If a blood test was ordered at a visit
- Whether any per rectal (PR), abdominal or pelvic vaginal examinations were performed at the visit
- Dates of diagnosis for irritable bowel syndrome (IBS) and diverticular conditions since they have been associated with misdiagnosis of ovarian cancer in the literature.[8] Diverticular conditions included diverticulosis, diverticulitis and diverticular disease. Diverticulosis was included even though it is often asymptomatic since all diverticular conditions appeared to be recorded interchangeably in the GP notes.
- If clinic letters mentioned any symptoms possibly related to ovarian cancer, the clinic date and type of symptom(s) were recorded. If a symptom onset or duration was described, these were also recorded.

For cases, the following additional data were extracted whenever possible:

- Dates of CA125 tests and ultrasound scans (pelvic, transvaginal or abdominal)
- Pre-treatment CA125 levels (result closest to treatment date if more than one)
- First referral made by GP (date and type)
- A 'cut-off date' for when ovarian aetiology was first suspected. 'Cut-off date' was defined as the date when suspicion of ovarian aetiology was first documented. For example if any of the following text was recorded; '?ovarian' '?ovarian cyst' '?pelvic mass' '?ovary'.
- Route of diagnosis
- Symptom details from multidisciplinary team (MDT) letters (if available). These were often helpful in deciphering route of diagnosis.

Consultations included face-to-face visits (surgery, home, out-of-hours) and telephone calls. Accident and emergency (A&E) visits (recorded in GP notes) were included even though they are not GP visits, since symptoms reported at these were of interest. Visits for blood drawing only, or telephone calls to report test results only, were

excluded. Each consultation was classified as either 'problem' or 'routine'. In brief, 'problems' were consultations for which the main reason was a new or undiagnosed problem. 'Routine' consultations included nurse visits or regularly scheduled GP visits such as monitoring of hypertension, diabetes or asthma, smears and vaccinations. 'Problem' visit coding overrode 'routine' if a problem was raised or discussed at a routine visit. Full details of the classification system are in Appendix VI.

The current project aims to bring forward diagnosis by expediting the date of appropriate referral (i.e. gynaecological-oncology). Therefore, frequency of visits and symptoms that developed after ovarian cancer was already suspected were of little interest. Hence, case GP visits were only databased up until a 'cut-off date' (as defined above).

Only symptoms considered to be possibly related or attributable to ovarian cancer were databased in detail (see 3.3.10 and Appendix VI). Initially, the list of possibly related symptoms was kept as broad as possible. Emphasis was placed upon abdominal, gastrointestinal, urinary, pelvic and gynaecological symptoms. If a diagnosis was recorded without specific symptom(s) (e.g. 'cystitis'), the diagnosis was databased. After an exhaustive list was created it was discussed with the clinical investigators to produce a final list of relevant symptoms (see Appendix VI). The same symptom classification was applied to all 3 data sources.

Data on clinic letters, blood tests, examinations by the GP and dates of CA125 and ultrasound were not analysed in the current thesis. However, these are planned to be included in a future more in-depth analysis.

3.3.7 Ethical Approval

Ethical approval for the study was granted from the joint UCL/UCLH (University College London/University College Hospital London) Ethics Committee. Subjects in the symptoms substudy were recruited between February 2006 and April 2008.

3.3.8 Sample Size Considerations

Formal power and sample size calculations were not performed since the study did not primarily aim to identify differences between two groups (case-control). Rather, the primary aim was to determine the proportion of women with ovarian cancer with 'early' symptoms. Thus, it was appropriate to discuss sample size in terms of sensitivity for the proportion of symptomatic cases that could be detected *x* months (e.g. 3, 6, 9) in advance, and the width of the associated 95% confidence intervals (CI).

The main study (UKOPS) initially set out to obtain 1000 newly diagnosed cases. Based on this estimate, 600 telephone interviews were planned in 500 cases and 100 controls. However, recruitment was much slower than expected and targets were adjusted accordingly to collect questionnaires from 250 newly diagnosed cases and 250 controls and perform telephone interviews in 150 cases and 150 controls. GP notes were requested for all women but were only received for a proportion of women (see 3.3.6).

For a sensitivity between 30% and 70% with the original sample size of 600 cases, the sensitivity could be estimated with a 95% confidence interval of ±4% (absolute sensitivity, see Table 3-3). With the actual sample size of 254 cases, the 95% confidence interval changed to ±6%. Hence the reduced sample size changed the expected width of the 95% confidence interval for the proportion of cases with the symptom from 0.07 to 0.12.

Table 3-3 95% Confidence Intervals (CI) Associated with 30% & 70% Sensitivity for Symptoms

30% Sensitivity	**95%CI**	**Approx. width of 95%CI**
Original Sample Size n=600	26.4% to 33.8%	±4%
Actual Sample Size n=250*	24.4% to 36.1%	±6%
70% Sensitivity		
Original Sample Size n=600	66.2% to 73.6%	±4%
Actual Sample Size n=250*	63.9% to 75.6%	±6%

*Rounded number (actual number of cases is 254)

Finally, given that the symptom data were collected from 3 different sources, the overall symptom data quality was expected to be much higher than that produced by previous studies.

3.3.9 Data Cleaning

Many issues arose with regard to data quality and cleaning. Some questionnaires had missing or nonsensical data (see Appendix VIII), some women gave vague interview answers (e.g. answers of 'not really' for symptom presence) and GP notes contained free text that required interpretation (e.g. symptom recorded as 'blown up').

Symptom classification was also more problematic than anticipated and decisions had to be made for symptoms that could reasonably belong under more than one category. For example, epigastric pain could be coded as indigestion or abdominal pain. As described in 3.3.6, all symptoms were discussed with the clinical investigators to set up a systematic coding frame and symptoms were grouped accordingly.

In order to address these issues, a set of databasing rules were drawn up (see Appendix VI). The following is a summary of how the major issues were dealt with:

- Missing or partial symptom variables (start dates, duration, end dates) were derived using existing data whenever possible (see Appendix VII)
- Nonsensical data were corrected when it could be clearly justified (e.g. if a questionnaire was completed in September 2006 but had a symptom recorded as starting in June 2007, with a recorded duration of 3 months; the onset date would be amended to June 2006). Otherwise, these were changed to missing.
- If the 'Y/N' answer was not completed on the questionnaire next to a symptom, the symptom was assumed not to be present (i.e. 'No') unless other data were completed for that symptom (e.g. severity, frequency) in which case the symptom was assumed to be present (i.e. 'Yes').
- On the questionnaire, ongoing status was assumed to be 'yes' if ongoing was missing
- Symptoms were classified and grouped to reduce the list of over 300 symptoms documented (see Appendix VI).
- Symptoms within each classification heading were assumed to be part of the same complex. E.g. if a woman reported epigastric pain and belching that occurred separately, both would be considered to be indigestion.
- If more than one symptom was recoded to a single symptom resulting in multiple symptom parameters (e.g. loose stools and frequent bowel motions were both coded as diarrhoea), the following rules were applied; use the
    - Earliest symptom onset date
    - Latest symptom end date
    - Maximum duration
    - Maximum severity
    - Maximum frequency
    - Earliest date of 'Yes' to GP visit
    - Answers of 'Yes' to ongoing overrides 'No' to ongoing
- Symptoms that started >2 years before diagnosis or recruitment were treated as a separate group (symptoms that started before this were considered to be unlikely to be related to ovarian cancer). Exact onset dates and duration for these symptoms were not derived if missing.
- Symptoms from all 3 sources were coded as 'other unrelated' without specific details (onset date, severity etc.) if they were not on the list of 'possibly related' symptoms
- Symptoms with a reported onset date after diagnosis or consent were dropped.

- If >1 unrelated symptom was reported, only one was databased as 'other unrelated'.
- Blank questionnaires were treated as missing data (i.e. no questionnaire).
- All symptoms on the list of 'possibly related' symptoms were databased regardless of medical history (i.e. no assumptions were made about symptom aetiology)

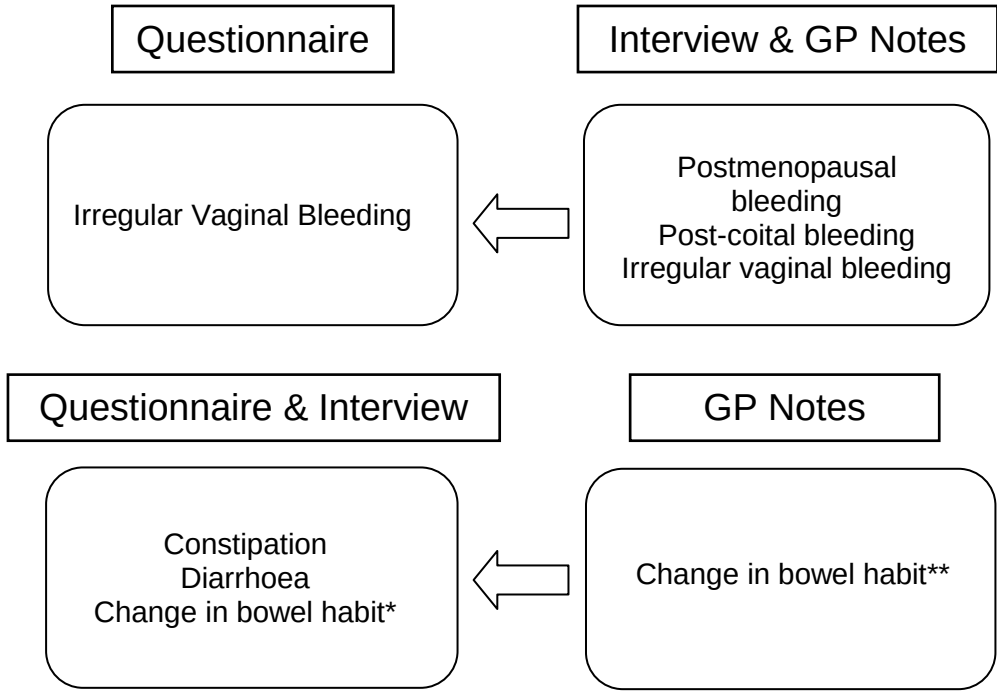
The list of 'possibly related' symptoms excludes some of the more unusual presentations of ovarian cancer (such as hirsutism or paraesthesia) as these symptoms were considered too rare to be included for the current study purposes.

Due to limited space on the questionnaire, irregular vaginal bleeding (IVB) was broadly defined to include all vaginal bleeding irregularities. However, on interview and GP notes postmenopausal bleeding (PMB) and post-coital bleeding were separated out from IVB since they were of special interest. A change in bowel habit is much more likely to be salient than any pre-existing tendency to constipation or diarrhoea that is 'normal' for a woman. Therefore, in GP note data, a change in bowel habit was recorded if it was clear that bowel symptoms (including constipation and diarrhoea) were unusual for that woman or represented a change from the norm. By contrast, change in bowel habit was only databased from the questionnaire and interview if women reported mucous in stool or change in stool consistency, shape or size. As such, some of the change in bowel habit data in GP notes will be represented by constipation or diarrhoea in the self-reported data (i.e. questionnaire and interview) and vice-versa. This issue was resolved by limiting the main results to symptoms that were new in the last two years before reference date. These differences in symptom definitions for the various sources are summarised in Figure 3-3.

Clinic letters were databased, but not formally analysed since their original purpose was to supplement the GP note data, especially for women in whom GP note data quality was poor. Some GP notes contained no consultation details over the period although it was clear a GP referral had been made before diagnosis.

Three women had already completed upfront chemotherapy at study entry. This was discovered after their data were already collected for the current study hence, they were included in the main analysis. See the Data Analysis section for details of an analysis which excluded these women (and other subjects).

Figure 3-3 Interchangeable Symptom Definitions in Questionnaire versus Other Data Sources



*Includes mucous in stool, changes in consistency, shape or size of stool
**Preferentially includes diarrhoea and constipation if these were indicated to be a change from normal bowel habit. Also includes mucous in stool, changes in frequency, consistency, shape or size of stool

3.3.10 Definition of Study Variables

Case Details

Case details such as date of diagnosis, FIGO staging, tumour histology and grade were sourced from the main UKOPS study. Date of diagnosis (DOD) was taken as date of first surgery or positive biopsy. Final staging and histology were confirmed by an independent pathologist. Histology was classified according to the International Classification of Diseases for Oncology, morphology for neoplasms, 3rd Edition (ICD-O-3).[130] 'Cut-off' date was the date when suspicion of ovarian aetiology was first documented. For example, if any of the following text was recorded; '?ovarian' '?ovarian cyst' '?pelvic mass' '?ovary'.

Early and advanced stage disease comprised FIGO stage I-II and III-IV tumours, respectively. This included borderline tumours and excluded unstaged tumours.

The route of diagnosis for cases was investigated by extracting data on first referral made by the GP. First referral was defined as the first referral made by the GP for any symptom(s) possibly related to ovarian cancer. For women who presented as

emergencies before a GP referral was made or referral appointment was undertaken, accident and emergency (A&E) was considered to be the first referral (patient- or GP-initiated). Women in whom disease was identified via screening studies or incidental findings were excluded from this classification system. Incidental diagnosis was defined as diagnosis solely by chance investigations (opportunistic diagnosis). This excluded women who presented with symptoms possibly related to ovarian cancer that were investigated for other diseases. Date of first referral was taken from the GP medical records. For women in whom no notes were received, date of first referral was taken from MDT letters, or interview and questionnaire whenever possible. Default dates of the 15th of the month were used if only month and year of referral were known.

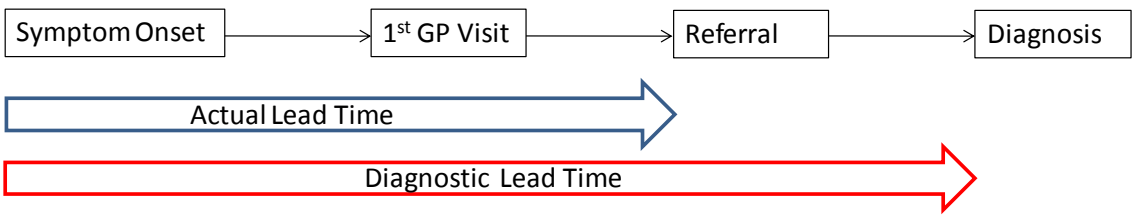
Timing Variables

Timing of prediagnostic symptoms was calculated relative to a reference date. For cases this was the date of diagnosis, and for controls this was the date of consent. The two years before reference date was divided into prediagnostic periods comprising 0-2, 3-5, 6-8 etc. Each period was equivalent to one quarter of a year. These were calculated such that '0-2' included symptoms that occurred anytime between 0 months up to, but excluding 3 months (i.e. 2.99 months) before reference date.

Timing of symptoms in the GP notes was problematic since there was no information on symptom duration. For this study, it was assumed that symptoms started from the earliest time recorded and continued up until diagnosis or consent.

In general, symptom lead time is considered to be the time between symptom onset and diagnosis. Symptoms at diagnosis were of little interest since they provide no information on the proportion of women who could benefit from any intervention based on symptoms. The greatest opportunity for intervention is the interval between symptom onset and referral. For clarity, two definitions of symptom lead time will be used in this thesis; actual lead time (time from symptom onset to referral) and diagnostic lead time (time from symptom onset to diagnosis) (see Figure 3-4).

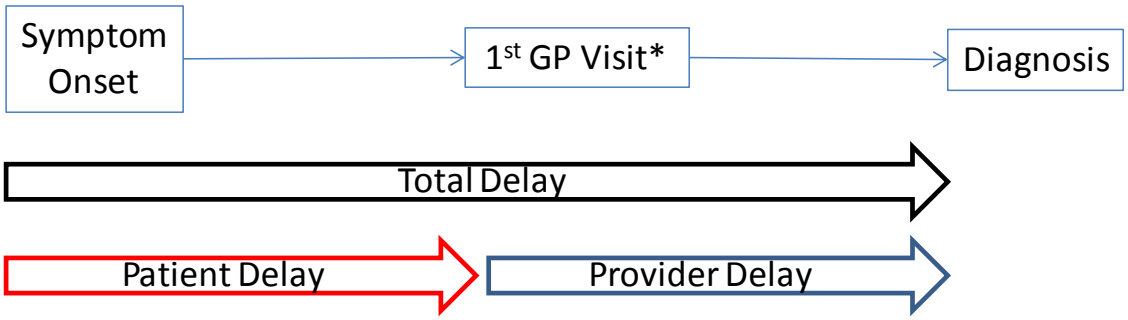
Figure 3-4 Symptom Lead Time Definitions



Delays in diagnosis

As discussed in the systematic review, the term 'delay' is not an appropriate label for the interval between symptom onset to diagnosis, however, it was used in this thesis to maintain consistency with the literature. Delays in diagnosis were evaluated in terms of total delay, patient delay and provider delay (see Figure 3-5). Total delay was defined as time from first symptom onset to date of diagnosis. Patient delay was defined as time from first symptom to first GP visit. Provider delay was defined as time from first GP visit to date of diagnosis.

Figure 3-5 Schematic Diagram showing Delays in Diagnosis



*GP visit for a symptom related to ovarian cancer

Symptom Variables

As described above, all symptoms were categorised as either 'possibly related' or 'unrelated' (i.e. 'other unrelated') to ovarian cancer, based on the literature[6-8, 11-13, 15-17, 19, 67-69, 75, 89, 91] and consultation with the clinical investigators. All of the symptoms featured on the questionnaire/interview were presented individually. Of the remaining 'possibly related' symptoms, only those that were reported by cases in significantly greater proportions than controls (any source) were presented individually (for symptoms reported at anytime, excluding those that started >2 years prior).

The symptoms were grouped (as per usual practice) under the key body systems affected by ovarian cancer. These included abdominal, gastrointestinal, urinary, gynaecological and general. 'General' included systemic symptoms such as weight loss, fatigue, loss of appetite. Details of each category are listed in Appendix VI under 'Major Symptom Groups'. As with individual symptoms, grouped symptom frequency was calculated for each quarter before diagnosis. This provided an overall impression of symptom development according to the different body systems of interest.

The Goff Symptom Index[108] was the first potential symptoms tool for ovarian cancer diagnosis to be developed and published. It is currently being evaluated prospectively

in a primary care setting. In the present study, the Goff Index was applied to the self-reported data for women aged 50 and above as per Table 3-4. GP notes did not have the duration and frequency information required for the index. Notably, difficulty eating and feeling full quickly were not specifically asked for in the present study which may reduce the sensitivity calculated for the index. Symptom frequency of 16-31 days per month was used in place of the >12 times per month criterion.

Table 3-4 Goff Symptom Index

	Goff Index	**Thesis Data**
Symptoms	Pelvic or Abdominal Pain	Pelvic/Abdominal Pain/Discomfort
	↑Abdominal Size or Bloating	↑Abdominal Size or Bloating
	Difficulty Eating or Feeling Full Quickly	Feeling Full Quickly*
Criteria	Any 1 of these symptoms	Any 1 of these symptoms
	New in the last year	New in the last year
	Occurred >12 times per month	Occurred 16-31 days per month

*Feeling full quickly was only specifically asked about on interview, however this was combined with loss of appetite (i.e. 'loss of appetite or feeling full quickly')

3.3.11 Data Analysis

All statistical analyses were performed using STATA for windows (version 10.0, StataCorp LP, College Station TX). Each data source was analysed separately, however combined data were used to investigate missing onset dates and in the calculation of total delays (in diagnosis). A number of statistical methods have been used in this thesis. In general, the analysis consisted of:

- Descriptive statistics
- Analysis of new symptoms in a specified period
- Analysis of new symptoms regarded as a process in time

Descriptive analysis included proportions (expressed as percentages) for categorical variables, and means and standard deviations as well as medians and interquartile range, and range, for numeric variables.

The timing of symptoms was considered using several different approaches depending on the variable of interest (see Table 3-5). The cumulative distributions of symptoms that were new over the 24 months prior to diagnosis were plotted for individual symptoms and symptom groups.

Symptoms were considered as a process in time in order to investigate and quantify the potential lead time (diagnostic and actual). Starting from two years prior to diagnosis, the hazard of a new symptom was calculated as a function of time. These

were presented with the two year period before diagnosis split into quarters. For some of the tables, 15-23 months were combined since case-control differences were negligible over this period. The estimated continuation odds ratio was based on the two-by-two table of cases and controls with symptoms during the quarter, and without symptoms at the beginning of the quarter. Only a small number of controls had symptoms in any one quarter, therefore we assumed that the proportion of new symptoms in controls in each quarter was the same (i.e. constant), and used $1/8^{th}$ of the proportion with a new symptom in the 24 months. This approach had the advantage of reducing the variance for the continuation odds ratios, which can be large with small numbers. Analysis of the data (and in particular the data on combined symptoms) showed that this was a very reasonable assumption. Confidence intervals were based on normal approximation to log-odds ratio using a specially derived formula for the standard deviation taking into account the averaging of the control data (see section on continuation odds ratios).

As described below, the decision was made to concentrate on symptoms incident in one year during 3-14 (i.e. 3 to 14.99) months prior to diagnosis. These binary 'exposure' variables were considered in terms of the proportion of cases with symptoms (sensitivity), the proportion of controls with symptoms (1-specificity), and the odds ratios of ovarian cancer associated with symptoms. The latter was estimated via logistic regression both univariately and multivariately. Statistical significance for univariate variables was assessed using a Chi-Squared test (or Fisher's Exact test for small expected values).

Table 3-5 Timing Categories for Analysis

Symptom Variable	Timing (months prior to diagnosis)
Symptom Incidence	0 to 24
Symptom Lead Time	0 to 24
Sensitivity & Specificity	3 to 14.99 (i.e. 1 year)
Delays	0 to 15
Symptoms at Diagnosis	Anytime*

*Must have been present during 12 months prior to diagnosis

Two separate analyses groups were formed. The main analysis group included all symptoms reported (regardless of onset date). The second analysis excluded symptoms that started more than two years before diagnosis or consent. This was based on the assumption that symptoms starting more than two years before diagnosis were unlikely to be associated with ovarian cancer.

Symptom Lead Time

Symptom frequencies, and crude odds ratios (OR) with 95% confidence intervals (95%CI) were calculated for symptoms listed on the questionnaire and some additional 'possibly related' symptoms for each prediagnostic period. Logistic regression was used to estimate the probability of symptoms occurring in each period. ORs (crude) were also calculated for symptom groups. The number of borderline cancers (n=34) was considered too small for any subgroup analysis.

As mentioned previously, symptom lead time was classified as either diagnostic lead time or actual lead time. In this study, the median time from referral to diagnosis was 1.6 months (IQR 1.1-2.5). To ensure that our results would be applicable to a clinical setting, this period was accounted for in the assessment of actual lead time and symptom incidence (over one year). For actual symptom lead time, this period was assumed to be 2 months. For symptom incidence over one year, this was assumed to be 3 months. This additional month was added to be conservative because it is unlikely that one month or less of (actual) lead time would offer any significant benefit. As such, any symptoms that started within 3 months (2.99 months) of diagnosis were excluded and only symptoms present over 3 to 14.99 months were included.

Sensitivity & Specificity for Detecting Ovarian Cancer 1 Year Before Diagnosis – A Novel Definition Based on Cumulative Symptom Incidence

The sensitivity and specificity of identifying women with ovarian cancer using new symptoms cannot be easily estimated from retrospective data. The terms sensitivity and specificity typically relate to the performance characteristics of a screening test which is usually cross-sectional. For the purposes of this study, we were interested in calculating cumulative symptom incidence over 1 year since 'targeted screening' is most likely to be offered for incident not prevalent symptoms. If women in the general population were to be asked at random if they had symptoms, then the proportion of 'controls' with symptoms on any given day would be 1-specificity, however this would refer to symptom prevalence not incidence. The sensitivity would be the proportion of women with (as yet) undetected ovarian cancer who had symptoms, however this is currently unknown. Hence, sensitivity in this study has been defined as the proportion of women with ovarian cancer who had any given symptom(s) in question. Generally, this was applied to symptom incidence over 1 year. 1-Specificity has been defined as the proportion of controls who reported incident symptoms over 1 year. In practice, one would not pick women from the general population at random, but rather, wait until they present to primary care with new symptoms. One might then ask what proportion

of women aged 50-74 would present each year, and what proportion of the cases that would have been diagnosed during the next 12 months would have symptoms.

Sensitivity and specificity (for cumulative symptom incidence) were calculated for symptoms present over 1 year, excluding the first 3 months before reference date (i.e. 3-14 months). As mentioned above, this period was felt to provide a better estimate of symptom sensitivity with 'real' potential lead time.

Early versus Late Stage Symptoms

Differences in symptom incidence for early versus late stage disease were analysed using a log-rank test.

Delays in Diagnosis

Delays were assessed individually for each source as patient delay, provider delay and total delay (see section 3.3.10 for definitions). Calculations were performed using an approach that aimed to exclude pre-existing symptoms that were unrelated to ovarian cancer. Specifically, only 'possibly related' symptoms that appeared for the first time in the 15 months before diagnosis were included in delay calculations. Hence, first GP visit was the date of first GP visit for a 'possibly related' symptom that started in the 15 months before diagnosis. This cut-off of 15 months prior to diagnosis was based on the longest interval before diagnosis that case-control differences appeared to diverge in the main results.

Note that by using this approach, women who had no 'possibly related' symptoms that started in the 15 months before diagnosis, were excluded from the analysis of delay (i.e. were considered to have no delay). Likewise, women who had no visits to the GP or no GP notes available were excluded from the calculation of provider delays. In addition, the first GP visit for interview and questionnaire was self-reported, hence, if this date was missing; patient and provider delays could not be calculated. Screen-detected and incidentally diagnosed women were excluded from delay calculations.

Total delay was also derived using combined data; this was calculated using the first symptom onset date (earliest from all sources) to date of diagnosis. Provider delays were further examined by quantifying the interval between first GP visit for a symptom and date of (1) first referral to any specialist and (2) first appropriate referral (i.e. referral to gynaecology). This allowed closer examination of how long it took from initial GP visit to referral. The most appropriate referral is rapid referral G/O however, for the purposes of this analysis, appropriate referral was defined as first referral to a gynaecology department. If this date was not known, the date that ovarian aetiology

was first suspected (i.e. 'cut-off' date) was used instead. Only GP note data were used to estimate time to referral since self-reported data were considered to be too unreliable. If date of referral was not known, time to referral was not derived.

Goff Symptoms Index

The Goff symptoms Index[108] rules were applied to self-reported data for symptoms of any onset date and symptoms present 3-14 months before diagnosis.

Symptoms Indices

In an exploratory analysis, two crude symptom scores were developed. The first was based on univariate analysis results, the second utilised multivariate analysis. For the multivariate analysis, backwards stepwise selection regression was performed on the questionnaire data (for symptoms present at anytime) using p=0.05 as the significance level for entry into the model and p=0.1 as the criteria for removal from the model. Symptoms that were dropped because they predicted case status perfectly were forced back into the model.

Continuation Odds Ratios

Any potential symptoms tool would be based on new symptoms since one would not test repeatedly (within a given time frame) for the same symptom. A continuation odds ratio[131] was calculated to examine the risk associated with having a (new) symptom and developing disease within the next *x* number of months (similar to a hazard). In this scenario, the timing element is crucial and women who already have the symptom are not of interest. Continuation odds ratios (cOR) were calculated based on the women who were still 'at risk' of developing a particular symptom at each time point. As such, women who had already reported the symptom at any previous time point (including longstanding symptoms and symptoms with missing start dates) were excluded from the total number of women 'at risk'. Given that the rate of new symptoms in controls was assumed to be constant, an average rate was calculated for each symptom and used to calculate the cORs and 95%CIs, in the place of the actual numbers of controls in each period. Half a control was added to each calculation to provide stability and allow derivation of odds ratios for symptoms that were not reported by any controls. The cORs for women who never reported the symptoms (i.e. 'none') were not calculated. The following is an example of how a continuation odds ratio is calculated.

Months Before Diagnosis	Cases	Controls
None	50	210
0-2	100	5
3-5	***50***	***5***
6-8	20	5
9-11	10	5
12+	20	20
TOTAL	**250**	**250**

Assuming we wish to calculate a continuation odds ratio (cOR) for the period 3-5 months before diagnosis:

$$cOR = \frac{50/150}{5/215}$$

The continuation odds is the number who develop the symptom in the interval (3-5 months), divided by the number (of those who have not already developed the symptom) who do not develop the symptom in the interval. Note that the odds are calculated using the number of women still 'at risk' of developing the symptom, as the denominator (i.e. women who have not reported the symptom plus the number of women who report the symptom at the next time point of 0-2 months). Also, for the calculations in the present study the equation was slightly more complicated as women with missing onset dates and symptoms present for >2 years were subtracted from total number of cases and controls, and half a control was added when calculating the average rate. A specific example using actual data from the study is in Appendix IX.

Severity & Frequency

Analyses of symptom severity and frequency were limited to self-reported data only as these data were not readily available in GP notes. This was performed for 'possibly related' symptoms only and was done on an individual and group level. Symptom frequency was not analysed for symptoms with unusual timing patterns or for symptoms in which frequency is irrelevant (in the context of 'targeted screening' based on incident symptoms). Specifically these symptoms were lump in abdomen, change in bowel habit, irregular vaginal bleeding, vaginal discharge, weight loss and postmenopausal bleeding (PMB). Symptom severity was examined with the exception of irregular vaginal bleeding, PMB, vaginal discharge, change in bowel habit, and lump in abdomen.

Urinary Symptoms

Urinary frequency and urgency are the most common urinary symptoms reported in previous studies of ovarian cancer (presumably due to tumour pressure effects).[6-8, 12, 13, 69, 75, 90] Hence, these were separated out in the surveys and analysis. Urinary symptoms such as incontinence, retention, dysuria or haematuria were also databased but as a separate group named 'urinary other'. Urge incontinence, even in the presence of urgency, was classified as 'urinary other'. This was based on the rationale that investigations to confirm detrusor instability or overactive bladder were likely to have been performed if documented as 'urge incontinence' rather than 'urgency'. If the terms 'UTI' (urinary tract infection) and 'cystitis' were recorded without specific symptom details, they were also grouped under 'urinary other' even though both are usually associated with frequency. In order to check that this did not create any bias; an analysis with these terms grouped with 'urinary frequency/urgency' was performed but produced similar results. Both classification rules produced highly significant case-control differences.

Screen-Detected Women

Women who were detected via screening studies were included in the main analysis but were excluded from delays in diagnosis calculations. This included women from both the population-based screening study (UKCTOCS) and high-risk women from the UKFOCS study (UK Familial Ovarian Cancer Study). The rationale for their inclusion in the main study was that almost all screen-detected women reported at least one possibly related symptom on at least one data source. In addition, often their data was already collected before it was known that they were screen-detected.

Exclusions

In a subgroup analysis, cases were restricted to postmenopausal women aged 45-74, with primary invasive ovarian or tubal cancer who had consented before or on date of diagnosis. Controls were also restricted to women aged 45-74. Data on the following were dropped:

- Screen-detected women (via UKCTOCS or UKFOCS)
- Questionnaires that were completed after diagnosis
- Version 2 of the questionnaire (most women completed version 3)
- Interviews performed after chemotherapy started or >3 months after diagnosis (cases) or consent (controls)

Results from this restricted analysis did not differ greatly from those in the main analysis, hence were not presented.

Missing Data

There was a sizeable amount of missing data on the self-reported data (i.e. questionnaire and interview). Most notably these included symptom onset dates, severity and frequency (see Table 3-6). Missing onset dates were derived from duration data (as per Appendix VII) whenever possible. Unfortunately, this was not possible if duration data were missing or if duration was recorded but symptoms were not ongoing. Data in Table 3-6 show symptoms with missing data that could not be derived in the data cleaning process. A small proportion of missing data could be ascribed to various symptom attributes such as symptoms that fluctuated with activity or meteorological seasons. However, the majority were missing with no explanatory information, and did not appear to be missing at random. If the data were missing at random we would have found the distribution of onset dates (for cases and controls separately) in the interview to be similar for those with missing and non-missing dates on questionnaire. Rather, the missing onset date tended to be more than 2 years before reference date, particularly for controls. This was of greatest concern for control questionnaire data since 34% of symptom onset dates were missing, and therefore would be unaccounted for in calculations involving lead time. By contrast, missing onset dates were much rarer in cases (8%).

Although one might expect there to be minimal missing data from the interview, this did not occur in our study as some women simply could not remember some of the symptom details requested. Severity and frequency data were missing on interview in slightly high proportions for cases compared with controls, whereas the converse was true on questionnaire. It is not clear why this occurred, although interviews in cases were certainly more complicated and drawn out than those in controls.

As mentioned above, combining the data sources for analysis was felt to be beyond the scope of this thesis. However, for missing data, interview answers were used to gain insight on possible reasons why symptom onset dates were missing on the questionnaire.

Interviews were performed in 52 of the missing case symptoms and 108 of the missing control symptoms. Of these, 17% (9/52) of the symptoms for cases and 11% (12/108) for controls, were not reported on interview. For those that did have the symptom recorded on interview, 6% (3/52) still had missing onset dates for cases and 10%

(11/108) for controls. Table 3-8 shows data on the informative interview answers for symptoms that had a missing start date on the questionnaire.

Table 3-6 Proportion of Symptoms with Missing Symptom Onset Date, Severity or Frequency for Self-reported Data

	Number Missing/Total Number (percentage)			
	Missing Data		Unexplained Missing Data*	
	Cases	Controls	Cases	Controls
Missing Onset Date				
Questionnaire	115/1487 (8)	161/479 (34)	110/1482 (7)	161/470 (34)
Interview	28/973 (3)	73/461 (16)	12/957 (1)	16/404 (4)
Severity**				
Questionnaire	148/1372 (11)	110/470 (23)	133/1357 (10)	90/450 (20)
Interview	155/897 (17)	58/449 (13)	143/885 (16)	44/435 (10)
Frequency***				
Questionnaire	155/1288 (12)	114/461 (24)	134/1267 (11)	88/435 (20)
Interview	209/858 (24)	154/445 (35)	150/799 (19)	56/347 (16)

*Excludes symptoms that were mild and transient, activity-dependent, episodic, seasonal, intermittent or started more than 2 years prior.
**Excludes change in bowel habit, irregular vaginal bleeding (IVB), vaginal discharge, abdominal lump, postmenopausal bleeding (PMB)
***Excludes change in bowel habit, IVB, vaginal discharge, abdominal lump, PMB, weight loss
N.B. Excludes unrelated and possibly related symptoms that were considered too rare to present individually

Table 3-7 Proportion of Missing Onset Dates on Questionnaire for Any Symptom*

Any Symptom*	Cases	Controls
Total number of symptoms	1487	479
Total number of symptoms with missing onset date on questionnaire	115 (8%)	161 (34%)

*Excludes unrelated symptoms and possibly related symptoms that were too rare to present individually (e.g. pain during intercourse)

Table 3-8 Symptoms with Missing Onset Date on Questionnaire with Onset Date on Interview

Number of Symptoms (percentage)		
Months before reference date	Cases	Controls
0-5	16 (40%)	11 (13%)
6-11	7 (18%)	7 (8%)
12-23	4 (10%)	4 (5%)
>24	13 (33%)	63 (74%)
Total	**40 (100%)**	**85 (100%)**

'0-5' months is 0 to 5.99 months etc.

In Table 3-6, it can be seen that cases completed the questionnaire better than controls (in terms of there being less missing data). For controls, of the symptoms for which a start date was recorded on interview but missing on questionnaire, 74% were reported as starting >2 years before consent. These data indicate that most control symptoms with missing onset dates on questionnaire were either trivial or longstanding. Conversely, case symptoms with no onset dates on questionnaire were usually reported on interview to be either within 0-5 months of diagnosis (40%) or longstanding (33%).

Based on these findings, the decision was made to treat symptoms with missing onset dates as if they were longstanding (i.e. present for >2 years) and exclude them from the symptom lead time calculations. The rationale for excluding longstanding symptoms from the main study results is detailed in section 4.1 Results. This was a relatively conservative approach given that symptoms in cases were actually more likely to start within the two years. Symptoms with missing onset dates were still included in the results showing symptoms present at anytime (i.e. not split into prediagnostic periods). Since the data were collected from multiple sources, these issues with missing values will be addressed more formally in a future combined analysis.

Missing severity and frequency were only a minor concern since they were not a main outcome. According to linear tests for trend, missing severity and frequency were more likely to be mild or infrequent (1-4 days per month) on both data sources. This demonstrated that severity and frequency were not missing at random. However, the fact that symptoms with missing severity and frequency were more likely to be mild or

infrequent was considered to be sufficient to justification to assume that these symptoms were less important.

Combined Analysis

Combining the three data sources for analysis would require development of a special algorithm to take into account the various advantages and disadvantages of each source. In addition, construct validity for the questionnaire and interview and intraindividual variation between the sources would need to be considered. Such an analysis was felt to be beyond the scope of this thesis and is planned to be completed as part of continued work on the study data.

4 CHAPTER 4: Case-Control Results

4.1 Results

A total of 263 newly diagnosed cases and 268 controls were initially included in the study. Nine of the cases were later confirmed to be primary peritoneal, leaving a total of 254 eligible cases. Demographic details are shown in Table 4-1. Median age at study entry was 64 years for both groups and the majority of women were White. One case was diagnosed 13 days before her 45th birthday (aged 44), which is below the age threshold, but was included since her data had already been collected and the age difference was marginal.

Figure 4-1 provides details of the distribution of data sources collected. Almost all subjects completed a questionnaire, and interview and GP note data were obtained for a similar proportion of cases and controls in each group.

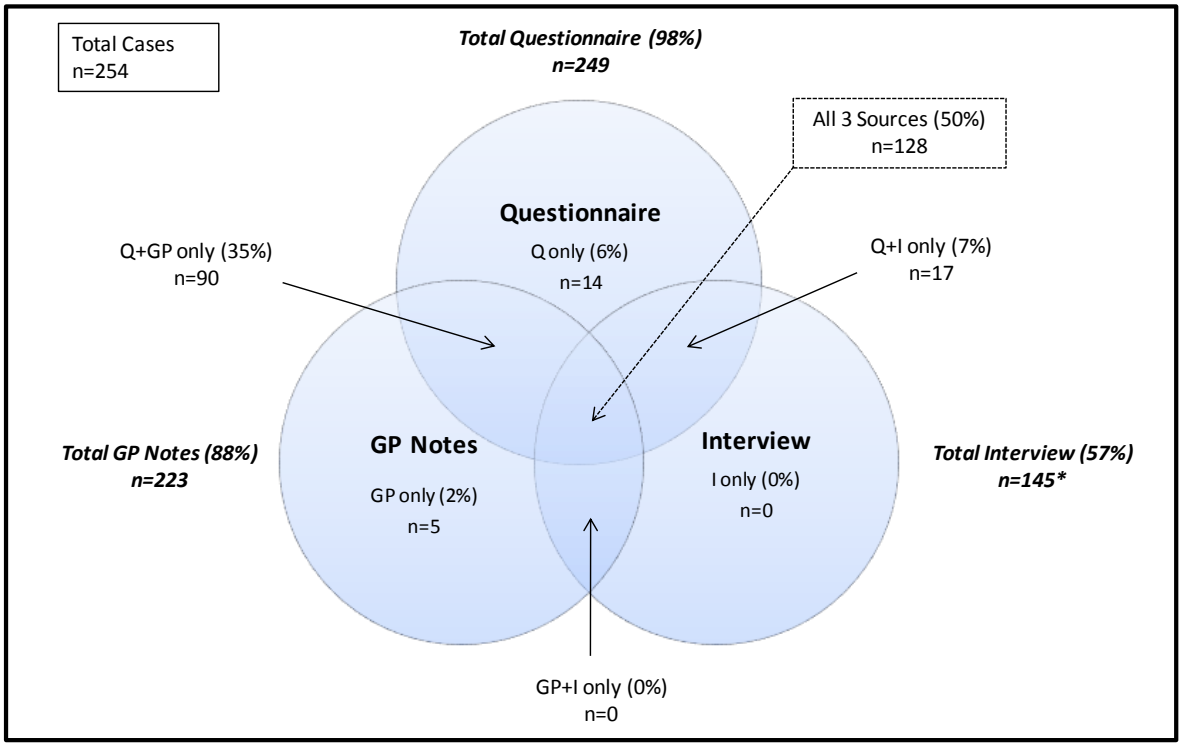
Table 4-1 Demographics

	Number (%)	
	Cases n=254	Controls n=268
Age in years (at diagnosis or consent)		
Median (IQR)	64 (56-72)	64 (57–72)
Mean (±SD)	64 (10)	65 (8)
Range	44 - 90	52 – 78
Age Distribution		
44-50 years	30 (12)	-
51-55 years	26 (10)	33 (12)
56-60 years	39 (15)	63 (24)
61-65 years	41 (16)	50 (19)
66-70 years	46 (18)	42 (16)
71-75 years	33 (13)	52 (19)
76-80 years	23 (9)	28 (10)
81-85 years	13 (5)	-
86-90 years	3 (1)	-
Ethnicity		
White	247 (97)	263 (98)
Other	7 (3)	5 (2)
Centre		
Belfast	30 (12)	39 (15)
Bristol	43 (17)	67 (25)
East Kent	39 (15)	-
Gateshead	17 (7)	16 (6)
Manchester	2 (1)	6 (2)
Middlesbrough	3 (1)	8 (3)
North Wales	12 (5)	50 (19)
Portsmouth	6 (2)	3 (1)
Southend	38 (15)	-
London*	64 (25)	79 (29)
Menopausal Status		
Post-menopausal	215 (85)	268 (100)
Pre- or Peri- menopausal	30 (4)	-
Unknown	9 (12)	-
Relevant Medical History (GP notes)		
IBS	14/223 (6)	16/227 (7)
Diverticular conditions**	23/223 (10)	21/227 (9)

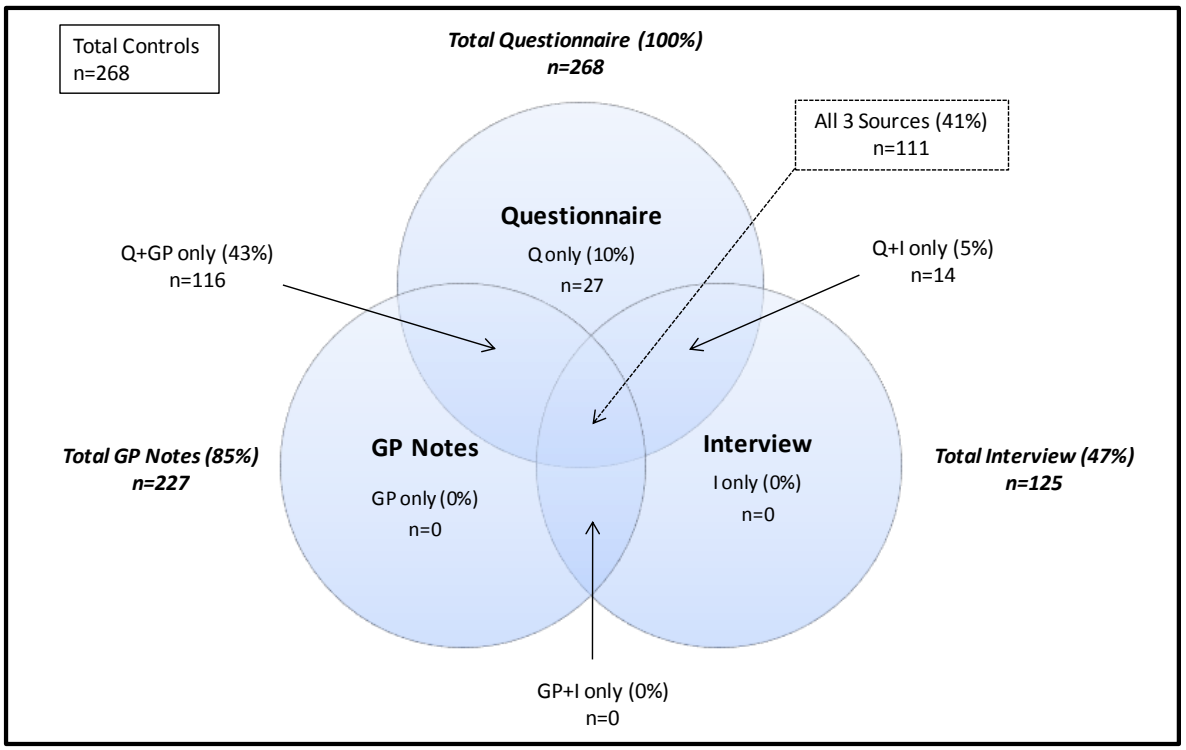
*Cases were from University College London Hospital, Controls were from Barts and The London NHS Trust

**Includes diverticular disease, diverticulosis, diverticulitis

Figure 4-1 Distribution of Data Sources Collected



*Includes 5 face-to-face interviews



Cases included 254 women with primary ovarian or fallopian tube cancer, 34 of which were borderline tumours. Most of the tumours were epithelial (98%). As shown in Table 4-2, 40% of cases were early stage (I-II) at diagnosis and 53% were advanced stage (III-IV). Seven percent of cases were unstaged, and these were predominantly borderline cancers (12/17). As mentioned previously, date of diagnosis was taken as the date of surgery or date of first positive biopsy. Most cases were consented before or on the date of diagnosis (198/254, 78%). Of the remaining cases, 18 were recruited more than one month after diagnosis. Two cases were consented more than 4 months after diagnosis but this included (2 of the 3) women who had already completed upfront chemotherapy (see 3.3.9 Data Cleaning). The median time between 'cut-off' date (i.e. date ovarian aetiology was first suspected) and diagnosis was 1.3 months (IQR 0.7-1.9).

CA125 data were available for 92% of cases. Serum levels were raised in 87% of early stage and 98% of late stage cancers (see Table 4-3). The proportion of cases and controls with documented IBS or diverticular conditions was similar. Dates of diagnosis for IBS or diverticular conditions were available for only a proportion of cases (IBS n=7, diverticular n=17) and controls (IBS n=5, diverticular n=14). Of the cases with, only 2 of the 7 had a diagnosis of IBS that was new in the year before cancer diagnosis. For diverticular conditions this was 6 out of 17 cases.

Table 4-2 Case Details

	Number (percentage)	
FIGO Stage		
I	85 (33)	Early
II	17 (7)	102 (40)
III	107 (42)	Late
IV	28 (11)	135 (53)
Unstaged*	17 (7)	
Tumour Grade		
1	25 (10)	
2	47 (19)	
3	111 (44)	
Ungraded	181 (71)	
Tumour Type		
EOC**	250 (98)	
Invasive	216 (86)	
Borderline	34 (14)	
Non-EOC	4 (2)	
Histology ICD-O Classification		
Ovary (C56)		
Carcinoma		
Serous	118 (46)	
Mucinous	37 (15)	
Endometrioid	24 (9)	
Clear Cell	17 (7)	
Adenocarcinoma NOS	11 (4)	
Other Specified Carcinoma	27 (11)	
Unspecified Carcinoma	4 (2)	
Other Specified Malignant Neoplasm***	10 (4)	
Sex Cord-Stromal	5 (1)	
Fallopian Tube (C57.0)		
Carcinoma		
Serous	1 (0)	

*Borderline (n=12), granulosa cell (n=2), primary invasive ovarian cancer (n=3, two of whom were receiving upfront chemotherapy so staging was not yet performed)
**Includes synchronous cancers (n=3)
***Mixed mullerian tumours (n=10)
Abbreviation: EOC is epithelial ovarian cancer

Table 4-3 CA125 Levels (Cases)

	Not Available	CA125 Available		Total
		Normal	Elevated	
Early	5 (5%)	13 (13%)	84 (87%)	102
Late	12 (9%)	3 (2%)	120 (98%)	135
Unstaged	3 (18%)	3 (21%)	11 (79%)	17
Total	**20 (8%)**	**19 (8%)**	**215 (92%)**	**254**

Note: CA125 levels ≥30 units/ml (postmenopausal) and ≥35 units/ml (pre-menopausal) were considered to be raised

Three important attributes of the symptom data were noticeable immediately:

1) Relative to cases, controls reported more symptoms that started >2 years before reference date (i.e. longstanding symptoms), particularly on interview (see Figure 4-2).

2) The proportion of symptoms 12-24 months before diagnosis (or consent) was extremely similar amongst cases and controls (see Figure 4-2),

3) Controls demonstrated a relatively constant rate of new symptoms over the two years prior to consent (see Figure 4-3),

Implications of 1) are that symptom specificity is underestimated if longstanding symptoms are included. The effects of this were greatest in the interview data. Table 4-4 shows the dramatic change in odds ratios when longstanding symptoms are omitted from interview data. Notably, the proportion of controls with bloating at 0-2 months before consent is >3 times larger if longstanding bloating is included (7% versus 23%, see Figure 4-2).

Table 4-4 Crude Case-Control Odds Ratios (95%CI) Including versus Excluding Longstanding Symptoms Reported on Interview (Anytime)

Symptom	Incl. >2 Year*	Excl. >2 Year**
Abdominal Bloating	**2.44 (1.49, 3.98)**	**5.93 (3.37, 10.43)**
Constipation	1.62 (0.97, 2.71)	**3.95 (1.93, 8.09)**
Indigestion	1.04 (0.64, 1.67)	**2.03 (1.18, 3.50)**
Fatigue	1.61 (0.99, 2.62)	**3.43 (2.05, 5.74)**

*Including symptoms that started >2 years before reference date
**Excluding symptoms that started >2 years before reference date
Bolded values have 95% confidence intervals above 1.0

Figure 4-2 Interview Symptom Prevalence Per Prediagnostic Quarter With & Without Longstanding Symptoms (Vertical Bars are 95% Confidence Intervals)

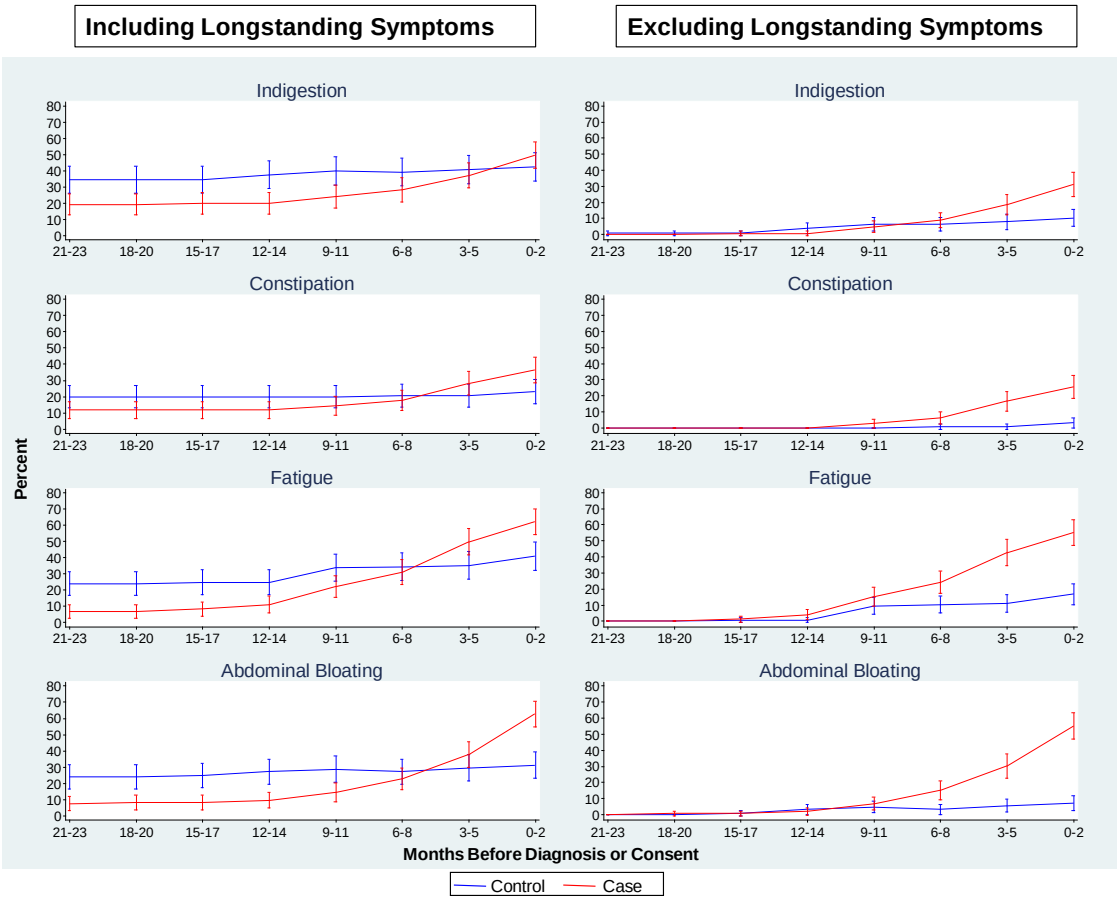
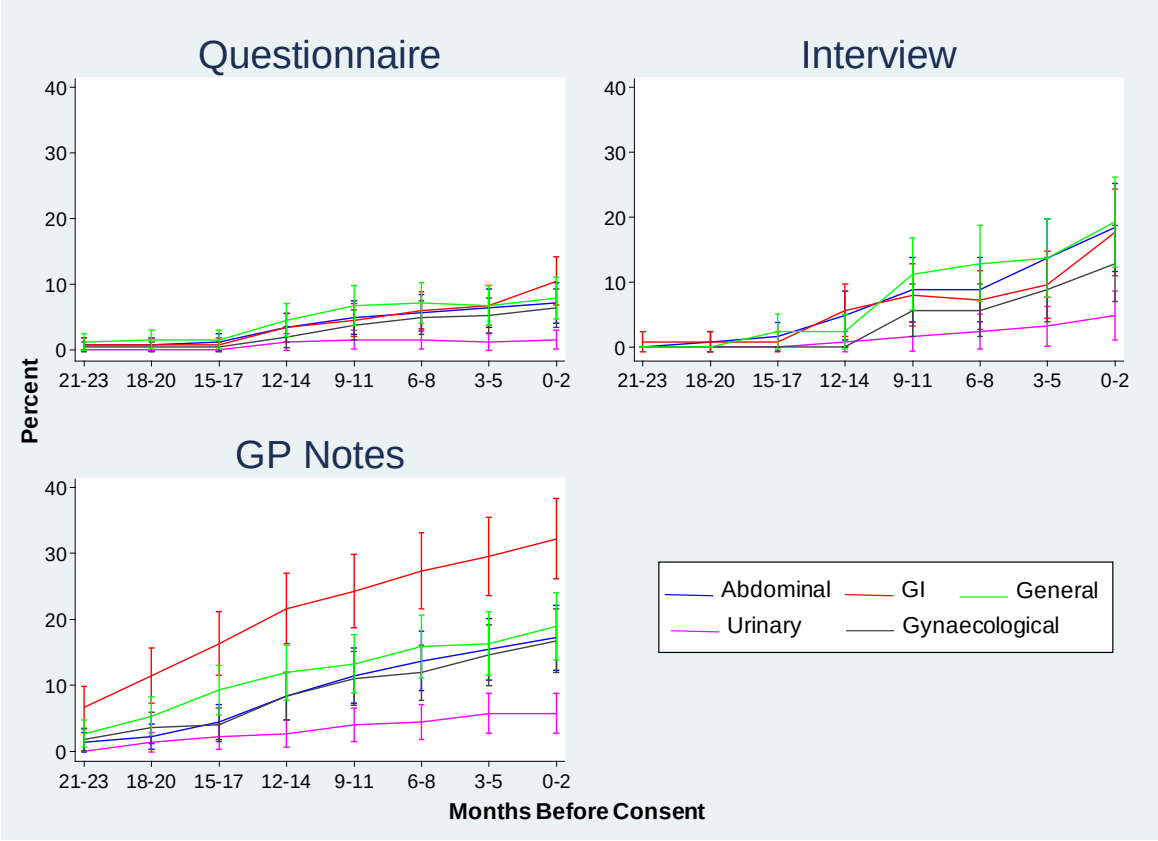
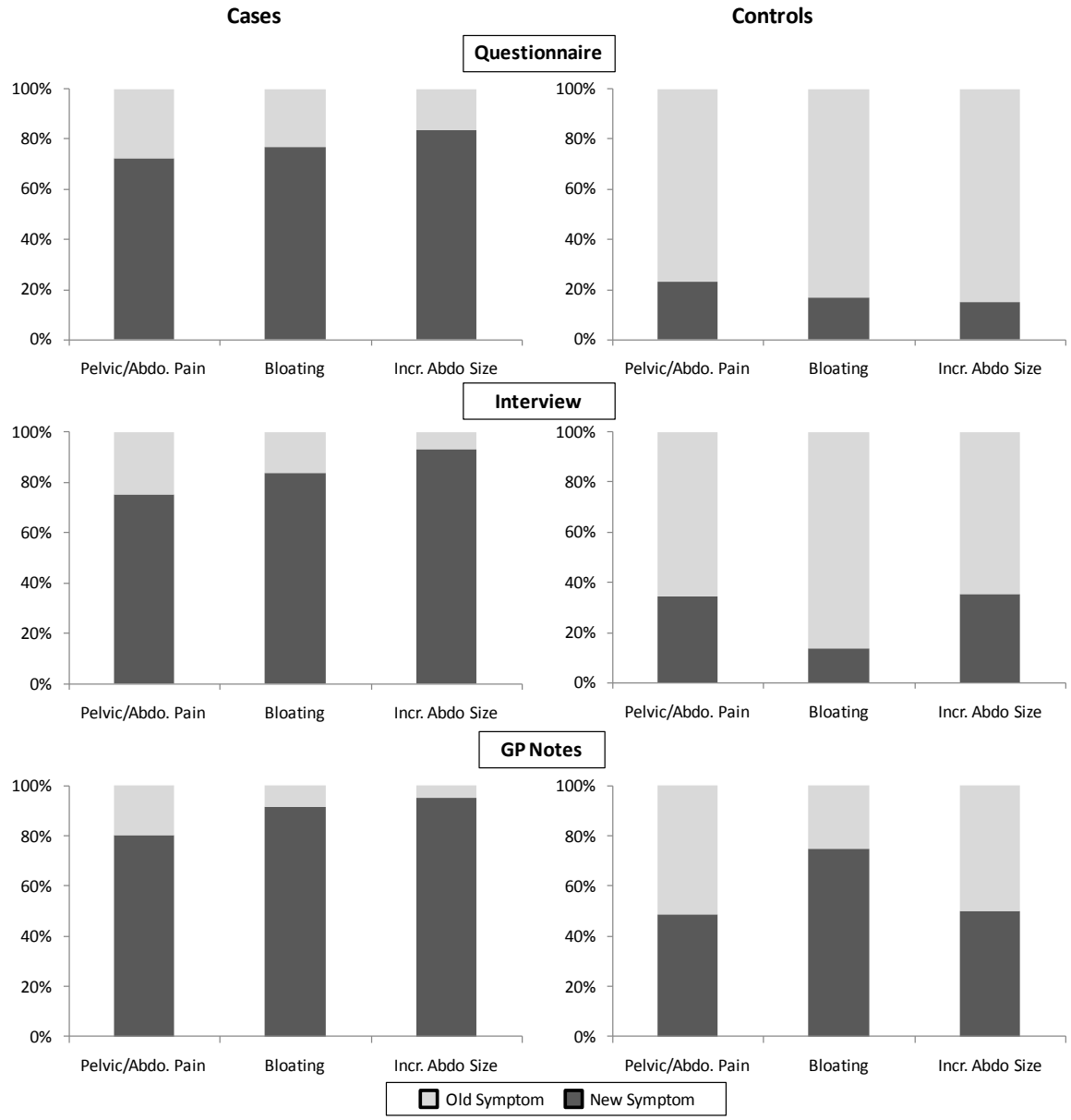


Figure 4-3 Control Grouped Symptom Prevalence Per Prediagnostic Quarter Excluding Longstanding Symptoms for Each Data Source (Vertical Bars are 95% Confidence Intervals)



Another point of interest was that the majority of salient case symptoms were new in the year before diagnosis according to all data sources (Figure 4-4). In contrast, the same symptoms in controls typically started more than one year before the reference date.

Figure 4-4 Proportion of 'New' versus 'Old' Symptoms in Cases & Controls for 3 Common Symptoms Per Data Source



Note: Note: All symptoms were present during the year before reference date, including longstanding symptoms. 'New' denotes symptoms that first started during the year before reference date. 'Old' denotes symptoms that first started more than one year before the reference date.
Abbreviations: Incr. is increased

Any referral algorithm using symptoms would be based on the first presentation for any given symptom (or symptoms). That is, one would not test repeatedly for the same symptom (within a certain timeframe). For these reasons, symptoms that started more than two years prior to reference date were omitted from the main graphs and separated out in the main tables. Also, symptom graphs were plotted as the cumulative incidence of new symptoms (i.e. new within the two years prior to reference date). This approach has the advantage of showing the proportion of controls that would be 'included' in each year if one were to offer testing based on symptoms.

Finally, given that the rate of new symptoms in controls was approximately constant, an average rate per month was estimated for each symptom. These estimates were used (after appropriate weighting) in place of the actual proportions of controls when calculating odds ratios for different periods before diagnosis. This allowed for a more direct comparison of the risk, and therefore lead time associated with individual symptoms for each prediagnostic period.

4.1.1 Questionnaire

Although it was requisite for all study participants to fill out a questionnaire, these were not received for four cases. An additional case questionnaire was dropped due to unreliable data since all symptoms were recorded as starting after diagnosis despite several prediagnostic symptoms being present in the GP notes. This resulted in questionnaire data for a total of 249 (98%) cases and 268 (100%) controls. The majority of women completed the questionnaire before, or on the day of diagnosis (cases 77%) or consent (controls 99%). Of the 58 cases who completed the questionnaire after the date of diagnosis, 3 had already received neoadjuvant chemotherapy and the remaining 55 were either scheduled for neoadjuvant chemotherapy (i.e. had positive biopsy but not treatment) or had malignant biopsy results before the primary site of cancer was known. Eighty-nine percent of cases and 90% of controls completed version 3 of the questionnaire, and the remainder completed version 2. Version 2 was used between February 2006 and August 2006, in 26 cases and 20 controls. Version 3 was used from June 2006 to April 2008, in 223 cases and 248 controls.

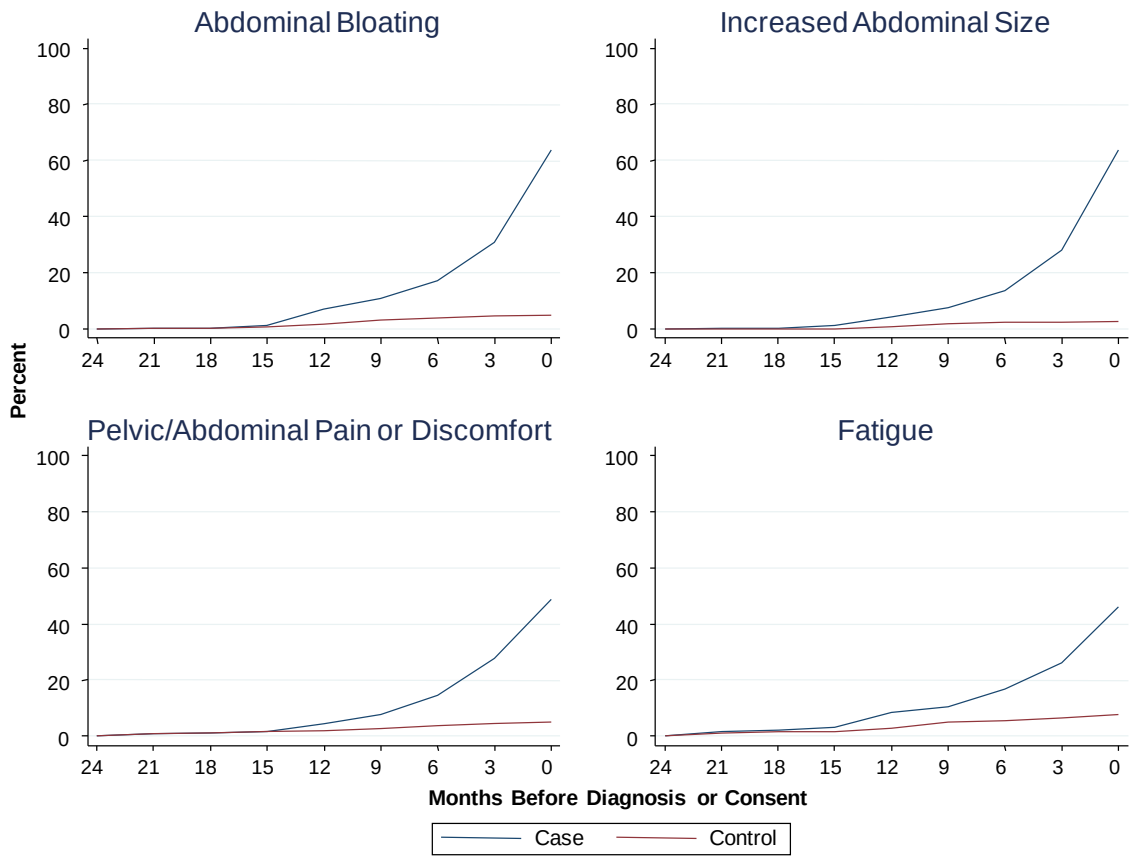
The proportion of women who received assistance from the study nurses to complete the questionnaire was unknown, and it was difficult to control this aspect of data collection, especially given that the main drive for UKOPS was to obtain biological rather than epidemiological endpoints. Study nurses were under pressure to process blood samples for the main study within strict time constraints.

Among those who completed a questionnaire, 98% of cases versus 62% controls reported at least one symptom in the year before diagnosis or consent. Once longstanding symptoms were excluded, this changed to 96% and 49% for cases and controls, respectively. Of the cases that reported no symptoms on questionnaire, all had symptoms possibly related to ovarian cancer in their GP notes or interview. Symptom reporting was generally higher in cases than controls. Overall, the most commonly reported symptoms at anytime (excluding longstanding) in cases were abdominal bloating (69%), increased abdominal size (67%), pelvic/abdominal pain or

discomfort (53%) and fatigue (50%). In contrast, the most frequent control symptoms were back pain (16%), indigestion (16%), fatigue (15%) and abdominal bloating (13%). Note that these percentages refer to symptoms reported at anytime, and therefore do not relate directly to the cumulative frequency graphs.

Figure 4-5 shows cumulative frequency for the most common case symptoms over the two years before diagnosis. Longstanding symptoms are excluded, so each line represents symptoms that were new in the two years before reference date. Strikingly, cases and controls start to diverge around 15 months prior to reference date for all four symptoms. Constipation, loss of appetite, diarrhoea, nausea/vomiting, irregular vaginal bleeding and weight loss started to diverge from 9 months before reference date (data not shown but see Table 4-7). Urinary frequency/urgency and indigestion appeared to separate later from around 6 months before reference date.

Figure 4-5 Cumulative Symptom Frequency for the Most Common Symptoms in Cases on Questionnaire for Two Years Before Reference Date



4.1.2 Telephone Interview

Telephone interview response rates were high in both groups with 79% of newly diagnosed cases and 70% controls indicating consent on the questionnaire. Of the

cases who consented, interviews were not performed in 46. This was because women were either deceased (n=6), not well enough (n=6), not contactable within reasonable time from diagnosis (n=29), or had changed their mind (n=5). Attrition of control interviews was due to failure to make contact within timelines (n=64) and women changing their mind (n=1). This resulted in interviews conducted with 145 cases and 125 controls.

Interviews took place within 3 months of consent or diagnosis. However, there were some exceptions to this (cases n=11, controls n=7). These were mostly due to oversight, and all but two cases were interviewed within 5 months of diagnosis. An attempt was made to contact all cases who agreed to telephone interview, unless the details of consent for interview were received more than 3 months after diagnosis or consent. The majority of interviews were conducted during working hours which potentially created bias against women who were in full-time employment, however women were also contacted at work whenever possible.

Case interviews lasted a median of 18 minutes (IQR 13-24), whereas control interviews elapsed over a shorter median period of 8 minutes (IQR 5-11). Interviews ideally took place before initiation of chemotherapy however, logistically this was difficult and 32% (47/145) of women had already started treatment by the time of the interview (see Table 4-5). Of these women, the median time from starting chemotherapy to interview was 0.6 months (IQR 0.2-1.1). Median time between diagnosis or consent and interview was one month for both cases (IQR 0.8-1.7) and controls (IQR 0.6-2.0).

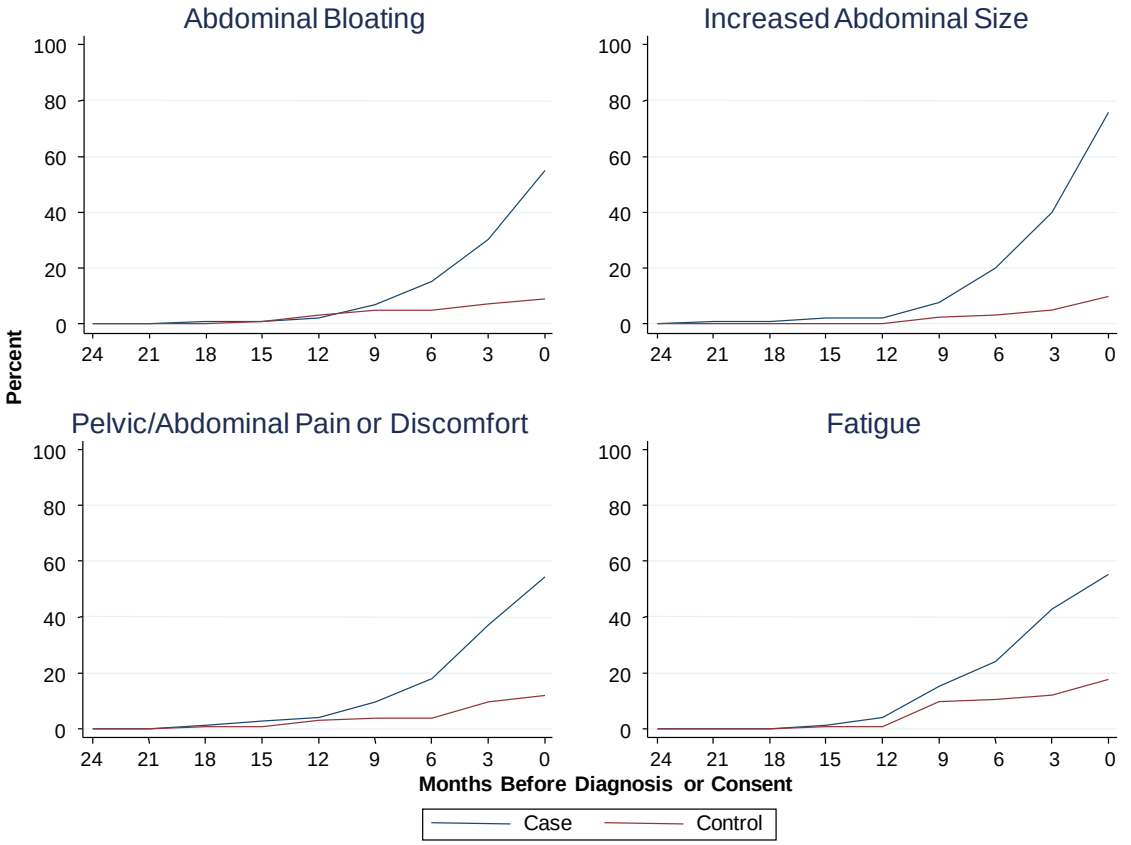
Table 4-5 Timing of Interviews

	Cases (n=145)	**Controls (n=125)**
Time from DOD/ICF Months		
Median (IQR)	1.2 (0.8-1.7)	0.9 (0.6-2.0)
Before Chemotherapy Start	98/145 (68%)	-
After Chemotherapy Start	47/145 (32%)	-

All cases reported at least one relevant symptom on interview, as did 90% of controls. Interestingly, this included more than half of the women who had recorded no symptoms on questionnaire. If longstanding symptoms were ignored, 99% of cases and 73% of controls reported at least one possibly related symptom. As with the questionnaire, increased abdominal size (77%), pelvic/abdominal pain or discomfort (57%), abdominal bloating (56%) and fatigue (55%) were the most frequently experienced symptoms for cases over any time period (excluding longstanding). For controls, these were fatigue (26%), indigestion (22%), abdominal bloating (18%), and increased abdominal size (16%).

Figure 4-6 shows cumulative incidence of the most common case symptoms reported on interview (excluding longstanding symptoms). In general, case-control differences started closer to diagnosis than those observed for questionnaire data. Pelvic/abdominal pain or discomfort started to show case-control separation from 12 months prior to diagnosis. For abdominal bloating, increased abdominal size and fatigue divergence appeared later at around 9 months before diagnosis, as did loss of appetite (feeling full quickly). Weight loss, indigestion, urinary frequency or urgency, nausea/vomiting began to separate 6 months before diagnosis.

Figure 4-6 Cumulative Symptom Frequency for the Most Common Symptoms in Cases on Interview for Two Years Before Reference Date

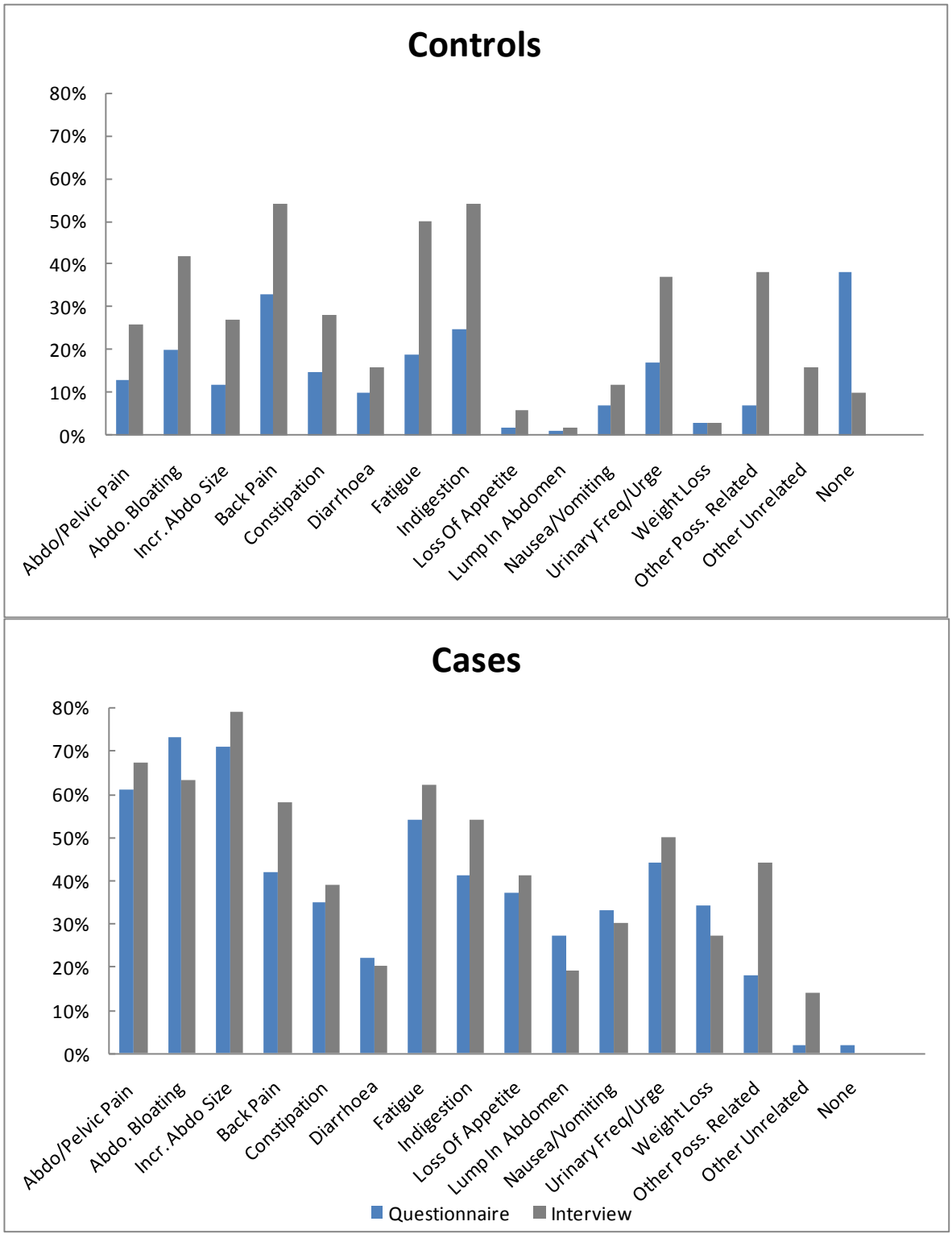


Overall, symptom reporting was higher on interview than questionnaire even after omission of longstanding symptoms (see Figure 4-7 and Figure 4-8). Back pain, bloating, pelvic/abdominal pain or discomfort, increased abdominal size, constipation, fatigue, indigestion and urinary frequency or urgency were reported in substantially higher proportions on interview for controls. However, most of these appeared to be longstanding.

This pattern of increased symptom reporting on interview was also observed in cases but to a much smaller extent. Weight loss, bloating, diarrhoea, nausea or vomiting, and

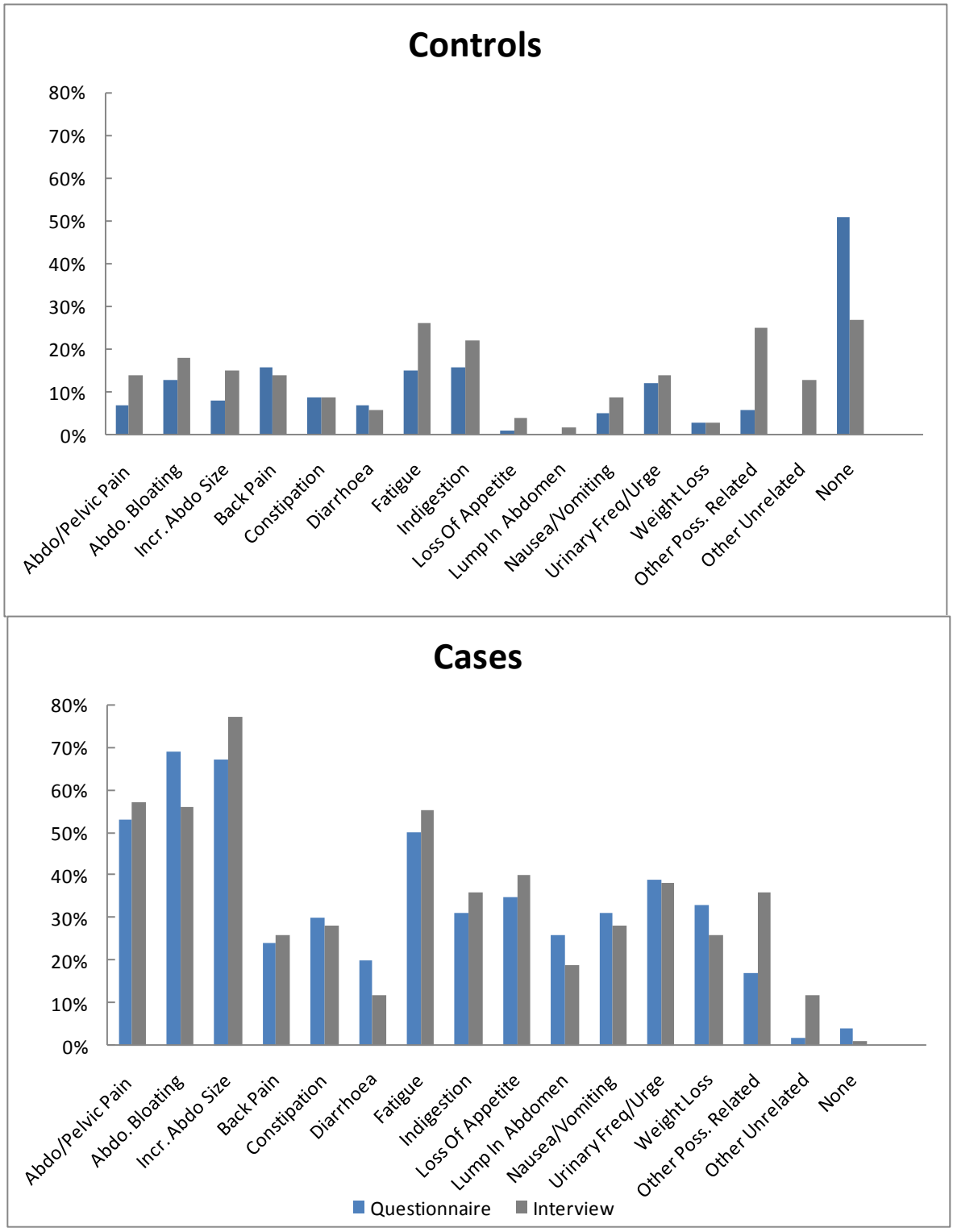
lump in abdomen were the only case symptoms less frequent on interview but all others increased.

Figure 4-7 Questionnaire versus Interview Symptom Prevalence for Cases and Controls *Including* Longstanding Symptoms



Note: 'Loss of appetite' is 'Loss of appetite or feeling full quickly' on interview

Figure 4-8 Questionnaire versus Interview Symptom Prevalence for Cases and Controls *Excluding* Longstanding Symptoms



Note: 'Loss of appetite' is 'Loss of appetite or feeling full quickly' on interview

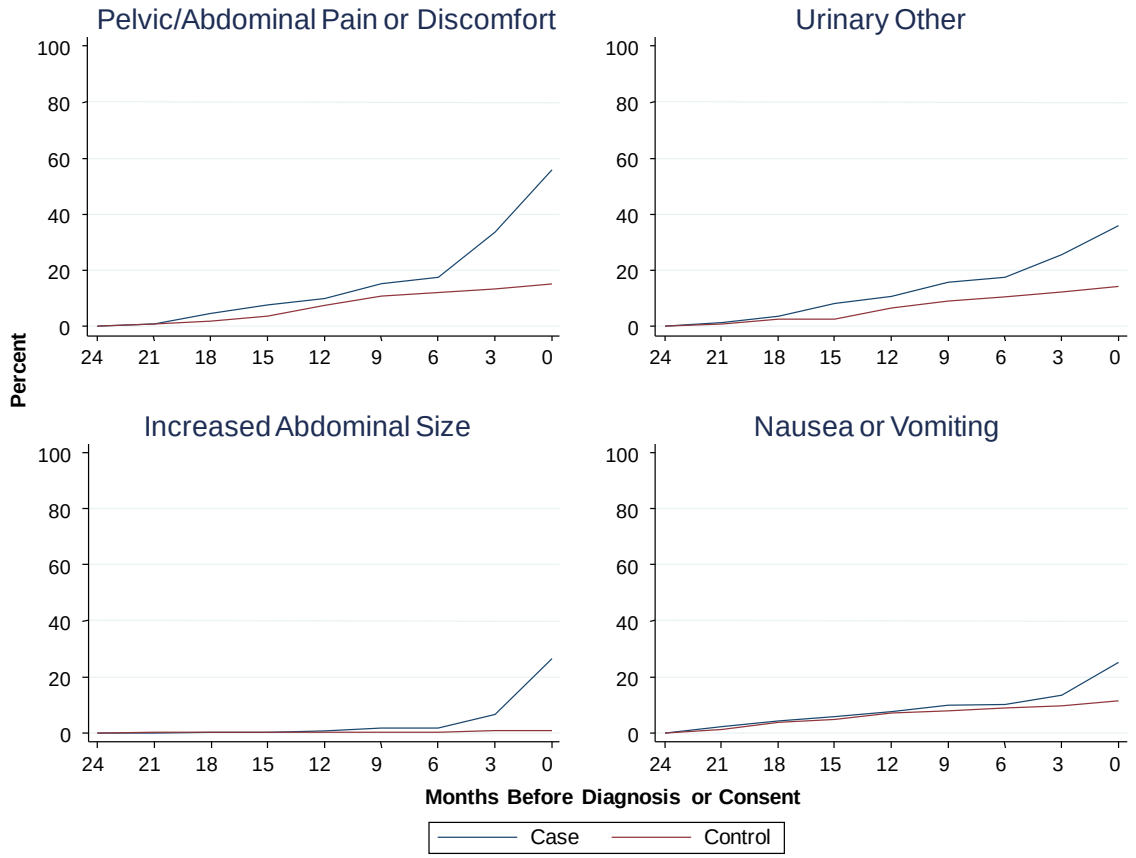
4.1.3 GP Notes

GP notes were received for 223 (88%) cases and 227 (85%) controls. Only 3% of cases had no potentially related symptoms recorded in their GP notes during the two years before diagnosis. This proportion was higher in controls with 32% of women without any potentially related ovarian cancer symptoms documented. Unlike self-reported data, persistence of symptoms was difficult to establish since there was insufficient information on symptom duration. Clearly, a single recording of a symptom 12 months before diagnosis does not necessarily mean that the symptom continued until diagnosis. However, for the purposes of this study, symptom persistence was irrelevant since women would be tested based on the first appearance of symptoms. An important point to note is that GP note data were only extracted up until a 'cut-off point' for cases (i.e. date that ovarian or pelvic aetiology was first suspected). However, the date of diagnosis was still used as the reference date for calculating the two year prediagnostic periods. Therefore, the data between 0-3 months before diagnosis do not necessarily represent symptoms at diagnosis, and overall symptom prevalence at diagnosis will be underestimated for cases.

As expected, symptom reporting/recording was lower in GP notes than the self-reported data sources. Recording of fatigue, indigestion and diarrhoea was extremely similar for cases and controls, and divergence was small even 3 months prior to diagnosis. Another distinguishing factor from the self-reported data was that urinary symptoms featured more heavily for both cases and controls. For cases, the symptoms most frequently recorded (excluding longstanding) were pelvic/abdominal pain or discomfort (58%), urinary other symptoms (37%), increased abdominal size (27%) and nausea/vomiting (25%). The most common symptoms (excluding longstanding) in control women were back pain (18%), indigestion (16%), fatigue (16%) and urinary other (16%).

Cumulative incidence of the most common case symptoms are shown in Figure 4-9. In general, symptom sensitivity was low with less than 40% of women reporting three of the most common symptoms. Separation of case-control symptoms was evident around 6 months before diagnosis for pelvic/abdominal pain or discomfort and increased abdominal size (see Figure 4-9). Constipation, change in bowel habit, bloating also showed signs of case-control divergence at this time point, however most of the remaining possibly related ovarian cancer symptoms began to separate 3 prior months to diagnosis, including nausea/vomiting.

Figure 4-9 Cumulative Symptom Frequency for the Most Common Symptoms Recorded in GP Notes for Cases over Two Years Before Reference Date



Note: 'Urinary other' includes symptoms such as urinary retention, incontinence, dysuria, haematuria, and change in urine smell or colour.

4.1.4 Symptoms At Diagnosis

Although it is of clinical interest to document the symptoms present at diagnosis of ovarian cancer, this was not a primary aim of the study. On the questionnaire, all women were asked to specify whether each symptom was 'ongoing', but this field was not always completed.

Symptoms present within one month of reference date were considered to be representative of symptoms at diagnosis. Since this was to examine all symptoms present at diagnosis (regardless of when they started), symptoms that had a missing onset date (if the symptom was ongoing) and symptoms that started more than two years before reference date were included. GP note data are not shown since visits were only examined up to a 'cut-off' date and therefore did not accurately reflect symptoms at diagnosis.

The most common symptoms at diagnosis for cases were increased abdominal size, abdominal bloating and pelvic/abdominal pain or discomfort for both sources. The vast majority of cases (97%) reported at least one relevant symptom on the questionnaire as did all cases did so on interview. This included screen-detected and incidentally-diagnosed women. According to both sources, most cases also had at least 3 symptoms. In contrast, controls reported the highest symptom frequencies for bloating, back pain and indigestion on questionnaire. This pattern was repeated on interview with the exception of bloating, which was replaced by fatigue. Almost half of control women had no symptoms on the questionnaire, but only 13% claimed the same during interview. Note that on the questionnaire, the term 'irregular vaginal bleeding' included postmenopausal bleeding and other bleeding irregularities such as post-coital bleeding and irregular menses. Odds ratios were not calculated since case-control differences at diagnosis are of lesser interest than those at least 3 months prior to diagnosis.

Table 4-6 Symptoms At Diagnosis

	Questionnaire		Interview	
Symptoms	**Cases (n=249)**	**Controls (n=268)**	**Cases (n=145)**	**Controls (n=125)**
Pelvic/Abdo. Pain Discomfort	**57%**	12%	**63%**	23%
Abdominal Bloating	**71%**	**17%**	**62%**	39%
↑Abdominal Size	**70%**	12%	**79%**	27%
Back Pain	38%	**31%**	54%	**52%**
Change in Bowel Habit	1%	0%	10%	1%
Constipation	34%	15%	36%	26%
Diarrhoea	20%	9%	18%	15%
Indigestion	38%	**21%**	52%	**52%**
Nausea or Vomiting	30%	4%	27%	9%
Lump in Abdomen	25%	1%	19%	2%
Urinary Freq./Urgency	43%	15%	50%	37%
Urinary Other*	4%	0%	26%	3%
PMB	1%	0%	8%	2%
IVB	10%	1%	5%	0%
Vaginal Discharge	2%	0%	6%	4%
Fatigue	52%	16%	61%	**49%**
Loss of Appetite**	34%	1%	41%	6%
Weight Loss	31%	1%	27%	2%
Other Possibly Related	*15%*	*7%*	*42%*	*38%*
None	*3%*	*43%*	*0%*	*13%*
Other Unrelated	*2%*	*0%*	*0%*	*0%*
Number of Symptoms				
0	8 (3%)	116 (43%)	0 (0%)	16 (13%)
1	23 (9%)	46 (17%)	3 (2%)	10 (8%)
2	16 (6%)	39 (15%)	7 (5%)	16 (13%)
3+	202 (81%)	67 (25%)	135 (93%)	83 (66%)

*'Urinary other' includes dysuria, retention, haematuria, incontinence, and change in urine colour or smell

**'Loss of appetite or feeling full quickly' on interview

Bolded values are the 3 most common symptoms per source

Abbreviations: Abdo. is abdominal, Freq. is frequency, IVB is irregular vaginal bleeding, PMB is postmenopausal bleeding.

4.1.5 Timing and Predictive Potential of Prediagnostic Symptoms

As previously inferred, any potential symptoms index would most sensibly be based on the first time specific symptoms appeared, or the first time a woman presented for those symptom(s). Table 4-7 contains continuation odds ratios (cOR) for individual symptoms which account for women who have already had the symptom, and only uses women who are still 'at risk' of developing a particular symptom at each time point. Note that each woman may have reported more than one symptom in the table. Also, the rate of new symptoms in controls was assumed to be constant (refer to Data Analysis for further details).

Overall, significantly elevated cORs were observed for most symptoms 6-8 months before diagnosis. The longest diagnostic lead time was seen on the questionnaire at 12-14 months before diagnosis for abdominal bloating (cOR 8.3 [95%CI 3.8, 18.2]), increased abdominal size (cOR 8.2, [95%CI 2.9, 22.8]), fatigue (cOR 5.7 [95%CI 2.8, 11.8]), pelvic/abdominal pain or discomfort (cOR 4.8 [95%CI 1.9, 12.4]), diarrhoea (cOR 3.7 [95%CI 1.1, 12.1]) and constipation (cOR 3.6 [95%CI 1.1, 12.3]). Urinary frequency/urgency was also present in significantly more cases than controls at the same time point in the GP notes (cOR 2.9 [95%CI 1.2, 6.8]). No symptoms were reported in significantly higher numbers of cases than controls for interview data at 12-14 months before diagnosis.

On GP notes, cORs were also significantly elevated 9-11 months before diagnosis for pelvic/abdominal pain or discomfort, urinary frequency/urgency, urinary other, loss of appetite, and irregular vaginal bleeding.

Table 4-7 'New' Symptom Frequency, Continuation Odds Ratios & 95% Confidence Intervals at Different Periods before Diagnosis

Months Before Diagnosis	Questionnaire				Interview				GP Notes			
	Cases (n=249)	Controls (n=268)	cOR	95% CI	Cases (n=145)	Controls (n=125)	cOR	95% CI	Cases (n=223)	Controls (n=227)	cOR	95% CI
Pelvic/Abdominal Pain or Discomfort												
None*	39%	87%	-	-	33%	74%	-	-	42%	85%	-	-
0-2	**21%**	0%	**73.9**	**(37.1, 146.9)**	**17%**	2%	**24.9**	**(11.2, 55.0)**	**22%**	2%	**23.9**	**(14.5, 39.4)**
3-5	**13%**	1%	**30.7**	**(15.2, 62.2)**	**19%**	6%	**18.7**	**(8.8, 39.8)**	**16%**	1%	**11.4**	**(6.9, 19.0)**
6-8	**7%**	1%	**13.1**	**(6.0, 28.3)**	**8%**	0%	**5.9**	**(2.5, 13.8)**	2%	1%	1.3	(0.5, 3.4)
9-11	**3%**	1%	**5.7**	**(2.3, 14.2)**	**6%**	1%	**3.6**	**(1.4, 9.1)**	**6%**	3%	**3.3**	**(1.7, 6.5)**
12-14	**3%**	0%	**4.8**	**(1.9, 12.4)**	1%	2%	0.9	(0.2, 3.9)	2%	4%	1.2	(0.5, 3.2)
15-23	2%	1%	0.9	(0.3, 2.8)	3%	1%	0.6	(0.2, 1.8)	8%	4%	1.4	(0.8, 2.5)
>2 Years**	8%	6%	1.6	(0.8, 3.1)	11%	12%	0.9	(0.4, 1.9)	2%	0%	4.1	(0.5, 37.2)
Missing***	4%	3%	1.6	(0.6, 4.2)	1%	2%	0.9	(0.1, 6.2)	0%	0%	-	-
Abdominal Bloating												
None*	27%	80%	-	-	37%	58%	-	-	79%	96%	-	-
0-2	**33%**	0%	**136.9**	**(68.8, 272.3)**	**25%**	2%	**34.3**	**(12.1, 96.9)**	**13%**	0%	**32.7**	**(14.8, 72.3)**
3-5	**14%**	1%	**25.5**	**(12.6, 51.8)**	**15%**	2%	**12.7**	**(4.5, 36.0)**	**5%**	1%	**12.2**	**(5.0, 29.8)**
6-8	**6%**	1%	**9.9**	**(4.5, 21.6)**	**8%**	0%	**5.7**	**(1.9, 16.9)**	1%	1%	1.9	(0.4, 9.0)
9-11	**4%**	1%	**5.1**	**(2.1, 12.5)**	5%	2%	3.0	(0.9, 9.9)	0%	0%	1.0	(0.1, 7.6)
12-14	**6%**	1%	**8.3**	**(3.8, 18.2)**	1%	2%	0.8	(0.2, 4.4)	0%	0%	-	-
15-23	1%	1%	0.5	(0.1, 1.9)	1%	1%	0.1	(0.0, 1.2)	2%	1%	1.3	(0.4, 4.2)
>2 Years**	4%	7%	0.6	(0.3, 1.3)	8%	24%	0.2	(0.1, 0.5)	0%	0%	-	-
Missing***	5%	8%	0.6	(0.3, 1.3)	1%	9%	0.1	(0.0, 0.6)	0%	0%	-	-

*Did not have symptom during the 12 months before reference date. **Symptom started >2 years before reference date. ***Symptom onset date missing.
Note: '0-2'=0-2.99 months etc. cOR 95%CI are calculated relative to the number of women at risk of symptom for each time point. An average rate of 'new' control symptom frequency was used for each symptom. Half a control was added to each calculation for stability. Bolded values have 95% confidence intervals above 1.0.

Months Before Diagnosis	Questionnaire				Interview				GP Notes			
	Cases (n=249)	Controls (n=268)	cOR	95% CI	Cases (n=145)	Controls (n=125)	cOR	95% CI	Cases (n=223)	Controls (n=227)	cOR	95% CI
Increase in Abdominal Size												
None*	29%	88%	-	-	21%	73%	-	-	73%	99%	-	-
0-2	**36%**	0%	**272.8**	**(120.5, 617.8)**	**36%**	5%	**100.4**	**(42.7, 235.9)**	**20%**	0%	**193.9**	**(53.4, 704.6)**
3-5	**14%**	0%	**49.6**	**(21.5, 114.2)**	**20%**	2%	**20.8**	**(9.1, 47.8)**	**5%**	0%	**38.2**	**(9.6, 152.7)**
6-8	**6%**	0%	**17.0**	**(6.8, 42.3)**	**12%**	1%	**9.7**	**(4.1, 23.0)**	0%	0%	-	-
9-11	**3%**	1%	**8.4**	**(3.0, 23.6)**	**6%**	2%	**3.8**	**(1.4, 10.2)**	1%	0%	6.6	(1.0, 42.7)
12-14	**3%**	1%	**8.2**	**(2.9, 22.8)**	0%	0%	-	-	0%	0%	3.3	(0.3, 33.5)
15-23	1%	0%	1.0	(0.3, 3.8)	2%	0%	0.5	(0.1, 1.7)	0%	0%	1.1	(0.1, 11.1)
>2 Years**	4%	4%	1.1	(0.4, 2.6)	3%	12%	0.2	(0.1, 0.6)	0%	0%	-	-
Missing***	3%	6%	0.6	(0.2, 1.3)	1%	6%	0.1	(0.0, 1.0)	0%	0%	-	-
Back Pain												
None*	58%	67%	-	-	42%	46%	-	-	78%	82%	-	-
0-2	**10%**	1%	**12.6**	**(5.9, 26.5)**	**10%**	0%	**13.8**	**(2.7, 71.4)**	2%	2%	0.8	(0.3, 2.3)
3-5	**4%**	1%	**4.2**	**(1.7, 10.1)**	2%	1%	2.5	(0.4, 16.4)	4%	1%	1.7	(0.8, 3.6)
6-8	2%	1%	2.0	(0.7, 5.8)	4%	0%	4.8	(0.9, 26.3)	2%	3%	0.8	(0.3, 2.3)
9-11	2%	1%	2.4	(0.9, 6.5)	4%	2%	4.5	(0.8, 24.3)	4%	2%	1.8	(0.9, 3.8)
12-14	2%	1%	1.5	(0.5, 4.9)	1%	1%	0.7	(0.1, 8.2)	3%	1%	1.4	(0.6, 3.2)
15-23	2%	2%	0.5	(0.2, 1.6)	0%	2%	-	-	7%	9%	1.1	(0.6, 1.9)
>2 Years**	17%	17%	0.9	(0.6, 1.5)	32%	40%	0.7	(0.4, 1.1)	0%	0%	-	-
Missing***	2%	8%	0.3	(0.1, 0.7)	5%	9%	0.5	(0.2, 1.4)	0%	0%	-	-

*Did not have symptom during the 12 months before reference date. **Symptom started >2 years before reference date. ***Symptom onset date missing.
Note: '0-2'=0-2.99 months etc. cOR 95%CI are calculated relative to the number of women at risk of symptom for each time point. An average rate of 'new' control symptom frequency was used for each symptom. Half a control was added to each calculation for stability. Bolded values have 95% confidence intervals above 1.0.

Months Before Diagnosis	Questionnaire				Interview				GP Notes			
	Cases (n=249)	Controls (n=268)	cOR	95% CI	Cases (n=145)	Controls (n=125)	cOR	95% CI	Cases (n=223)	Controls (n=227)	cOR	95% CI
Constipation												
None*	65%	85%	-	-	61%	72%	-	-	75%	91%	-	-
0-2	**12%**	0%	**36.8**	**(16.1, 84.2)**	**10%**	2%	**25.0**	**(6.3, 100.1)**	**11%**	1%	**12.0**	**(6.5, 22.3)**
3-5	**6%**	1%	**15.0**	**(6.1, 36.9)**	**10%**	0%	**23.3**	**(5.9, 91.8)**	**6%**	0%	**5.9**	**(2.9, 12.0)**
6-8	**3%**	0%	**7.5**	**(2.7, 20.6)**	**3%**	1%	**6.8**	**(1.5, 31.8)**	2%	1%	1.6	(0.5, 4.7)
9-11	0%	1%	0.9	(0.1, 7.3)	**3%**	0%	**5.3**	**(1.1, 25.9)**	1%	2%	1.2	(0.4, 4.0)
12-14	**2%**	1%	**3.6**	**(1.1, 12.3)**	0%	0%	-	-	2%	1%	2.0	(0.7, 5.3)
15-23	2%	1%	1.2	(0.4, 4.0)	1%	0%	0.4	(0.0, 4.4)	2%	4%	0.7	(0.2, 1.7)
>2 Years**	5%	5%	0.9	(0.4, 2.0)	11%	20%	0.5	(0.2, 0.9)	0%	0%	-	-
Missing***	5%	6%	0.8	(0.4, 1.7)	1%	5%	0.1	(0.0, 1.2)	0%	0%	-	-
Diarrhoea												
None*	78%	90%	-	-	81%	82%	-	-	85%	90%	-	-
0-2	**8%**	1%	**20.0**	**(8.7, 46.3)**	**3%**	2%	**6.3**	**(1.6, 24.1)**	**5%**	0%	**4.2**	**(2.0, 8.9)**
3-5	**3%**	0%	**7.7**	**(2.9, 20.7)**	2%	0%	3.6	(0.8, 16.6)	2%	0%	1.5	(0.5, 4.3)
6-8	**2%**	0%	**5.6**	**(1.9, 16.2)**	**3%**	1%	**4.8**	**(1.2, 19.4)**	1%	2%	0.7	(0.2, 3.1)
9-11	0%	1%	0.9	(0.1, 7.3)	1%	0%	1.2	(0.1, 10.5)	1%	0%	1.1	(0.3, 3.7)
12-14	**2%**	1%	**3.7**	**(1.1, 12.1)**	1%	2%	1.2	(0.1, 10.4)	2%	2%	1.5	(0.5, 4.3)
15-23	0%	0%	0.3	(0.0, 2.4)	1%	0%	0.4	(0.0, 3.5)	4%	4%	1.1	(0.5, 2.4)
>2 Years**	2%	2%	1.1	(0.3, 3.4)	8%	10%	0.7	(0.3, 1.7)	0%	0%	-	-
Missing***	4%	4%	1.1	(0.5, 2.5)	1%	5%	0.1	(0.0, 1.2)	0%	0%	-	-

*Did not have symptom during the 12 months before reference date. **Symptom started >2 years before reference date. ***Symptom onset date missing.
Note: '0-2'=0-2.99 months etc. cOR 95%CI are calculated relative to the number of women at risk of symptom for each time point. An average rate of 'new' control symptom frequency was used for each symptom. Half a control was added to each calculation for stability. Bolded values have 95% confidence intervals above 1.0.

Months Before Diagnosis	Questionnaire				Interview				GP Notes			
	Cases (n=249)	Controls (n=268)	cOR	95% CI	Cases (n=145)	Controls (n=125)	cOR	95% CI	Cases (n=223)	Controls (n=227)	cOR	95% CI
Change Bowel Habit												
None*	-	-	-	-	89%	99%	-	-	82%	94%	-	-
0-2	-	-	-	-	**5%**	0%	**35.7**	**(6.0, 212.1)**	**8%**	1%	**11.6**	**(5.6, 23.8)**
3-5	-	-	-	-	**3%**	0%	**19.4**	**(2.9, 128.5)**	**4%**	0%	**4.7**	**(2.0, 11.4)**
6-8	-	-	-	-	**1%**	0%	**9.4**	**(1.1, 79.2)**	1%	0%	1.7	(0.5, 6.0)
9-11	-	-	-	-	0%	0%	0.0	-	2%	1%	2.3	(0.7, 7.0)
12-14	-	-	-	-	1%	1%	4.7	(0.4, 58.9)	1%	2%	1.7	(0.5, 5.9)
15-23	-	-	-	-	0%	0%	-	-	2%	2%	0.7	(0.2, 2.3)
>2 Years**	-	-	-	-	1%	0%	-	-	0%	0%	-	-
Missing***	-	-	-	-	0%	0%	-	-	0%	0%	-	-
Indigestion												
None*	59%	75%	-	-	46%	46%	-	-	78%	84%	-	-
0-2	**12%**	1%	**20.3**	**(9.5, 43.3)**	**14%**	2%	**11.2**	**(3.2, 38.8)**	**4%**	1%	**2.3**	**(1.1, 4.9)**
3-5	**6%**	0%	**9.4**	**(4.2, 21.4)**	**9%**	2%	**5.7**	**(1.6, 20.1)**	4%	2%	2.2	(1.0, 4.8)
6-8	2%	1%	2.2	(0.7, 7.1)	4%	0%	2.4	(0.6, 9.2)	2%	2%	1.2	(0.5, 3.1)
9-11	2%	2%	2.7	(0.9, 8.0)	4%	2%	2.3	(0.6, 8.7)	1%	1%	0.7	(0.2, 2.4)
12-14	**3%**	1%	**4.3**	**(1.7, 10.9)**	0%	3%	-	-	3%	3%	1.5	(0.6, 3.5)
15-23	0%	0%	0.2	(0.0, 1.3)	1%	1%	0.1	(0.0, 1.1)	8%	7%	1.4	(0.8, 2.5)
>2 Years**	10%	9%	1.0	(0.6, 1.9)	19%	34%	0.4	(0.2, 0.7)	0%	1%	0.5	(0.0, 5.6)
Missing***	6%	10%	0.5	(0.3, 1.0)	3%	10%	0.3	(0.1, 0.9)	0%	0%	-	-

*Did not have symptom during the 12 months before reference date. **Symptom started >2 years before reference date. ***Symptom onset date missing.
Note: '0-2'=0-2.99 months etc. OR 95%CI are calculated relative to the number of women at risk of symptom for each time point. An average rate of 'new' control symptom frequency was used for each symptom. Half a control was added to each calculation for stability. Bolded values have 95% confidence intervals above 1.0.

Months Before Diagnosis	Questionnaire				Interview				GP Notes			
	Cases (n=249)	Controls (n=268)	cOR	95% CI	Cases (n=145)	Controls (n=125)	cOR	95% CI	Cases (n=223)	Controls (n=227)	cOR	95% CI
Nausea/Vomiting												
None*	67%	93%	-	-	70%	88%	-	-	75%	89%	-	-
0-2	18%	3%	**52.7**	**(25.8, 107.8)**	**16%**	4%	**26.6**	**(10.7, 66.2)**	**11%**	2%	**9.1**	**(5.1, 16.2)**
3-5	6%	0%	**12.6**	**(5.5, 29.0)**	**8%**	2%	**10.5**	**(3.9, 28.4)**	3%	1%	2.2	(1.0, 5.3)
6-8	2%	0%	**4.2**	**(1.4, 12.6)**	1%	0%	0.9	(0.1, 7.3)	0%	1%	0.3	(0.0, 2.3)
9-11	0%	0%	0.0	-	1%	0%	1.8	(0.4, 8.7)	2%	1%	1.6	(0.6, 4.2)
12-14	1%	1%	1.7	(0.4, 7.7)	0%	0%	-	-	2%	2%	1.3	(0.4, 3.6)
15-23	1%	0%	0.6	(0.1, 2.6)	1%	0%	0.3	(0.0, 2.4)	6%	5%	1.5	(0.8, 2.8)
>2 Years**	2%	1%	1.7	(0.5, 6.0)	3%	3%	0.8	(0.2, 3.4)	0%	0%	-	-
Missing***	3%	1%	2.2	(0.7, 7.4)	1%	3%	0.4	(0.1, 2.3)	0%	0%	-	-
Lump in Abdomen												
None*	73%	99%	-	-	81%	98%	-	-	86%	100%	-	-
0-2	**17%**	0%	**998.2**	**(59.8, 16661.9)**	**12%**	2%	**59.8**	**(15.6, 229.9)**	**11%**	0%	**474.3**	**(28.7, 7847.4)**
3-5	**4%**	0%	**188.0**	**(10.7, 3301.1)**	**2%**	0%	**8.7**	**(1.6, 47.2)**	**2%**	0%	**83.9**	**(4.6, 1544.0)**
6-8	**2%**	0%	**90.0**	**(4.8, 1688.0)**	**2%**	0%	**8.5**	**(1.6, 46.1)**	0%	0%	16.4	(0.5, 490.9)
9-11	0%	0%	-	-	1%	0%	5.6	(0.9, 36.0)	0%	0%	-	-
12-14	0%	0%	-	-	1%	0%	2.8	(0.3, 28.1)	0%	0%	-	-
15-23	0%	0%	5.9	(0.2, 178.7)	0%	0%	-	-	0%	0%	5.4	(0.2, 162.7)
>2 Years**	0%	1%	0.5	(0.0, 6.1)	0%	0%	-	-	0%	0%	-	-
Missing***	2%	0%	-	-	0%	0%	-	-	0%	0%	-	-

*Did not have symptom during the 12 months before reference date. **Symptom started >2 years before reference date. ***Symptom onset date missing.
Note: '0-2'=0-2.99 months etc. cOR 95%CI are calculated relative to the number of women at risk of symptom for each time point. An average rate of 'new' control symptom frequency was used for each symptom. Half a control was added to each calculation for stability. Bolded values have 95% confidence intervals above 1.0.

Months Before Diagnosis	Questionnaire				Interview				GP Notes			
	Cases (n=249)	Controls (n=268)	cOR	95% CI	Cases (n=145)	Controls (n=125)	cOR	95% CI	Cases (n=223)	Controls (n=227)	cOR	95% CI
Urinary Frequency/Urgency												
None*	56%	83%	-	-	50%	63%	-	-	77%	91%	-	-
0-2	**21%**	1%	**37.9**	**(20.1, 71.4)**	**14%**	2%	**12.9**	**(5.0, 33.2)**	**7%**	1%	**7.1**	**(3.5, 14.2)**
3-5	**9%**	0%	**11.8**	**(5.9, 23.7)**	**12%**	3%	**9.3**	**(3.7, 23.7)**	**4%**	3%	**3.9**	**(1.8, 8.8)**
6-8	**4%**	1%	**4.4**	**(1.9, 10.3)**	3%	0%	2.2	(0.7, 7.2)	1%	0%	0.8	(0.2, 3.7)
9-11	2%	2%	1.9	(0.6, 5.8)	**6%**	6%	**3.9**	**(1.4, 10.7)**	**4%**	1%	**4.2**	**(2.0, 9.2)**
12-14	1%	1%	0.9	(0.2, 4.1)	1%	0%	0.4	(0.0, 3.3)	**3%**	2%	**2.9**	**(1.2, 6.8)**
15-23	0%	0%	0.2	(0.0, 1.2)	0%	0%	-	-	4%	2%	1.1	(0.5, 2.4)
>2 Years**	5%	5%	1.0	(0.4, 2.1)	12%	24%	0.4	(0.2, 0.8)	0%	0%	-	-
Missing***	3%	6%	0.5	(0.2, 1.2)	1%	2%	0.6	(0.1, 3.5)	0%	0%	-	-
Urinary Other												
None*	-	-	-		73%	96%	-	-	63%	84%	-	-
0-2	-	-	-	-	**12%**	2%	**46.4**	**(14.1, 152.5)**	**10%**	2%	**7.4**	**(4.2, 13.2)**
3-5	-	-	-	-	**5%**	0%	**15.5**	**(4.1, 57.8)**	**8%**	2%	**5.1**	**(2.8, 9.4)**
6-8	-	-	-	-	**3%**	0%	**8.4**	**(1.9, 36.3)**	2%	1%	1.3	(0.5, 3.4)
9-11	-	-	-	-	1%	0%	4.1	(0.7, 23.7)	**5%**	3%	**2.9**	**(1.4, 5.8)**
12-14	-	-	-	-	1%	0%	2.0	(0.2, 18.9)	3%	4%	1.5	(0.6, 3.6)
15-23	-	-	-	-	0%	0%	-	-	8%	3%	1.5	(0.8, 2.7)
>2 Years**	-	-	-	-	3%	1%	4.5	(0.5, 38.7)	0%	1%	0.3	(0.0, 3.3)
Missing***	-	-	-	-	1%	1%	1.7	(0.2, 19.4)	0%	0%	-	-

*Did not have symptom during the 12 months before reference date. **Symptom started >2 years before reference date. ***Symptom onset date missing.
Note: '0-2'=0-2.99 months etc. cOR 95%CI are calculated relative to the number of women at risk of symptom for each time point. An average rate of 'new' control symptom frequency was used for each symptom. Half a control was added to each calculation for stability. Bolded values have 95% confidence intervals above 1.0.

Months Before Diagnosis	Questionnaire				Interview				GP Notes			
	Cases (n=249)	Controls (n=268)	cOR	95% CI	Cases (n=145)	Controls (n=125)	cOR	95% CI	Cases (n=223)	Controls (n=227)	cOR	95% CI
Fatigue												
None*	46%	81%	-	-	38%	50%	-	-	77%	84%	-	-
0-2	**20%**	1%	**35.5**	**(19.5, 64.7)**	**12%**	6%	**7.2**	**(2.8, 18.1)**	**8%**	2%	**4.4**	**(2.4, 8.0)**
3-5	**9%**	1%	**11.5**	**(5.9, 22.2)**	**19%**	2%	**8.5**	**(3.6, 19.8)**	**4%**	1%	**2.3**	**(1.1, 4.7)**
6-8	**6%**	0%	**7.1**	**(3.5, 14.4)**	**9%**	1%	**3.1**	**(1.3, 7.7)**	1%	2%	0.7	(0.2, 2.2)
9-11	2%	2%	2.1	(0.8, 5.7)	**11%**	9%	**3.5**	**(1.5, 8.2)**	2%	1%	0.9	(0.3, 2.5)
12-14	**6%**	1%	**5.7**	**(2.8, 11.8)**	3%	0%	0.8	(0.2, 2.6)	1%	2%	0.4	(0.1, 1.9)
15-23	4%	1%	1.2	(0.5, 2.6)	2%	1%	0.2	(0.1, 0.7)	7%	7%	1.1	(0.6, 2.1)
>2 Years**	3%	4%	0.8	(0.3, 1.9)	6%	24%	0.2	(0.1, 0.4)	0%	0%	-	-
Missing***	4%	7%	0.6	(0.3, 1.3)	0%	9%	-	-	0%	0%	-	-
Weight Loss												
None*	66%	97%	-	-	73%	97%	-	-	85%	97%	-	-
0-2	**16%**	0%	**77.1**	**(32.6, 182.2)**	**8%**	1%	**24.3**	**(8.0, 73.9)**	**9%**	0%	**26.0**	**(11.1, 61.2)**
3-5	**8%**	0%	**32.7**	**(13.2, 80.8)**	**12%**	1%	**31.0**	**(10.7, 90.0)**	**2%**	0%	**5.6**	**(1.8, 17.6)**
6-8	**4%**	0%	**15.6**	**(5.8, 41.9)**	**3%**	1%	**8.0**	**(2.2, 29.1)**	0%	0%	1.1	(0.1, 8.9)
9-11	1%	0%	4.1	(1.0, 16.2)	1%	1%	3.1	(0.6, 16.6)	0%	0%	-	-
12-14	1%	1%	2.7	(0.5, 13.2)	1%	0%	1.5	(0.2, 13.5)	0%	0%	1.1	(0.1, 8.9)
15-23	0%	0%	0.4	(0.1, 3.7)	0%	0%	-	-	2%	2%	1.5	(0.4, 5.0)
>2 Years**	0%	0%	1.1	(0.1, 17.5)	1%	0%	-	-	1%	0%	-	-
Missing***	2%	1%	2.7	(0.5, 14.2)	1%	0%	-	-	0%	0%	-	-

*Did not have symptom during the 12 months before reference date. **Symptom started >2 years before reference date. ***Symptom onset date missing.
Note: '0-2'=0-2.99 months etc. cOR 95%CI are calculated relative to the number of women at risk of symptom for each time point. An average rate of 'new' control symptom frequency was used for each symptom. Half a control was added to each calculation for stability. Bolded values have 95% confidence intervals above 1.0.

Months Before Diagnosis	Questionnaire				Interview				GP Notes			
	Cases (n=249)	Controls (n=268)	cOR	95% CI	Cases (n=145)	Controls (n=125)	cOR	95% CI	Cases (n=223)	Controls (n=227)	cOR	95% CI
Loss of Appetite												
None*	63%	98%	-	-	54%	85%	-	-	80%	98%	-	-
0-2	**18%**	0%	244.6	**(66.5, 898.9)**	**17%**	2%	**16.4**	**(8.2, 32.9)**	**12%**	1%	**48.6**	**(19.0, 124.1)**
3-5	**9%**	0%	90.7	**(23.9, 343.9)**	**15%**	1%	**11.1**	**(5.6, 22.3)**	**3%**	0%	**9.4**	**(2.9, 30.3)**
6-8	**4%**	0%	41.0	**(10.2, 165.3)**	**6%**	0%	**3.4**	**(1.4, 8.2)**	**1%**	0%	**4.6**	**(1.1, 18.9)**
9-11	**2%**	0%	14.2	**(2.9, 70.3)**	**6%**	10%	**3.7**	**(1.6, 8.5)**	**1%**	0%	**4.5**	**(1.1, 18.7)**
12-14	0%	0%	3.5	(0.3, 36)	1%	0%	0.4	(0.1, 3.0)	0%	0%	-	-
15-23	0%	0%	1.2	(0.1, 12)	0%	1%	-	-	2%	1%	2.5	(0.7, 8.4)
>2 Years**	2%	0%	4.4	(0.5, 39.3)	1%	2%	0.6	(0.1, 3.5)	0%	0%	-	-
Missing***	1%	1%	1.1	(0.2, 7.7)	0%	0%	-	-	0%	0%	-	-
Vaginal Discharge												
None*	-	-	-	-	-	-	-		89%	96%	-	-
0-2	-	-	-	-	-	-	-	-	1%	0%	3.5	(0.9, 13.6)
3-5	-	-	-	-	-	-	-	-	**4%**	0%	**10.4**	**(3.9, 28.0)**
6-8	-	-	-	-	-	-	-	-	0%	0%	1.1	(0.1, 9.0)
9-11	-	-	-	-	-	-	-	-	0%	0%	1.1	(0.1, 9.0)
12-14	-	-	-	-	-	-	-	-	1%	0%	3.3	(0.9, 12.8)
15-23	-	-	-	-	-	-	-	-	3%	2%	2.2	(0.7, 6.5)
>2 Years**	-	-	-	-	-	-	-	-	0%	0%	1.0	(0.1, 16.4)
Missing***		-	-	-	-	-	-	-	0%	0%	-	-

*Did not have symptom during the 12 months before reference date. **Symptom started >2 years before reference date. ***Symptom onset date missing.
Note: '0-2'=0-2.99 months etc. cOR 95%CI are calculated relative to the number of women at risk of symptom for each time point. An average rate of 'new' control symptom frequency was used for each symptom. Half a control was added to each calculation for stability. Bolded values have 95% confidence intervals above 1.0.
'Loss of appetite' is 'Loss of appetite or feeling full quickly' on interview

Months Before Diagnosis	Questionnaire				Interview				GP Notes			
	Cases (n=249)	Controls (n=268)	cOR	95% CI	Cases (n=145)	Controls (n=125)	cOR	95% CI	Cases (n=223)	Controls (n=227)	cOR	95% CI
Irregular Vaginal Bleeding												
None*	84%	97%	-	-	94%	99%	-	-	95%	100%	-	-
0-2	**6%**	0%	**21.2**	**(8.7, 51.4)**	0%	0%	0.0	-	0%	0%	5.7	(0.4, 71.7)
3-5	**3%**	0%	**9.9**	**(3.6, 27.1)**	**1%**	0%	9.6	**(1.1, 81.0)**	0%	0%	5.7	(0.4, 71.4)
6-8	**2%**	0%	**7.2**	**(2.4, 21.2)**	**1%**	0%	9.5	**(1.1, 79.8)**	**1%**	0%	**11.3**	**(1.3, 94.1)**
9-11	1%	0%	3.5	(0.9, 13.5)	1%	0%	4.7	(0.4, 59.3)	**1%**	0%	**16.7**	**(2.3, 119.5)**
12-14	1%	1%	2.3	(0.5, 11.1)	1%	1%	4.7	(0.4, 58.9)	0%	0%	0.0	-
15-23	0%	0%	0.4	(0.0, 3.1)	0%	0%	-	-	1%	0%	5.5	(0.8, 39.3)
>2 Years**	1%	0%	-	-	1%	0%	-	-	0%	0%	-	-
Missing***	0%	0%	-	-	0%	0%	-	-	0%	0%	-	-
Postmenopausal Bleeding												
None*	-	-	-	-	89%	98%	-	-	88%	97%	-	-
0-2	-	-	-	-	**6%**	1%	**19.4**	**(5.5, 68.3)**	**8%**	0%	**20.3**	**(8.4, 49.0)**
3-5	-	-	-	-	**3%**	1%	**8.1**	**(1.9, 34.5)**	**3%**	1%	**6.6**	**(2.2, 19.7)**
6-8	-	-	-	-	0%	0%	-	-	0%	0%	1.1	(0.1, 8.7)
9-11	-	-	-	-	0%	1%	-	-	1%	1%	2.2	(0.4, 10.3)
12-14	-	-	-	-	1%	0%	2.0	(0.2, 18.4)	0%	0%	-	-
15-23	-	-	-	-	0%	0%	-	-	0%	0%	-	-
>2 Years**	-	-	-	-	1%	0%	-	-	0%	0%	-	-
Missing***	-	-	-	-	1%	0%	-	-	0%	0%	-	-

*Did not have symptom during the 12 months before reference date. **Symptom started >2 years before reference date. ***Symptom onset date missing.
Note: '0-2'=0-2.99 months etc. cOR 95%CI are calculated relative to the number of women at risk of symptom for each time point. An average rate of 'new' control symptom frequency was used for each symptom. Half a control was added to each calculation for stability. Bolded values have 95% confidence intervals above 1.0.

4.1.6 Symptom Sensitivity & Specificity 3-14 Months Before Reference Date

One of the key objectives of this study was to assess the potential for using symptoms as a tool for earlier diagnosis by way of minimising the time between symptom onset to (appropriate) referral. In this scenario, it was appropriate to consider symptoms in terms of sensitivity (proportion of cases with the symptom) and specificity (proportion of controls without the symptom). In order to quantify the proportion of cases that would potentially be referred earlier; cumulative symptom incidence over one year was examined along with the time between symptom onset and referral. For a more conservative (and clinically relevant) approach, symptoms that started within 3 months (2.99 months) of diagnosis were excluded. As detailed in the Methods section, this allowed calculation of symptom incidence with at least 1 month of actual lead time. In addition, Table 4-7 demonstrated that symptom incidence for cases and controls were similar 15 months before reference date. Thus, symptoms that were present during 3-14 (i.e. 3 to 14.99) months before reference date were used to examine cumulative symptom incidence over one year. Again, symptoms that started more than 2 years prior and symptoms with missing onset dates were excluded. Hence, in Table 4-8 to Table 4-10 sensitivity was calculated as the annual incidence of each symptom in cases, and specificity was equivalent to (1-annual incidence) of each symptom in controls. It is important to note that symptom specificity will not necessarily be representative of women in the general population since control women were apparently healthy volunteers.

Crucially, sensitivity was greatly reduced by restricting symptoms to 3-14 months. Symptom sensitivity was comparable for questionnaire and interview. The most sensitive symptoms according to the questionnaire were abdominal bloating (30%), increase in abdominal size (27%) and pelvic/abdominal pain or discomfort (26%). Importantly, 64% of cases reported at least one relevant symptom (i.e. 'possibly related' to ovarian cancer, see Appendix VI) 3-14 months before diagnosis. For interview, sensitivity was highest for fatigue (41%), increased abdominal size (38%), and pelvic/abdominal pain or discomfort (34%). As many as 81% of cases reported at least one ovarian cancer related symptom. Again, symptom sensitivity on GP notes was lower with the greatest proportions observed for pelvic/abdominal pain or discomfort (26%), urinary other (17%) and urinary frequency or urgency (13%). 70% of cases had least one relevant symptom recorded in their notes over the year examined.

Specificity was relatively high on the questionnaire. According to the questionnaire, specificity for at least one relevant symptom was 80%. The corresponding specificity on interview was only 37%. Lump in abdomen and loss in appetite both had a specificity of 100%, and the lowest specificity for any symptom was 95%. Other symptoms such as vaginal discharge and change in bowel habit also had a specificity of 100%, however these were not specifically asked about on the questionnaire or interview. Specificity was worse on interview for bloating (94%), fatigue (89%), pelvic/abdominal pain (91%), urinary frequency/urgency (91%) and indigestion (93%). Otherwise, values were similar to those in the questionnaire. GP note symptom specificity was 90% for urinary other and pelvic/abdominal pain; however specificity was higher for the remaining symptoms.

All of the odds ratios for each of the 14 symptoms that appeared on the questionnaire were significant. Case-control differences were also significant on interview for all individual symptoms except for diarrhoea, PMB and vaginal discharge. Ten of the 18 (individual) symptoms were recorded significantly more in cases than controls according to GP notes

Table 4-8 Questionnaire Symptom Incidence 3-14 Months Before Reference Date with Crude Odds Ratios (OR) & 95% Confidence Intervals

Symptom	Cases n=249	Controls n=268	OR (95%CI)	P value
Lump In Abdomen	**15 (6%)**	0 (0%)	**- (4.5, -)**	**<0.0001**
Loss Of Appetite	**38 (15%)**	1 (0%)	**48.1 (6.5, 353.0)**	**0.0001**
Increase in Abdominal Size	**67 (27%)**	6 (2%)	**16.1 (6.8, 37.8)**	**<0.0001**
Pelvic/Abdo. Pain/Discomfort	**65 (26%)**	8 (3%)	**11.5 (5.4, 24.5)**	**<0.0001**
Abdominal Bloating	**74 (30%)**	11 (4%)	**9.9 (5.1, 19.2)**	**<0.0001**
Weight Loss	**37 (15%)**	5 (2%)	**9.2 (3.5, 23.8)**	**<0.0001**
Nausea Or Vomiting	**21 (8%)**	3 (1%)	**8.1 (2.4, 27.6)**	**0.0008**
Constipation	**28 (11%)**	6 (2%)	**5.5 (2.2, 13.6)**	**0.0002**
Fatigue	**57 (23%)**	14 (5%)	**5.4 (2.9, 10.0)**	**<0.0001**
Diarrhoea	**19 (8%)**	6 (2%)	**3.6 (1.4, 9.2)**	**0.0071**
IVB	**19 (8%)**	6 (2%)	**3.6 (1.4, 9.2)**	**0.0071**
Urinary Frequency/Urgency	**37 (15%)**	13 (5%)	**3.4 (1.8, 6.6)**	**0.0002**
Indigestion	**33 (13%)**	12 (4%)	**3.3 (1.6, 6.5)**	**0.0007**
Back Pain	**25 (10%)**	11 (4%)	**2.6 (1.3, 5.4)**	**0.0102**
Change In Bowel Habit*	1 (0%)	0 (0%)	-	
Urinary Other*	2 (1%)	0 (0%)	-	
PMB*	0 (0%)	0 (0%)	-	
Vaginal Discharge*	0 (0%)	0 (0%)	-	
Other Possibly Related	***12 (5%)***	*3 (1%)*	***4.5 (1.2, 16.0)***	***0.0215***
None	*89 (36%)*	***214 (80%)***	***0.1 (0.1, 0.2)***	***<0.0001***

*Symptoms were not specifically listed on the questionnaire.
Symptoms are in order of highest to lowest odds ratios for individual symptoms. Bolded values are significant at p<0.05 level.
PMB is postmenopausal bleeding. IVB is irregular vaginal bleeding (includes all IVB except for PMB).

Table 4-9 Interview Symptom Incidence 3-14 Months Before Reference Date with Crude Odds Ratios (OR) & 95% Confidence Intervals

Symptom	Cases n=145	Controls n=125	OR (95%CI)	P value
Urinary Other*	**14 (10%)**	0 (0%)	**- (3.4, -)**	**0.0001**
Lump In Abdomen	**9 (6%)**	0 (0%)	**- (2.1, -)**	**0.0041**
Loss Of Appetite**	**33 (23%)**	1 (1%)	**36.5 (4.9, 271.5)**	**0.0004**
Constipation	**24 (17%)**	1 (1%)	**24.6 (3.3, 184.7)**	**0.0018**
Increase in Abdominal Size	**55 (38%)**	6 (5%)	**12.1 (5.0, 29.4)**	**<0.0001**
Weight Loss	**25 (17%)**	3 (2%)	**8.5 (2.5, 28.8)**	**0.0006**
Nausea Or Vomiting	**14 (10%)**	2 (2%)	**6.6 (1.5, 29.5)**	**0.014**
Change In Bowel Habit	7 (5%)	1 (1%)	6.3 (0.8, 51.8)	0.0875
Abdominal Bloating	**43 (30%)**	8 (6%)	**6.2 (2.8, 13.7)**	**<0.0001**
Fatigue	**60 (41%)**	14 (11%)	**5.6 (2.9, 10.7)**	**<0.0001**
Pelvic/Abdo. Pain/Discomfort	**50 (34%)**	11 (9%)	**5.5 (2.7, 11.1)**	**<0.0001**
IVB	6 (4%)	1 (1%)	5.4 (0.6, 45.1)	0.1228
Back Pain	**16 (11%)**	4 (3%)	**3.8 (1.2, 11.5)**	**0.0211**
Urinary Frequency/Urgency	**33 (23%)**	11 (9%)	**3.1 (1.5, 6.3)**	**0.0027**
Indigestion	**25 (17%)**	9 (7%)	**2.7 (1.2, 6.0)**	**0.016**
Diarrhoea	9 (6%)	3 (2%)	2.7 (0.7, 10.2)	0.1444
Vaginal Discharge*	6 (4%)	2 (2%)	2.7 (0.5, 13.4)	0.2371
PMB	5 (3%)	2 (2%)	2.2 (0.4, 11.5)	0.3522
Other Possibly Related	***24 (17%)***	*7 (6%)*	***3.3 (1.4, 8.1)***	***0.0071***
None	*27 (19%)*	***79 (63%)***	***0.1 (0.1-0.2)***	***<0.0001***

*Symptoms not specifically asked about on interview
**Loss of appetite or feeling full quickly'
Symptoms are in order of highest to lowest odds ratios for individual symptoms. Bolded values are significant at p<0.05 level.
PMB is postmenopausal bleeding. IVB is irregular vaginal bleeding (includes all IVB except for PMB).

Table 4-10 GP Notes Symptom Incidence 3-14 Months Before Reference Date with Crude Odds Ratios (OR) & 95% Confidence Intervals

Symptom	Cases n=223	Controls n=227	OR (95%CI)	P value
Lump In Abdomen	**6 (3%)**	0 (0%)	**- (1.6, -)**	**0.0143**
Increase in Abdominal Size	**14 (6%)**	1 (0%)	**15.1 (2.0, 116.1)**	**0.0089**
Loss Of Appetite	**12 (5%)**	1 (0%)	**12.9 (1.7, 99.7)**	**0.0146**
IVB	6 (3%)	1 (0%)	6.2 (0.7, 52.3)	0.091
Vaginal Discharge	**14 (6%)**	3 (1%)	**5.0 (1.4, 17.7)**	**0.0124**
Pelvic/Abdo. Pain/Discomfort	**58 (26%)**	22 (10%)	**3.3 (1.9, 5.6)**	**<0.0001**
Weight Loss	7 (3%)	2 (1%)	3.6 (0.7, 17.7)	0.1091
Abdominal Bloating	**15 (7%)**	5 (2%)	**3.2 (1.1, 9.0)**	**0.0267**
Constipation	**26 (12%)**	10 (4%)	**2.9 (1.3, 6.1)**	**0.0063**
Urinary Frequency/Urgency	**28 (13%)**	13 (6%)	**2.4 (1.2, 4.7)**	**0.014**
Change In Bowel Habit	18 (8%)	8 (4%)	2.4 (1.0, 5.6)	0.0442
Urinary Other	**39 (17%)**	22 (10%)	**2.0 (1.1, 3.5)**	**0.0171**
Back Pain	28 (13%)	16 (7%)	1.9 (1.0, 3.6)	0.0521
PMB	9 (4%)	5 (2%)	1.9 (0.6, 5.7)	0.2698
Nausea Or Vomiting	17 (8%)	11 (5%)	1.6 (0.7, 3.5)	0.2264
Fatigue	19 (9%)	14 (6%)	1.4 (0.7, 2.9)	0.3404
Indigestion	23 (10%)	17 (7%)	1.4 (0.7, 2.7)	0.2943
Diarrhoea	13 (6%)	11 (5%)	1.2 (0.5, 2.8)	0.6428
Other Possibly Related	*41 (18%)*	*37 (16%)*	*1.2 (0.7, 1.9)*	*0.5591*
None	*68 (30%)*	***121 (53%)***	***0.4 (0.3, 0.6)***	***<0.0001***

Symptoms are in order of highest to lowest odds ratios for individual symptoms. Bolded values are significant at p<0.05 level.
PMB is postmenopausal bleeding. IVB is irregular vaginal bleeding (includes all IVB except for PMB).

4.1.7 GP Visits

Since GP visit data for cases were only extracted up to a 'cut-off point', visits observed at 0 to 2 (i.e. 0 to 2.99) months prediagnosis exclude any visits made between 'cut-off' date and diagnosis (median 1.3 months, IQR 0.7-1.9). Examination of GP visits showed that cases began to consult for problems more frequently 6-8 months prior to diagnosis (see Figure 4-10). This effect became even more prominent when visits were restricted to those in which a possibly related symptom was recorded (see Figure 4-11). For cases, the number of routine visits decreased as problem visits increased (N.B. if problems were mentioned at routine visits, a problem visit code was assigned). About 40% of cases and controls had a problem GP visit in any quarter, until 6-8 months before diagnosis (at which point this proportion increased for cases).

The median number of problem visits over the 2 year period was 7 (IQR 3-11) for cases and 4 (IQR 2-8) for controls (p<0.0001). Routine visits were much more

comparable with a case median of 3 (IQR 1-7) and control median of 4 (IQR 2-6) (p=0.0783).

Figure 4-10 Proportion of Women with Problem versus Routine Visits Prior to Reference Date Per Prediagnostic Quarter

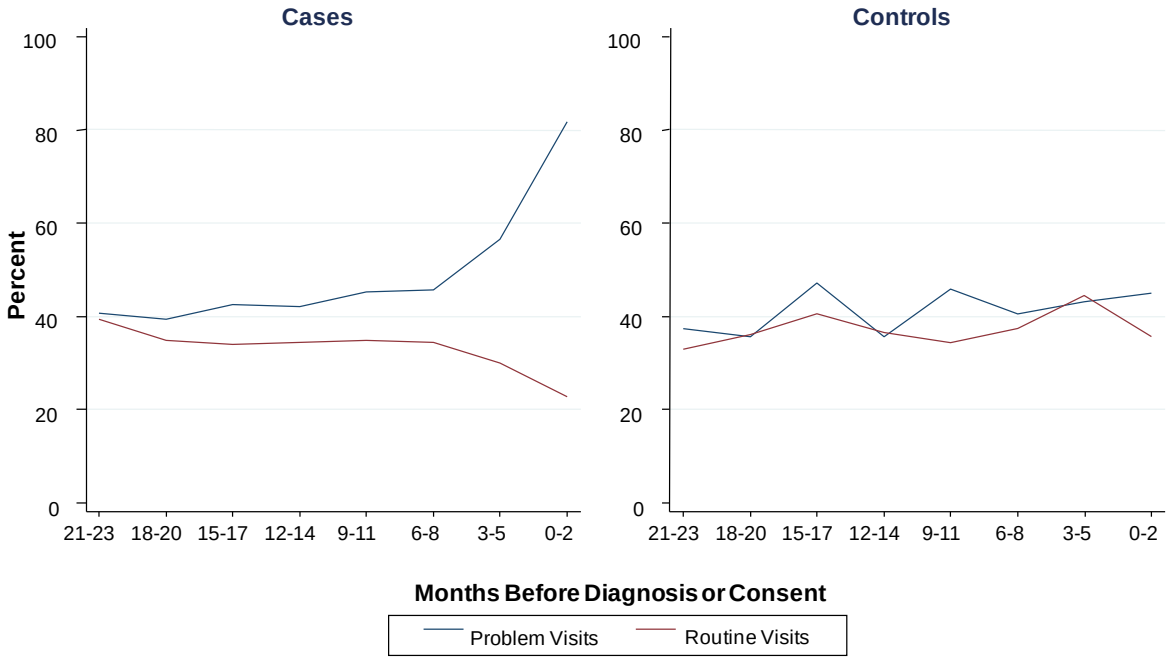
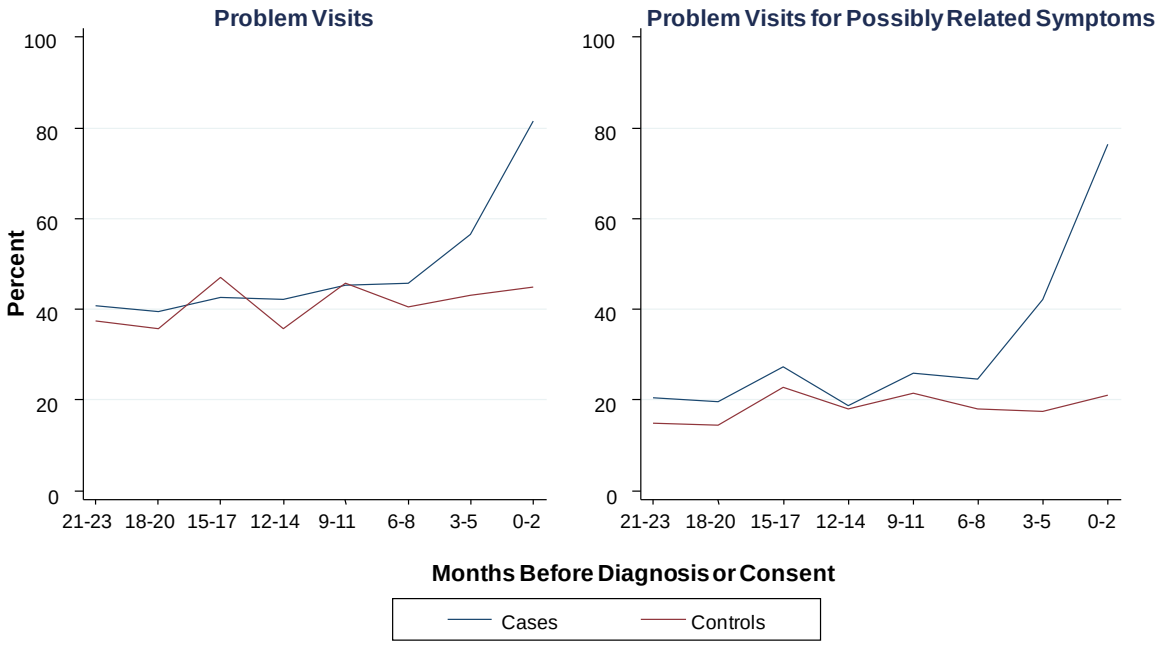


Figure 4-11 Proportion of Women with Problem GP Visits Prior to Diagnosis

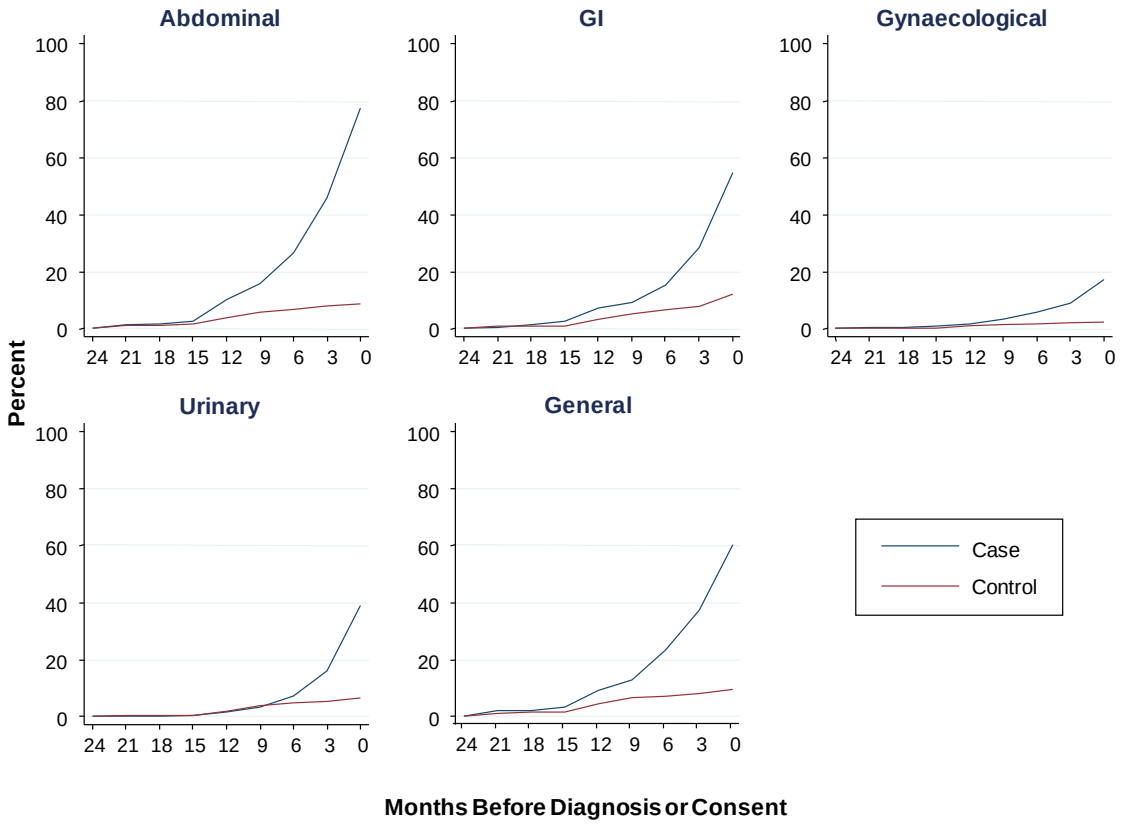


4.1.8 Grouped Symptoms

Symptoms possibly related to ovarian cancer were grouped into five categories including abdominal, GI, urinary, gynaecological and general (see 3.3.10 and Appendix VI for details). These provided a more general impression of the timing and extent to which each of the named body systems were affected. As with the individual symptoms, longstanding symptoms were excluded and graphs show cumulative incidence of 'new' symptoms in the two years prior to reference date. Grouped control symptom frequency was relatively constant over the two year period examined for all sources (see Figure 4-12, Figure 4-13, Figure 4-14). In contrast, cases exhibited sharp increases in grouped symptom incidence for all sources. This was most pronounced for abdominal, GI and general symptoms and least for gynaecological symptoms. Increases were evident earliest on questionnaire data (15 months prior to diagnosis), then interview (12 months before diagnosis) and closer to diagnosis on GP notes (6 months).

Self-reported data shared a similar pattern of symptom onset with abdominal, GI, general symptoms diverging first, followed by gynaecological and urinary symptoms. Abdominal symptoms were the most commonly reported, followed by GI and general symptoms. Conversely, symptom groups in GP note data all seemed to separate around the same time point of 6 months prior to reference date with the exception of urinary symptoms. These were recorded in slight excess for cases from 18 months before diagnosis, however there was a much sharper increase around 6 months. In corroboration with self-reported data, abdominal and GI symptoms were amongst the most prevalent. However, unlike self-reported data, sensitivity and specificity for general symptoms was lower in GP notes.

Figure 4-12 Questionnaire - Cumulative Grouped Symptoms Per Quarter Before Reference Date



Abdominal: Pelvic/abdominal pain or discomfort, bloating, increased abdominal size
Gastrointestinal (GI): indigestion, diarrhoea, constipation, change in bowel habit, GI other, nausea or vomiting
Gynaecological: Postmenopausal bleeding, irregular vaginal bleeding, pain with intercourse, post-coital bleeding, vaginal discharge
Urinary: Frequency, urgency, incontinence, dysuria, retention, change in smell or colour, haematuria
General: Loss of appetite, weight loss, fatigue

Figure 4-13 Interview - Cumulative Grouped Symptoms Per Quarter Before Reference Date

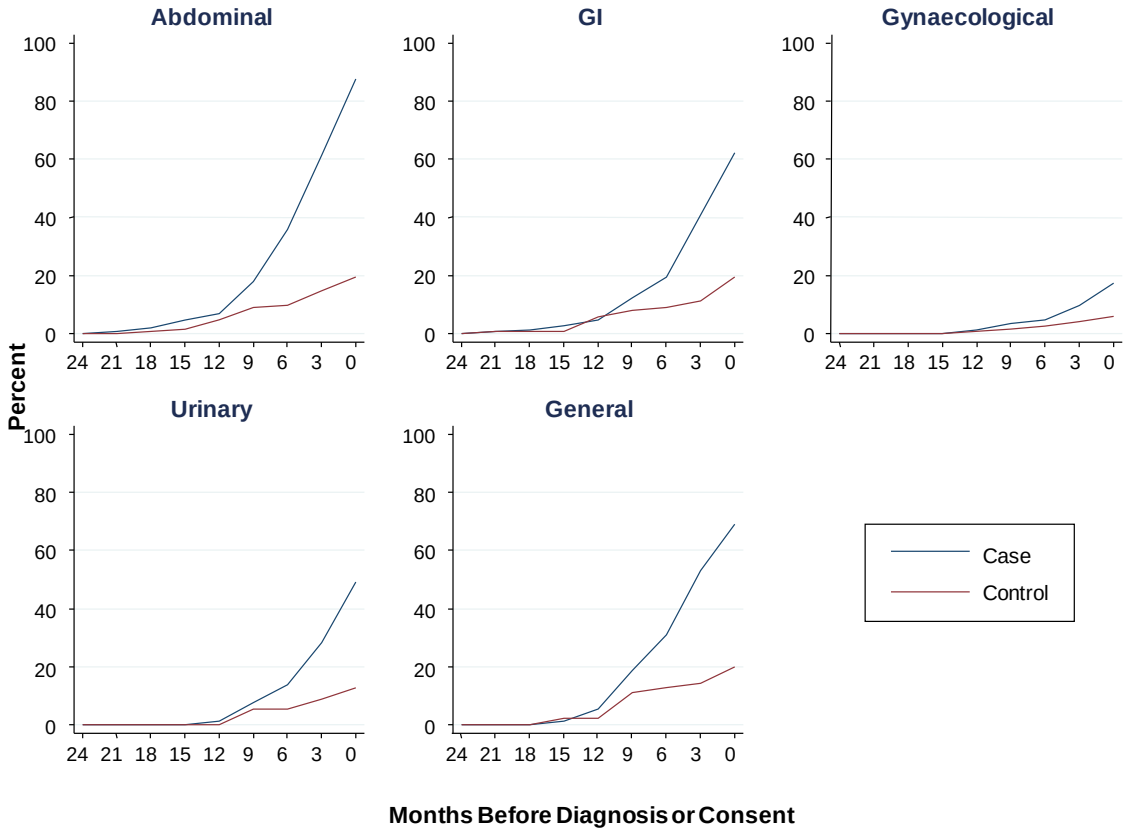
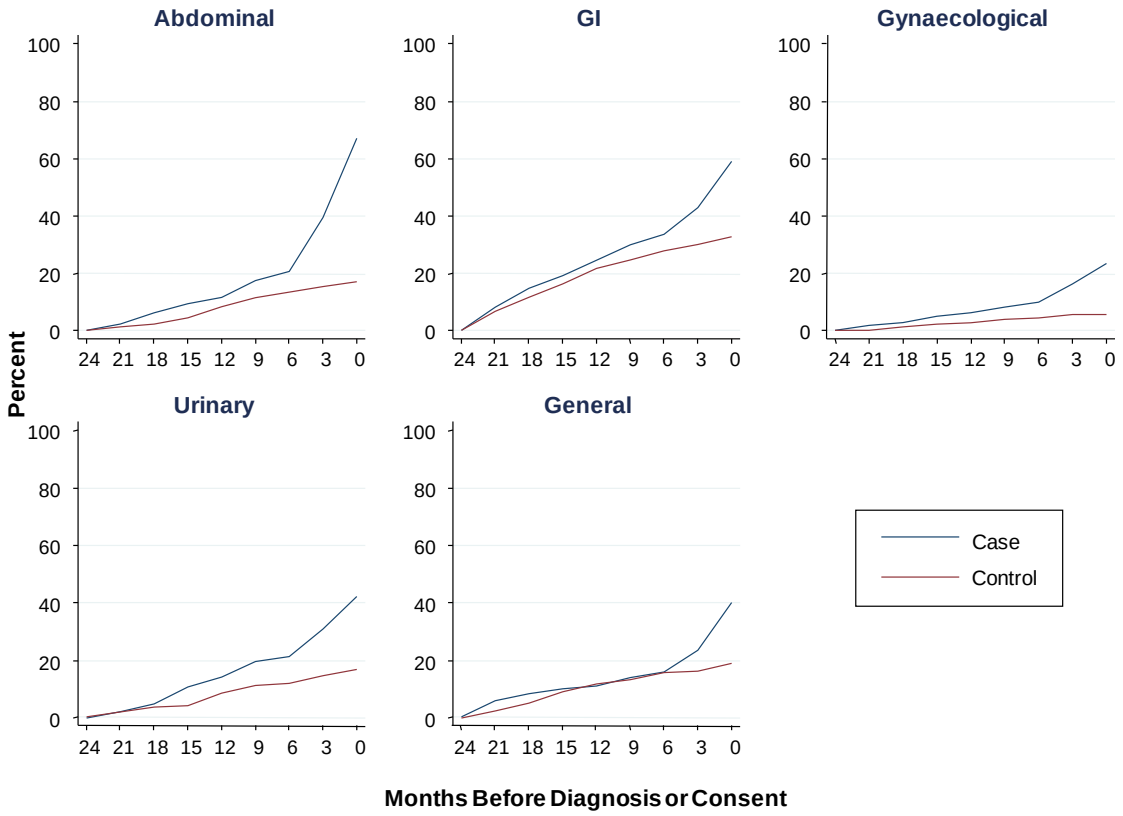


Figure 4-14 GP Notes - Cumulative Grouped Symptoms Per Quarter Before Reference Date



4.1.9 Symptom Severity & Frequency

Severity and frequency for symptoms present over one year (i.e. 3-14.99 months before reference date) were analysed for case-control differences. Longstanding symptoms and symptoms with missing onset were excluded. The maximum severity and frequency was taken for each woman for each data source. Rating of symptoms was intended to represent severity (mild, moderate, severe) and frequency (1-4, 5-15, 16-31 days per month) for when the symptom first started. A test for linear trend in severity was significant for questionnaire (p=0.0047) and interview (p=0.0493), showing that women reporting moderate or severe symptoms were more likely to be cases than women reporting mild symptoms (see Table 4-11).

As shown in Table 4-11, women with more frequent symptoms were more likely to be cases according to both questionnaire (p=0.0152) and interview (p=0.0024)

Table 4-11 Crude Odds Ratios with 95% Confidence Intervals for Symptom Severity & Frequency for 3-14 Months Before Reference Date

	Questionnaire		Interview	
Severity	**OR**	**95%CI**	**OR**	**95%CI**
Mild	1.0	-	1.0	-
Moderate	1.1	0.5, 2.6	2.4	1.0, 5.8
Severe	**3.3**	**1.2, 8.6**	2.4	1.0, 5.8
Frequency	**OR**	**95%CI**	**OR**	**95%CI**
1-4 days	1.0	-	1.0	-
5-15 days	1.3	0.3, 4.7	0.6	0.1, 3.9
16-31 days	2.9	0.9, 9.0	4.0	0.9, 17.5

Bolded values have 95% confidence intervals above 1.0

Table 4-12 and Table 4-13 show symptom severity and frequency for individual symptoms for symptoms reported at anytime (excluding longstanding symptoms). In general, for cases symptom severity was rated in higher proportions as moderate on questionnaire, and mild on interview. Frequency ratings were typically 16-31 days per month for most symptoms reported by cases on both sources. Numbers of individual control symptoms with severity and frequency data were too small to study trends. However, overall the proportion of symptoms that were rated as mild was generally larger. At least 80% of cases who reported bloating, increased abdominal size and loss of appetite had the respective symptom with a frequency of 16-31 days per month.

Table 4-12 Symptom Severity on Self-Reported Data* (Anytime)

	Questionnaire		Interview	
	Cases	Controls	Cases	Controls
Pelvic/Abdo Pain	n=131	n=20	n=82	n=17
Mild	32/124 (26%)	2/15 (13%)	31/67 (46%)	4/16 (25%)
Moderate	57/124 (46%)	11/15 (73%)	22/67 (33%)	8/16 (50%)
Severe	35/124 (28%)	2/15 (13%)	14/67 (21%)	4/16 (25%)
Missing	7/131 (5%)	5/20 (25%)	15/82 (18%)	1/17 (6%)
Bloating	n=171	n=35	n=81	n=22
Mild	26/156 (17%)	10/24 (42%)	40/71 (56%)	16/20 (80%)
Moderate	73/156 (47%)	11/24 (46%)	23/71 (32%)	3/20 (15%)
Severe	57/156 (37%)	3/24 (13%)	8/71 (11%)	2/20 (10%)
Missing	15/171 (9%)	11/35 (31%)	10/81 (12%)	2/22 (9%)
↑**Abdo Size**	n=167	n=22	n=111	n=19
Mild	15/151 (10%)	5/11 (45%)	66/99 (67%)	10/14 (71%)
Moderate	54/151 (36%)	4/11 (36%)	23/99 (23%)	3/14 (21%)
Severe	82/151 (54%)	2/11 (18%)	10/99 (10%)	1/14 (7%)
Missing	16/167 (10%)	11/22 (50%)	12/111 (11%)	5/19 (26%)
Back Pain	n=59	n=42	n=37	n=18
Mild	22/53 (42%)	5/33 (15%)	19/32 (59%)	11/17(65%)
Moderate	21/53 (40%)	21/33 (64%)	10/32 (31%)	2/17 (12%)
Severe	10/53 (19%)	7/33 (21%)	3/32 (9%)	4/17 (24%)
Missing	6/59 (10%)	9/42 (21%)	5/37 (14%)	1/18 (6%)
Constipation	n=74	n=25	n=40	n=11
Mild	9/63 (14%)	2/14 (14%)	10/33 (30%)	6/8 (75%)
Moderate	37/63 (59%)	9/14 (64%)	8/33 (24%)	2/8 (25%)
Severe	17/63 (27%)	3/14 (21%)	15/33 (45%)	0/8 (0%)
Missing	11/74 (15%)	11/25 (44%)	7/40 (18%)	3/11 (27%)
Diarrhoea	n=49	n=20	n=18	n=8
Mild	9/38 (24%)	5/15 (42%)	5/14 (36%)	6/7 (86%)
Moderate	19/38 (50%)	5/15 (42%)	4/14 (29%)	0/7 (0%)
Severe	10/38 (26%)	2/15 (17%)	5/14 (36%)	1/7 (14%)
Missing	11/49(22%)	8/20 (40%)	4/18 (22%)	1/8 (13%)
Indigestion	n=77	n=43	n=52	n=27
Mild	21/64 (33%)	14/32 (44%)	24/48 (50%)	18/24 (75%)
Moderate	29/64 (45%)	14 (44%)	12 (25%)	2 (10%)
Severe	14/64 (22%)	4 (13%)	12 (25%)	4 (17%)
Missing	13/77 (17%)	11/43 (26%)	4/52 (8%)	3/27 (11%)
Loss of Appetite	n=87	n=4	n=58	n=5
Mild	16/79 (20%)	1/3 (33%)	25/48 (52%)	2/5 (40%)
Moderate	38/79 (48%)	0/3 (0%)	16/48 (33%)	1/5 (20%)
Severe	25/79 (32%)	2/3 (67%)	7/48 (15%)	2/5 (40%)
Missing	8/87 (9%)	1/4 (25%)	10/58 (11%)	0/5 (0%)

*Severity rated as mild, moderate or severe

Lump in abdomen, PMB, IVB, change in bowel habit and vaginal discharge excluded from calculations.

Boxes shaded in grey show symptoms that have noticeably different severity for cases versus controls.

Table 4-12 continued

	Questionnaire		Interview	
	Cases	Controls	Cases	Controls
Nausea/Vomiting	n=77	n=14	n=40	n=11
Mild	19/68 (28%)	5/12 (42%)	10/22 (45%)	4/6 (67%)
Moderate	31/68 (46%)	4/12 (33%)	8/22 (36%)	0/6 (0%)
Severe	18/68 (26%)	3/12 (25%)	4/22 (18%)	2/6 (33%)
Missing	9/77 (12%)	2/14 (14%)	18/40 (45%)	5/11 (25%)
Urinary Freq/Urge	n=96	n=32	n=55	n=18
Mild	20/87 (23%)	5/30 (17%)	27/48 (56%)	10/16 (63%)
Moderate	31/87 (36%)	4/30 (13%)	12/48 (25%)	4/16 (25%)
Severe	18/87 (21%)	3/30 (10%)	9/48 (19%)	2/16 (13%)
Missing	9/96 (10%)	2/32 (6%)	7/55 (13%)	2/18 (11%)
Weight Loss	n=83	n=9	n=38	n=4
Mild	12/73 (16%)	0/6 (0%)	17/29 (59%)	0/1 (0%)
Moderate	31/73 (42%)	6/6 (100%)	7/29 (24%)	1/1 (100%)
Severe	30/73 (41%)	0/6 (0%)	5/29 (17%)	0/1 (0%)
Missing	10/83(12%)	3/9 (33%)	9/38 (24%)	3/4 (75%)
Fatigue	n=125	n=39	n=80	n=33
Mild	22/108 (20%)	9/26 (35%)	39/73 (53%)	23/29 (79%)
Moderate	64/108 (59%)	11/26 (42%)	26/73 (36%)	5/29 (17%)
Severe	22/108 (20%)	6/26 (23%)	8/73 (11%)	1/29 (3%)
Missing	17/125 (14%)	13/39 (33%)	7/80 (9%)	4/33 (12%)

*Severity rated as mild, moderate or severe

Lump in abdomen, PMB, IVB, change in bowel habit and vaginal discharge excluded from calculations

Boxes shaded in grey show symptoms that have noticeably different severity for cases versus controls.

Table 4-13 Symptom Frequency (days per month) on Self-Reported Data* (Anytime)

	Questionnaire		Interview	
	Cases	Controls	Cases	Controls
Pelvic/Abdo Pain	n=131	n=21	n=82	n=17
1-4	17/121 (14%)	5/15 (33%)	8/60 (13%)	3/13 (23%)
5-15	30/121 (25%)	7/15 (47%)	7/60 (12%)	4/13 (31%)
16-31	74/121 (61%)	3/15 (20%)	45/60 (75%)	6/13 (46%)
Missing	10/131 (8%)	6/21 (29%)	22/82 (27%)	4/17 (24%)
Bloating	n=171	n=35	n=81	n=22
1-4	4/147 (3%)	8/26 (31%)	0/65 (0%)	2/9 (22%)
5-15	16/147 (11%)	13/26 (50%)	7/65 (11%)	4/9 (44%)
16-31	127/147 (86%)	5/26 (19%)	58/65 (89%)	3/9 (33%)
Missing	24/171 (14%)	9/35 (26%)	16/81 (20%)	13/22 (59%)
↑**Abdo Size**	n=167	n=22	n=111	n=19
1-4	0/146 (0%)	2/14 (14%)	0/98 (0%)	1/10 (10%)
5-15	11/146 (8%)	9/14 (64%)	3/98 (3%)	0/10 (0%)
16-31	135/146 (92%)	3/14 (21%)	95/98 (97%)	10/10 (90%)
Missing	21/167 (13%)	8/22 (36%)	13/111 (12%)	9/19 (47%)
Back Pain	n=59	n=42	n=37	n=18
1-4	10/53 (19%)	4/30 (13%)	2/21 (10%)	0/7 (0%)
5-15	26/53 (49%)	10/30 (33%)	5/21 (24%)	2/7 (29%)
16-31	17/53 (32%)	16/30 (53%)	14/21 (67%)	5/7 (71%)
Missing	6/59 (10%)	12/42 (29%)	16/37 (43%)	11/18 (61%)
Constipation	n=74	n=25	n=40	n=11
1-4	10/61 (16%)	5/15 (33%)	2/30 (7%)	0/2 (0%)
5-15	24/61 (39%)	5/15 (33%)	3/30 (10%)	1/2 (50%)
16-31	27(44%)	5/15 (33%)	25/30 (83%)	1/2 (50%)
Missing	13/74 (18%)	10/25 (40%)	10/40 (25%)	9/11 (82%)
Diarrhoea	n=49	n=20	n=18	n=8
1-4	12/43 (28%)	7/14 (50%)	2/10 (20%)	1/3 (33%)
5-15	16/43 (37%)	4/14 (29%)	1/10 (10%)	1/3 (33%)
16-31	15 (35%)	3/14 (21%)	7/10 (70%)	1/3 (33%)
Missing	6/49(12%)	6/20 (30%)	8/18 (44%)	5/8 (63%)
Indigestion	n=77	n=43	n=52	n=27
1-4	16/65 (25%)	13/15 (33%)	5/40 (13%)	2/11 (18%)
5-15	22/65 (34%)	7/15 (33%)	10/40 (25%)	5/11 (45%)
16-31	27/65 (42%)	8/15 (33%)	25/40 (63%)	4/11 (36%)
Missing	12/77 (16%)	15/43 (35%)	12/52 (23%)	16/27 (59%)
Loss of Appetite	n=87	n=4	n=58	n=5
1-4	7/81 (9%)	1/3 (33%)	0/45 (0%)	0/4 (0%)
5-15	9/81 (11%)	0/3 (0%)	3/45 (7%)	0/4 (0%)
16-31	65/81 (80%)	2/3 (67%)	42/45 (93%)	4/4 (100%)
Missing	6/87(9%)	1/4 (25%)	13/58 (22%)	1/5 (20%)

*Frequency is rated as days per month

Lump in abdomen, PMB, IVB, change in bowel habit and vaginal discharge excluded from calculations

Boxes shaded in grey show symptoms that have noticeably different frequency for cases versus controls.

Table 4-13 continued

	Questionnaire		Interview	
	Cases	Controls	Cases	Controls
Nausea/Vomiting	n=77	n=14	n=40	n=11
1-4	20/68 (29%)	6/11 (55%)	2/18 (11%)	0/3 (0%)
5-15	17/68 (25%)	2/11 (18%)	6/18 (33%)	2/3 (67%)
16-31	31/68 (46%)	3/11 (27%)	10/18 (56%)	1/3 (33%)
Missing	9/77 (12%)	3/14 (21%)	22/40 (55%)	8/11 (73%)
Urinary Freq/Urge	n=96	n=32	n=55	n=18
1-4	5/88 (6%)	6/22 (27%)	0/50 (0%)	0/13 (0%)
5-15	18/88 (20%)	4/22 (18%)	5/50 (10%)	1/13 (8%)
16-31	65/88 (74%)	12/22 (55%)	45/50 (90%)	12/13 (92%)
Missing	8/96 (8%)	10/32 (31%)	5/55 (9%)	5/18 (28%)
Fatigue	n=125	n=39	n=80	n=33
1-4	5/105 (5%)	2/26 (8%)	0/73 (0%)	3/22 (14%)
5-15	18/105 (17%)	9/26 (35%)	7/73 (10%)	2/22 (9%)
16-31	82/105 (78%)	15/26 (58%)	66/73 (90%)	17/22 (77%)
Missing	20/125 (16%)	13/39 (33%)	7/80 (9%)	11/33 (33%)

*Frequency is rated as days per month
Lump in abdomen, PMB, IVB, change in bowel habit and vaginal discharge excluded from calculations.
Boxes shaded in grey show symptoms that have noticeably different frequency for cases versus controls.

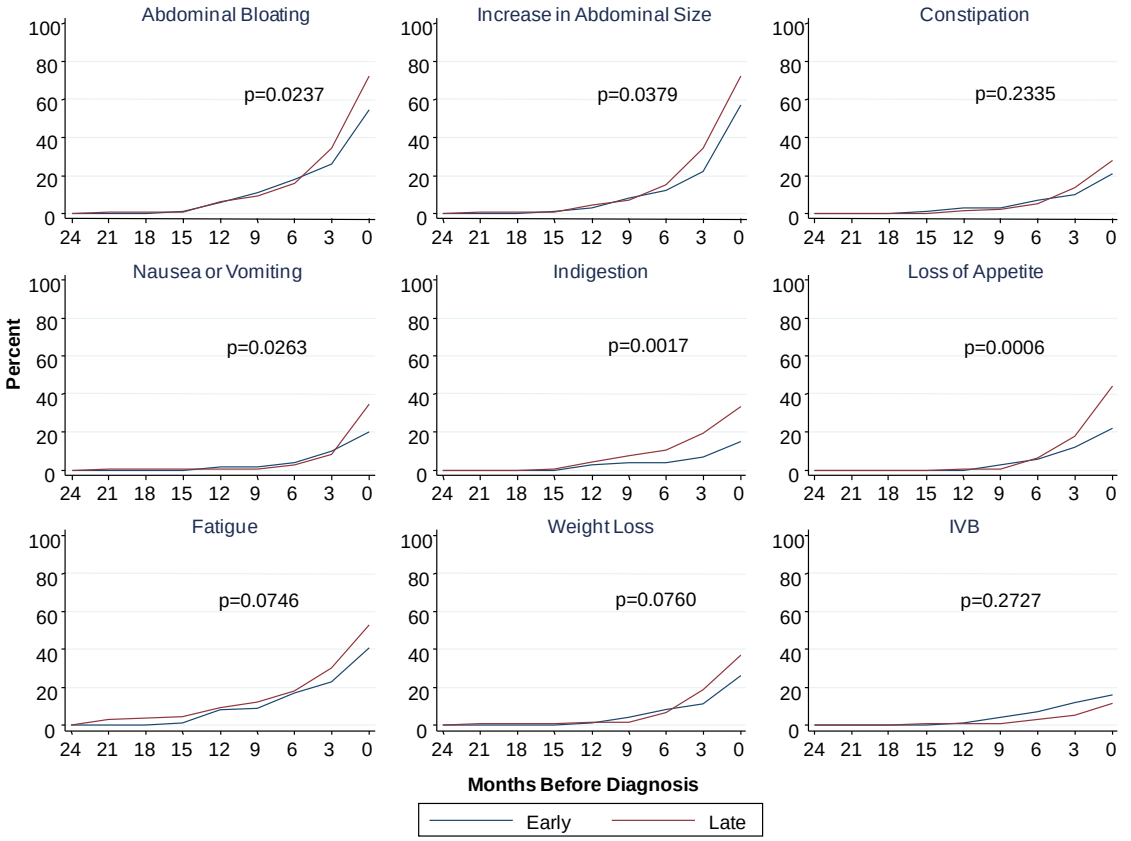
Changes in symptom severity were recorded on interview; however there was a considerable amount of missing responses. Data were available for 65% (469/723) of symptoms in cases and 42% (81/192) of symptoms in controls. The options were 'improved', 'same', or 'worsened'. Of the informative symptoms, a linear test for trend showed that women were significantly more likely to be a case if they reported worsening symptoms ($p<0.0001$).

4.1.10 Early versus Advanced Stage

Most women with early stage disease had at least one symptom that started at anytime before diagnosis (100% [GP], 96% [I], 95% [Q]). Similarly, women with late stage disease were also highly symptomatic (100% [both GP & I], 98% [Q]). Cumulative symptom frequency was calculated for cases according to stage. Overall, the symptoms reported by women with early versus late stage disease were comparable. However, there was a tendency for women with late stage disease to report higher symptom incidence close to diagnosis across all data sources. These included indigestion, fatigue, loss of appetite, weight loss, constipation and nausea or vomiting. This reached statistical significance at diagnosis for indigestion, loss of appetite, abdominal bloating, increased abdominal size, nausea/vomiting and weight loss (see Figure 4-15 to Figure 4-17). This remained significant at 3 months prior to diagnosis for indigestion (GP notes; $p=0.0299$ and questionnaire; $p=0.0067$), and weight loss (interview; $p=0.0068$). Increased abdominal size was more common in GP notes for cases at 3 months before ($p=0.0190$), but not at diagnosis ('cut-off').

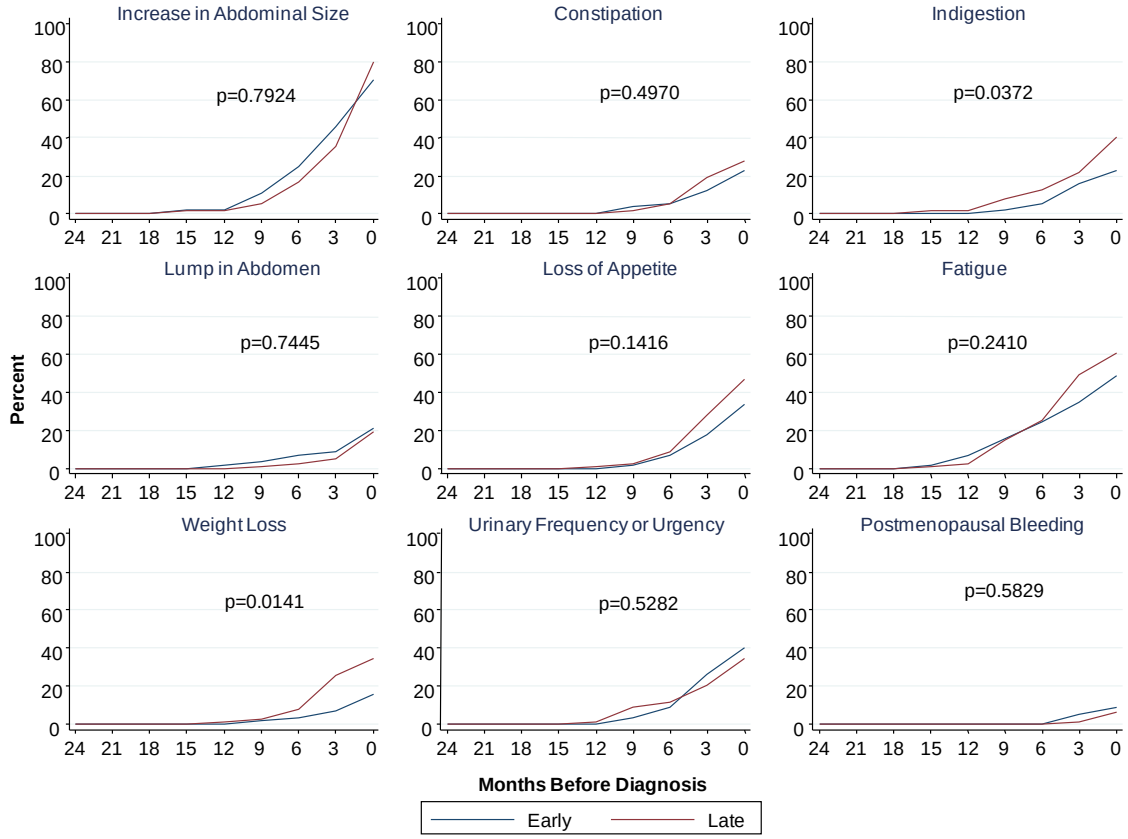
In contrast, urinary symptoms, lump in abdomen and gynaecological symptoms were more prominent in early stage disease however these differences were not statistically significant at any time point (see Figure 4-15, Figure 4-16, Figure 4-17). The figures shown contain the symptoms which showed the most pronounced differences between early and late disease for each data source.

Figure 4-15 Cumulative Symptom Frequency for Early versus Late Disease - Questionnaire



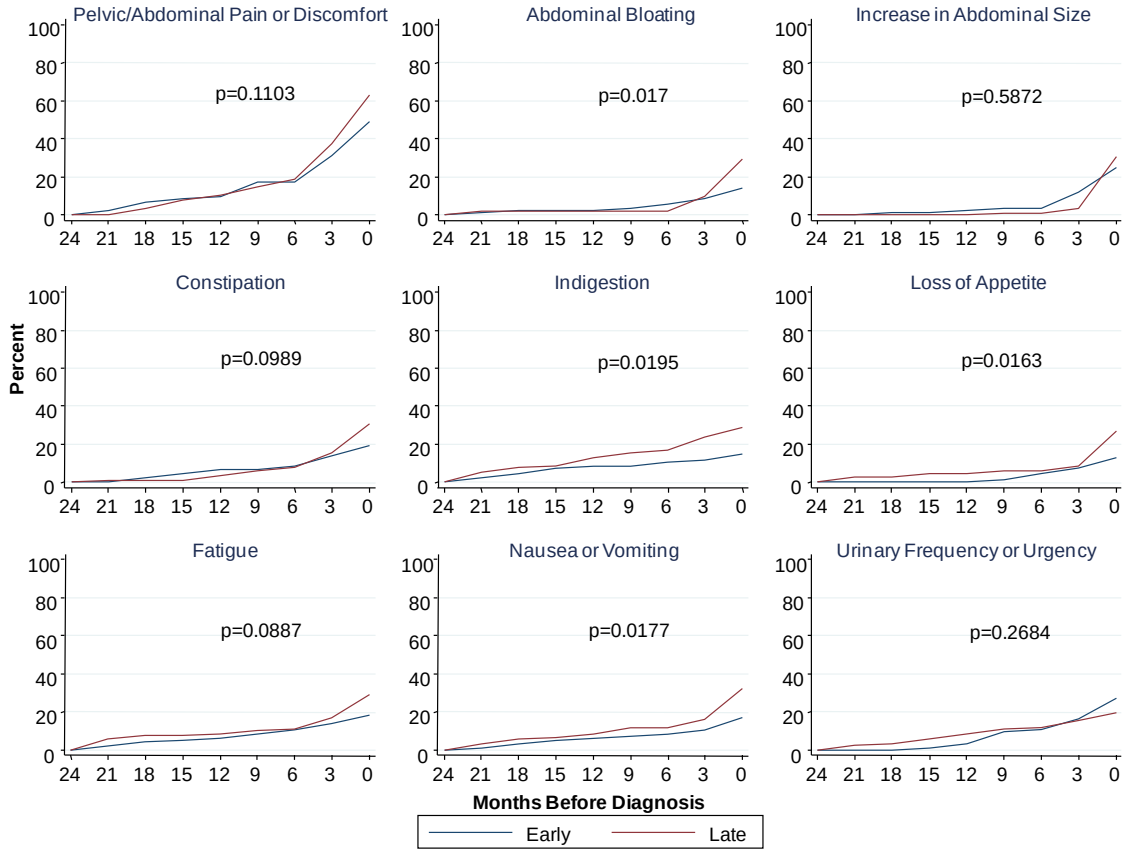
Note: Graphs show p-values using log rank test comparing rates in the two years before diagnosis.

Figure 4-16 Cumulative Symptom Frequency for Early versus Late Disease - Interview



Note: Graphs show p-values using log rank test comparing rates in the two years before diagnosis.
'Loss of appetite' is 'Loss of appetite or feeling full quickly' on interview

Figure 4-17 Cumulative Symptom Frequency for Early versus Late Disease - GP Notes



Note: Graphs show p-values using log rank test comparing rates in the two years before diagnosis.

The median number of symptoms for early versus late disease is shown in Table 4-14. Women with late stage disease reported more symptoms than their early stage counterparts for the three data sources.

Table 4-14 Early vs. Late Disease: Median Number of Symptoms (up to diagnosis) per Data Source

	Early	Late	P value
Questionnaire	n=100	n=132	
Median (IQR)	5 (2-7)	6 (3-9)	0.0036
Interview	n=57	n=79	
Median (IQR)	5 (3-7)	6 (4-9)	0.0425
GP Notes	n=93	n=117	
Median (IQR)	3 (2-6)	5 (3-7)	0.0006

4.1.11 Number of Symptoms

The number of symptoms possibly related to ovarian cancer was counted for each source (excluding longstanding symptoms). The median number of (new) symptoms reported was higher for cases in all three data sources. Symptom reporting was

greatest on questionnaire with a median of 6 (IQR 3-8) and 0 (IQR 0-2) for cases and controls, respectively. The next highest was for interview, then GP notes.

Table 4-15 Number of Symptoms

	Cases	Controls	P value
Questionnaire	n=249	n=268	
Median (IQR)	6 (3-8)	0 (0-2)	<0.0001
Telephone Interview	n=145	n=125	
Median (IQR)	5 (4-8)	2 (0-3)	<0.0001
GP Notes	n=223	n=227	
Median (IQR)	4 (2-7)	1 (0-3)	<0.0001

Excludes 'other unrelated' symptoms

4.1.12 Route of Diagnosis

Route of diagnosis was examined by looking at the first GP referral made for symptoms. Although accident and emergency (A&E) is not a true referral, this was included since these women had either:

1) Never been given a referral by the GP (despite consulting)

2) Never consulted the GP, or

3) Consulted the GP but symptoms became so severe that a referral appointment could not be executed within the referral timeframe.

The vast majority of women were diagnosed via symptoms (89%), the remaining 11% were identified by incidental findings (5.5%) or screening studies (5.9%) (see Table 4-16). Rapid referral to gynaecological-oncology (G/O) was the most common initial referral made, yet still only comprised less than half of first referrals. Only 49% of cases were initially referred to a gynaecology department. Twenty percent (44/225) of cases were sent to gastrointestinal clinics, 18 of the 44 (41%) were via rapid referral. One fifth of cases presented as emergencies before a GP referral was made or successfully executed. For some women, referral to A&E or rapid referral G/O may have been on recommendation by the radiographer after suspicious lesion(s) found on ultrasound, which had been ordered by their GP.

Interestingly, all incidentally diagnosed women reported potentially related symptoms on interview and GP notes in the year before diagnosis and only one woman reported having no symptoms on questionnaire. In contrast, 12 of the 15 women who were screen-detected were asymptomatic according to GP notes. However, all reported at least one symptom on interview and questionnaire, many of which were present months before diagnosis.

Table 4-16 First Referral

Referral Type	Number (%)
Rapid Referral Gynaecological Oncology	91 (40)
General Surgery	16 (7)
GI Physician*	10 (4)
General Gynaecology	20 (9)
Urgent to Other Specialties	24 (11)
Routine to Other Specialities**	5 (2)
A&E GP-initiated	28 (12)
A&E Patient-initiated	20 (9)
Unknown	11 (5)
Total Referrals	**225 (89)**
N/A***	29 (11)
TOTAL	**254 (100)**

*Private GI physician referrals (n=4)
**One woman had a routine referral to clinical oncology as she was still undergoing follow up for colon cancer
***Screen-detected (n=15), incidental diagnosis (n=14)

4.1.13 Delays in Diagnosis

Delays in diagnosis were examined using all 'possibly related' symptoms that started within 15 months prior to diagnosis (symptoms that started >15 months before diagnosis were omitted). Women who were screen-detected (n=15) or incidentally diagnosed (n=14) were excluded from the analyses. Of the remaining women, 9% (19/220), 1% (1/124) and 3% (5/197) had no eligible symptoms on questionnaire, interview and GP notes, respectively. These women were classified as having no delays in diagnosis (not zero delay). Note that delays could not be derived for women who had missing dates for symptom onset or first GP visit for all eligible symptoms, hence these women were also classified as having no delay.

Table 4-17 below shows that the longest median patient delay was reported on interview at 3.0 months (IQR 0.0-7.0). Conversely, patient delay was almost non-existent according to GP notes. Provider delays were similar for self-reported data, however the largest interval was 4.4 months (IQR 2.0-10.0) in GP notes. Total delay ranged between a median of 5.0 months (GP notes) and 8.4 months (interview). The overall total delay (combined data) was 9.8 months (IQR 5.0-12.3) (based on 222 women with available data, 3 women had no eligible symptoms in any of the 3 data sources).

Table 4-17 Delays in Diagnosis (Months)

	Questionnaire n=220	Interview n=124	GP Notes n=197
Women with No Delay	19 (9%)	1 (1%)	5 (3%)
Patient Delay	n=161 (73%)	n=107 (86%)	n=192 (97%)
Median	1.0	3.0	0.0
IQR	0.0-5.0	0.0-7.0	0.0-0.3
Mean	3.1	3.8	0.5
Range	0.0-13.0	0.0-12.0	0.0-9.0
Provider Delay	n=161 (73%)	n=107 (86%)	n=192 (97%)
Median	2.1	2.7	4.4
IQR	1.4-4.8	1.4-4.8	2.0-10.0
Mean	3.6	3.6	5.9
Range	0.0-13.2	0.1-12.5	0.1-14.9
Total Delay	n=201 (91%)	n=123 (99%)	n=192 (97%)
Median	5.6	8.4	5.0
IQR	2.8-11.4	4.0-10.4	2.6-10.4
Mean	6.6	7.3	6.5
Range	0.2-14.8	0.1-14.9	0.5-14.9

Table 4-18 shows that GPs took a median of 1.7 months to make a referral to any department. The median interval between first GP visit (for symptom) to appropriate referral was 2.7 months.

Of the women who self-presented to A&E with GP note data available, 3 had no prior GP visits for (eligible) symptoms (3/18 with GP notes). Similarly, of the women who were sent by their GP to A&E, only 3 had no (eligible) symptoms recorded in their notes (3/26 with GP notes).

The median time from first GP referral to date of diagnosis was similar for rapid referral G/O versus all other referrals (1.4 months, IQR 1.1-2.0 and 1.8 months IQR 1.0-3.2, respectively). When referrals to A&E were excluded, median time from first GP referral to diagnosis was slightly longer at 2.3 months (IQR 1.4-5.0).

Table 4-18 Time to Referral (Months)

	Any Referral n=195	Appropriate Referral* n=195
Provider Delay	n=191 (98%)	n=189 (97%)
Median	1.7	2.7
IQR	0.2-7.2	0.5-8.2

*Appropriate referral was considered to be any gynaecological referral. If this date was missing, 'cut-off' date was used
Women with unknown referral pathways were excluded from total delay except for one woman in whom first referral date was known but referral type was unknown

4.1.14 Goff's Symptom Index

The criteria for a positive "test" according to the symptom index developed by Goff *et al.* were applied to women in our study.[108] Specifically, this was reporting of any one of pelvic/abdominal pain, increased abdominal size/bloating, difficulty eating/feeling full quickly with a frequency of 16-31 days per month but for less than 1 year (see 3.3.10 Definition of Study Variables for full details). Symptoms that were not ongoing at time of questionnaire or interview completion were included since Goff's survey also included these. The index was tested on data for women aged ≥50, then on women of all ages including a split by stage.

Sensitivity and specificity (based on cumulative incidence over 1 year) for interview data in women aged ≥50 was very similar to Goff's confirmatory sample, as shown in Table 4-19. Sensitivity for early stages was almost identical to the Goff data based on questionnaire answers. Conversely, questionnaire sensitivity for women aged ≥50 and late stage disease were lower in comparison to Goff's Index.

To gain an impression of how the index would perform at least 3 months before diagnosis, the same criteria were applied to symptoms present 3-14 months before diagnosis. Based on cumulative incidence for 1 year (over 3-14 months), sensitivity for both sources was markedly reduced but specificity remained the same or better (Table 4-19).

Table 4-19 Sensitivity & Specificity Calculated According to Goff's Index for Questionnaire & Interview Data

	Goff Data[108]	Symptoms at Anytime				3-14 Months Before Reference Date			
		Questionnaire	95%CI	Interview	95%CI	Questionnaire	95%CI	Interview	95%CI
Women aged ≥50									
Sensitivity	**66.7%**	**58.5%** (131/224)	51.7%, 65.0%	**70.8%** (92/130)	62.2%, 78.4%	**27.2%** (61/224)	21.5%, 33.6%	**46.2%** (60/130)	37.4%, 55.1%
Specificity	**90.0%**	**98.9%** (265/268)	96.8%, 99.8%	**89.6%** (112/125)	82.9%, 94.3%	**98.9%** (265/268)	96.8%, 99.8%	**93.6%** (117/125)	87.8%, 97.2%
Women All Ages									
Sensitivity	-	**59.0%** (147/249)	52.6%, 65.2%	**71.7%** (104/145)	63.7%, 78.9%	**32.9%** (82/249)	27.1%, 39.1%	**48.9%** (71/145)	40.6%, 57.4%
Specificity	-	**98.9%** (265/268)	96.8%, 99.8%	**89.6%** (112/125)	82.9%, 94.3%	**98.5%** (264/268)	96.8%, 99.8%	**92.8%** (116/125)	86.8%, 96.7%
Early Stage									
Sensitivity	**56.7%**	**57.0%** (57/100)	46.7%, 66.9%	**68.4%** (39/57)	54.8%, 80.1%	**28.0%** (28/100)	19.5%, 37.9%	**49.1%** (28/57)	35.6%, 62.7%
Late Stage									
Sensitivity	**79.5%**	**65.2%** (86/132)	56.4%, 73.2%	**75.9%** (60/79)	65.0%, 84.9%	**37.1%** (49/132)	28.9%, 50.0%	**51.9%** (41/79)	40.4%, 63.3%

Note: 95% confidence intervals were not provided in the Goff paper. Feeling full quickly was only recorded on interview so in our data sensitivity may be lower, and specificity may be higher, than expected for questionnaire data

4.1.15 Development of a Potential Symptom Index

In order to explore the potential for developing a symptoms index, two different approaches were employed. The first approach used the results of the univariate analyses to produce an unweighted score. The second involved using results from a multivariate analysis (backwards stepwise selection regression) on the questionnaire data to develop a weighted score. Both techniques excluded longstanding symptoms from the analyses. Neither of these approaches were ideal, however both were felt to be adequate for exploratory purposes.

4.1.15.1 Unweighted Symptoms Index

Two (crude) symptoms indices were devised using two different OR thresholds from the univariate analyses. Both thresholds were derived from data on symptoms present over one year (3-14 months before reference date) as per Table 4-8 to Table 4-10. The first index was based on symptoms with ORs of ≥10 in at least two data sources (Index 1). The second was limited to symptoms with ORs of ≥5 (or symptoms with case frequency of ≥6% for infinite ORs) in at least two data sources (Index 2). Women who had at least one index symptom over the year examined were considered to be positive for the index (see Table 4-20). Median time from first index symptom to diagnosis was calculated for women who were positive, and this served as a measure of diagnostic lead time. Note that because this was based on symptoms 3-14 months before reference date, the minimum possible median was 3 and the maximum was 14.99. Index 1 included women with at least one of the following symptoms during 3-14 months before reference date:

- Increased abdominal size
- Loss of appetite

Index 2 included women with at least one of the following symptoms during 3-14 months before reference date:

- Increased abdominal size
- Loss of appetite
- Pelvic/abdominal pain or discomfort
- Lump in abdomen
- Constipation
- Fatigue
- Bloating
- Nausea or vomiting

- Weight loss

As expected, Index 2 produced the greatest sensitivity for cases and the longest median (diagnostic) lead times; however these were offset by a loss in specificity. The highest values (for both indices) were observed for interview data, followed by questionnaire and GP notes. According to interview, 74% (95%CI 67%, 81%) of cases were positive for Index 2, with a median of 8 months from first index symptom onset to diagnosis. Corresponding figures for questionnaire data were 57% (95%CI 51%, 63%) of cases with a median of 7 months (diagnostic) lead time. For GP notes, this was 45% (95%CI 38%, 52%) of cases positive for the index. The change in sensitivity between Index 1 and Index 2 was greatest for GP notes with a 34% increase from 1 to 2. Corresponding loss in specificity was 21%. Questionnaire and interview data had similar magnitudes of sensitivity increase, however the decrease in specificity was double that on interview compared to questionnaire.

In order to examine actual lead time, Table 4-21 shows the proportion of women that were positive for Index 2 at various intervals before diagnosis, with 2 months for referral to diagnosis taken into account. In the table, the proportion of women with an index symptom between 5-8 (i.e. 5 to 7.99) months before diagnosis is shown as the proportion of women with 3-5 months (actual) lead time. According to these calculations, between 9% (GP notes) and 18% (interview) of cases had an actual lead time of 3-5 months. Between 26% (GP notes) and 56% (interview) of cases had at least 3 months actual lead time (sum of 3-5 to 9-12 months lead time). Note that 43% (questionnaire), 26% (interview) and 55% (GP notes) of cases had no Index 2 symptoms between 3-14 months of diagnosis, and therefore no actual lead time.

Table 4-20 Number (%) [95%CI] of Women with Positive Index Over 1 Year (3-14 Months Before Reference Date), Median Time from Earliest Index Symptom Onset to Reference Date (Months) & Crude Odds Ratios (95% CI) Per Source

		Questionnaire			Interview			GP Notes		
		Cases n=249	**Controls n=268**	**OR (95%CI)**	**Cases n=145**	**Controls n=125**	**OR (95%CI)**	**Cases n=223**	**Controls n=227**	**OR (95%CI)**
Index 1*	**n (%)**	86 (35%)	7 (3%)	20 (9, 51)	68 (47%)	7 (6%)	15 (6, 40)	25 (11%)	2 (1%)	14 (3, 125)
	95%CI	[29%, 41%]	[1%, 5%]	-	[39%, 55%]	[2%, 11%]	-	[7%, 16%]	[0%, 3%]	-
Median time (IQR)		6 (4-9)	-	-	6 (4-9)	-	-	5 (4-8)	-	-
Index 2**	**n (%)**	142 (57%)	36 (13%)	9 (5, 14)	108 (74%)	33 (26%)	8 (5, 15)	100 (45%)	49 (22%)	3 (2, 5)
	95%CI	[51%, 63%]	[10%, 18%]	-	[67%, 81%]	[19%, 35%]	-	[38%, 52%]	[16%, 28%]	-
Median time (IQR)		7 (5-12)	12 (8-12)	-	8 (5-10)	11 (6-12)	-	6 (4-10)	10 (7-13)	-

*Symptoms present 3-14 months before diagnosis with case-control OR ≥10 in at least 2 sources (increased abdominal size & loss of appetite)
**Symptoms present 3-14 months before diagnosis with case-control OR ≥5 (or present in ≥6% of cases if OR was infinite) in at least 2 sources (increased abdominal size, lump in abdomen, loss of appetite, pelvic/abdominal pain or discomfort, constipation, fatigue, bloating, nausea or vomiting, weight loss)
Note: Median for controls with Index 1 not shown due to small numbers

Table 4-21 Number (%) of Cases Positive for Index 2 at Various *Actual* Lead times using Symptoms Over 1 Year (3-14 Months Before Diagnosis)

	Questionnaire		Interview		GP Notes	
Lead Time*	Cases n=249	95%CI	Cases n=145	95%CI	Cases n=223	95%CI
None	*107 (43%)*	*37% to 49%*	*37 (26%)*	*19% to 33%*	*123 (55%)*	*48% to 62%*
*≥1 month***	*142 (57%)*	*51% to 63%*	*108 (74%)*	*67% to 81%*	*100 (45%)*	*38% to 52%*
1-2 months	41 (16%)	12% to 22%	27 (19%)	13% to 26%	41 (18%)	14% to 24%
3-5 months	41 (16%)	12% to 22%	26 (18%)	12% to 25%	20 (9%)	6% to 14%
6-8 months	21 (8%)	5% to 13%	38 (26%)	19% to 34%	23 (10%)	7% to 15%
9-11 months	39 (16%)	11% to 21%	17 (12%)	7% to 18%	16 (7%)	4% to 11%

*Lead time is time from first index symptom (present 3-14.99 months before diagnosis) to 2 months prior to diagnosis
**Number (%) of women with at least 1 month of actual lead time for symptoms present over 3-14.99 months before diagnosis
Values are calculated as women positive for Index 2 using symptoms present 3-14.99 months before diagnosis.
Note: '1-2 months'=1.0 to 2.99 months, '3-5 months'=3.0 to 5.99 months, etc.

A 3-point scoring system was devised using the two indices:

- Score 0 if have none of the symptoms in Index 1 or Index 2
- Score 1 if have any of the symptoms in Index 2 that are *not* in Index 1 (5≥OR>10)
- Score 2 if have any of the symptoms in Index 1 (≥OR 10)

This was applied to data for symptoms present 3-14 months before diagnosis, the results are shown below in Table 4-22. Clinical application of a 3-point score might involve offering 'targeted screening' to women with a score of 1, and a more urgent test or referral for a score of 2. However, the score would have 'missed' 43%, 26% and 55% of cases for questionnaire, interview and GP notes, respectively, within 3 months of diagnosis. The score performed the best on interview data, with a sensitivity of 28% and specificity of 90%.

Table 4-22 Unweighted Score in Each Data Source for Symptoms Over 1 Year (3-14 Months Before Reference Date)

Score	Cases	Controls	Sensitivity	1-Specificity
Questionnaire	n=249	n=268		
0	107	232	-	-
1	56	29	22%	11%
2	86	7	35%	3%
Interview	n=145	n=125		
0	37	92	-	-
1	40	26	28%	10%
2	68	7	47%	3%
GP Notes	n=223	n=227		
0	123	178	-	-
1	75	47	34%	21%
2	25	2	11%	1%

N.B. A score of 1 in the table refers to the proportion of women with an Index 1 symptom that is not in Index 2 (i.e. not loss of appetite or increased abdominal size)

4.1.15.2 Weighted Score

A weighted scoring system based on questionnaire data for symptoms at anytime (excluding longstanding symptoms) was developed. All of the symptoms that were presented individually in the main tables were put into the model (18 symptoms). A significance level of p=0.1 was required for removal and p=0.05 for entry. Symptoms that were significantly associated with cases in the model were used to derive a weight for each criterion. This had the advantage of accommodating for the varying predictive power of each symptom and of accounting for multiple symptoms. The resultant model identified 12 symptoms that were independently associated with ovarian cancer. Table 4-23 shows the log odds ratios and derived weights for symptoms identified in the regression model output. In order to simplify the calculations for a weighted score,

derived weights were converted to integers by rounding. Symptoms that were dropped from the model because of perfect prediction (i.e. not reported by any controls), were automatically given a weight of +5. The range of possible scores was between 0 and 24.

Table 4-23 Log Odds Ratio and Derived Weights for Symptoms Identified in the Model

Variable	Log odds ratio (β Coefficient)	Derived Weight
Pelvic/Abdominal Pain or Discomfort	1.537135	+2
Abdominal Bloating	1.074012	+1
Increase in Abdominal Size	1.4292	+1
Back Pain	-0.67347	0
Loss of Appetite	2.19223	+2
Irregular Vaginal Bleeding	1.806627	+2
Weight Loss	1.133038	+1
Change in Bowel Habit*	-	+5
Lump in Abdomen*	-	+5
Postmenopausal Bleeding*	-	+5
Vaginal Discharge*	-	+5
Urinary Other*	-	+5

*Symptoms that were dropped from the model due to perfect prediction of case status

As this approach was more data-driven the model was applied only to data for symptoms present 3-14 months before reference date and was cross-validated (for questionnaire data only). The resultant sensitivity corresponds to the proportion of cases that could be 'identified' at least 3 months prior to (current) diagnosis, and 1-specificity corresponds to the proportion of women in general population who would be identified over 12 months.

Scores were derived for each source as shown in Table 4-24. Two thresholds for the score were selected based on the proportion of cases versus controls correctly identified at various cut-off points. For self-reported data, the same two thresholds were used, however for GP notes the spread of scores was different enough to justify devising separate thresholds.

For self-reported data, the first threshold was set at a score of 4 and above. The second threshold was a score between 1 and 3. Again, any women meeting the first threshold might be sent for an 'urgent' investigation or procedure, whereas anyone meeting the second threshold might be offered 'targeted screening' according to routine timelines. For scores of 0 no action would be taken. Based on these assumptions, for the questionnaire data 31.7% (95%CI 26.0%, 37.9%) of cases and 9.0% (95%CI 5.8%, 13.0%) of controls would be sent for 'targeted screening' and 24.1% (95%CI 18.9%, 29.9%) of cases and 0.7% (95%CI 0.1%, 2.7%) of controls would have an 'urgent' test or referral. Likewise, based on interview data 25.5%

(95%CI 18.6%, 33.4%) of cases and 16.0% (95%CI 10.0%, 23.6%) of controls would be ‘screened’, and 43.4% (95%CI 35.2%, 51.9%) of cases and 4.0% (95%CI 1.3%, 9.1%) of controls would have an ‘urgent’ test or referral (Table 4-24).

The two thresholds for GP notes were set at scores of 1-9 for ‘targeted screening’ and ≥10 for ‘urgent’ testing or referral. This translated into ‘screening’ on 46.2% (95%CI 39.5%, 53.0%) of cases and 20.7% (95%CI 15.6%, 26.6%) of controls (Threshold 1), and ‘urgent’ investigation for 6.7% (95%CI 3.8%, 10.9%) of cases and 1.3% (95%CI 0.3%, 3.8%) of controls (Threshold 2).

For questionnaire data, the advantage of adding the second threshold is a 31.7% gain in sensitivity for only a 9.0% loss in specificity. For interview, corresponding figures were 25.5% additional sensitivity, 16.0% loss in specificity, and for GP notes this was 46.2% sensitivity, and 20.7% specificity.

The sensitivity and specificity of the weighted index varied between the three data sources. As seen in Table 4-24 and Figure 4-18 to Figure 4-20, the weighted score performed better on interview data. Performance was noticeably reduced when applied to the GP data.

If a single threshold were to be used with a cut-off score of 1, reasonable sensitivity and specificity can still be achieved (Table 4-25). For the questionnaire, it achieves 55.8% sensitivity (95%CI 49.4%, 62.1%) with 90.3% (95%CI 86.1%, 93.6%) specificity. On interview, one could achieve 69.0% (95%CI 60.8%, 76.4%) but with only 80.0% (95%CI 71.9%, 86.6%) specificity. By raising the threshold of the index (for interview), one could increase the specificity to 96.0% (95%CI 90.9%, 98.7%) but the sensitivity would be reduced to 43.4% (95%CI 35.2%, 51.9%). For GP notes, the sensitivity is similar to that from the questionnaire (52.9% [95%CI 46.1%, 59.6%]), but the specificity is comparable to that of the interview (78.0% [95%CI 16.8%, 28.0%]).

Table 4-24 Weighted Score for Each Data Source for Symptoms Over 1 Year (3-14 Months Before Reference Date)

	Questionnaire		Interview		GP Notes	
Score	Cases n=249	Controls n=268	Cases n=145	Controls n=125	Cases n=223	Controls n=227
0	44.2%	90.3%	31.0%	80.0%	47.1%	80.0%
1	8.0%	3.7%	10.3%	6.4%	4.5%	1.8%
2	18.5%	3.0%	12.4%	6.4%	11.2%	4.0%
3	5.2%	2.2%	2.8%	3.2%	3.1%	0.4%
4	9.2%	0.7%	7.6%	0.8%	1.3%	0.4%
5	4.0%	0.0%	7.6%	2.4%	11.7%	9.3%
6	3.6%	0.0%	6.9%	0.0%	1.8%	0.4%
7	2.4%	0.0%	6.9%	0.0%	9.9%	3.5%
8-9	2.4%	0.0%	5.5%	0.0%	2.7%	0.9%
10-24	2.4%	0.0%	9.0%	0.8%	6.7%	1.3%
TOTAL	**100%**	**100%**	**100%**	**100%**	**100%**	**100%**

Table 4-25 Cumulative Performance of Weighted Scoring System for Symptoms Over 1 Year (3-14 Months Before Reference Date) for Each Data Source

	Questionnaire		Interview		GP Notes	
Score	Sensitivity n=249	1-Specificity n=268	Sensitivity n=145	1-Specificity n=125	Sensitivity n=223	1-Specificity n=227
0	100%	100%	100%	100%	100%	100%
1	**55.8%**	**9.7%**	**69.0%**	**20.0%**	**52.9%**	**22.0%**
2	**47.8%**	**6.0%**	**58.6%**	**13.6%**	**48.4%**	**20.3%**
3	**29.3%**	**3.0%**	**46.2%**	**7.2%**	**37.2%**	**16.3%**
4	*24.1%*	*0.7%*	*43.4%*	*4.0%*	**34.1%**	**15.9%**
5	*14.9%*	*0.0%*	*35.9%*	*3.2%*	**32.7%**	**15.4%**
6	*10.8%*	*0.0%*	*28.3%*	*0.8%*	**21.1%**	**6.2%**
7	*7.2%*	*0.0%*	*21.4%*	*0.8%*	**19.3%**	**5.7%**
8-9	*4.8%*	*0.0%*	*14.5%*	*0.8%*	**9.4%**	**2.2%**
10-24	*2.4%*	*0.0%*	*9.0%*	*0.8%*	*6.7%*	*1.3%*

Bolded values meet Threshold 1, italicised values meet Threshold 2

Figure 4-18 Receiver Operating Curve (ROC) for Questionnaire Data using Weighted Score on Symptoms Over 1 Year (3-14 Months Before Reference Date)

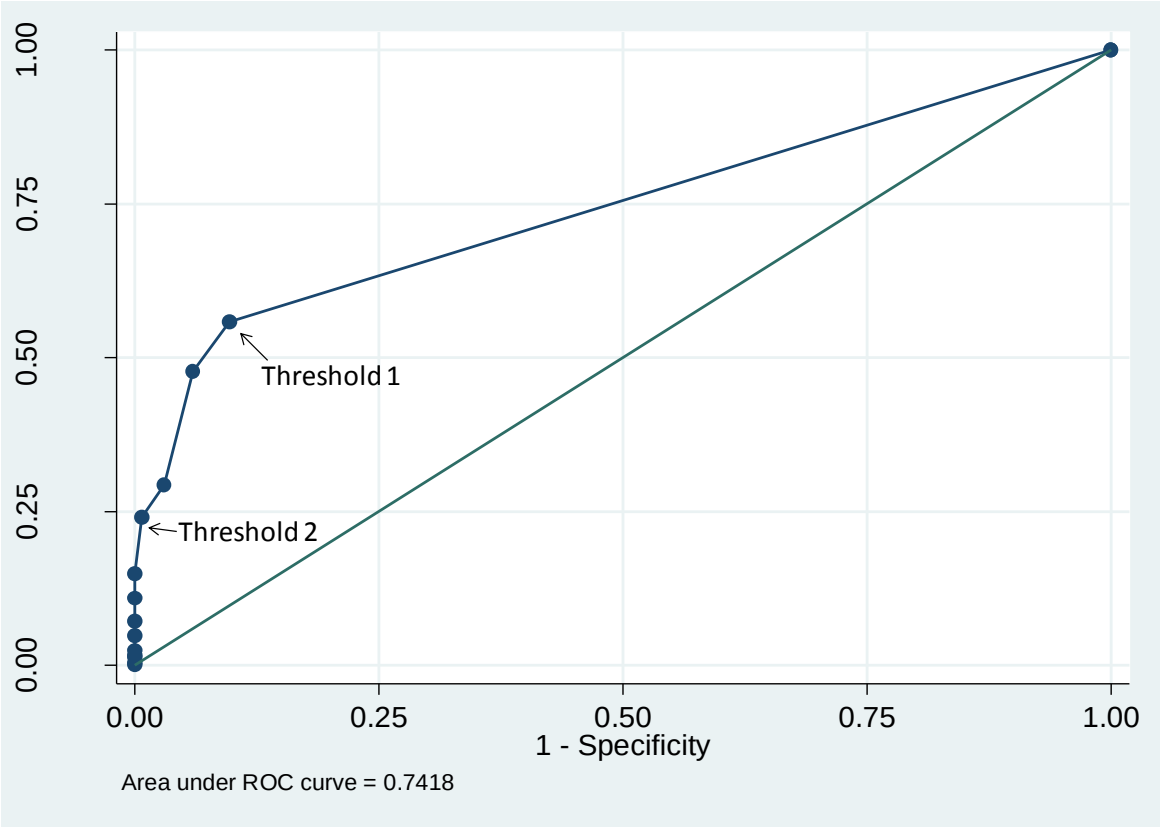


Figure 4-19 Receiver Operating Curve (ROC) for Interview Data using Weighted Score on Symptoms Over 1 Year (3-14 Months Before Reference Date)

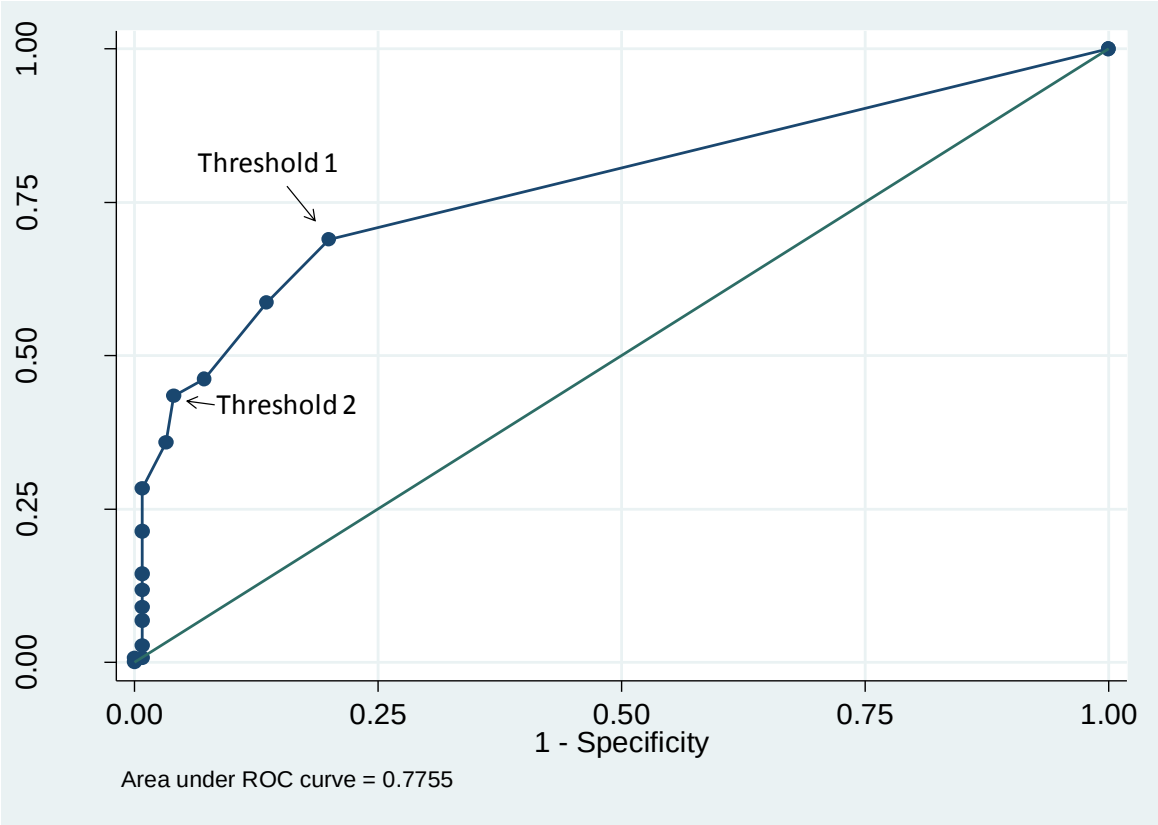


Figure 4-20 Receiver Operating Curve (ROC) for GP Note Data using Weighted Score on Symptoms Over 1 Year (3-14 Months Before Reference Date)

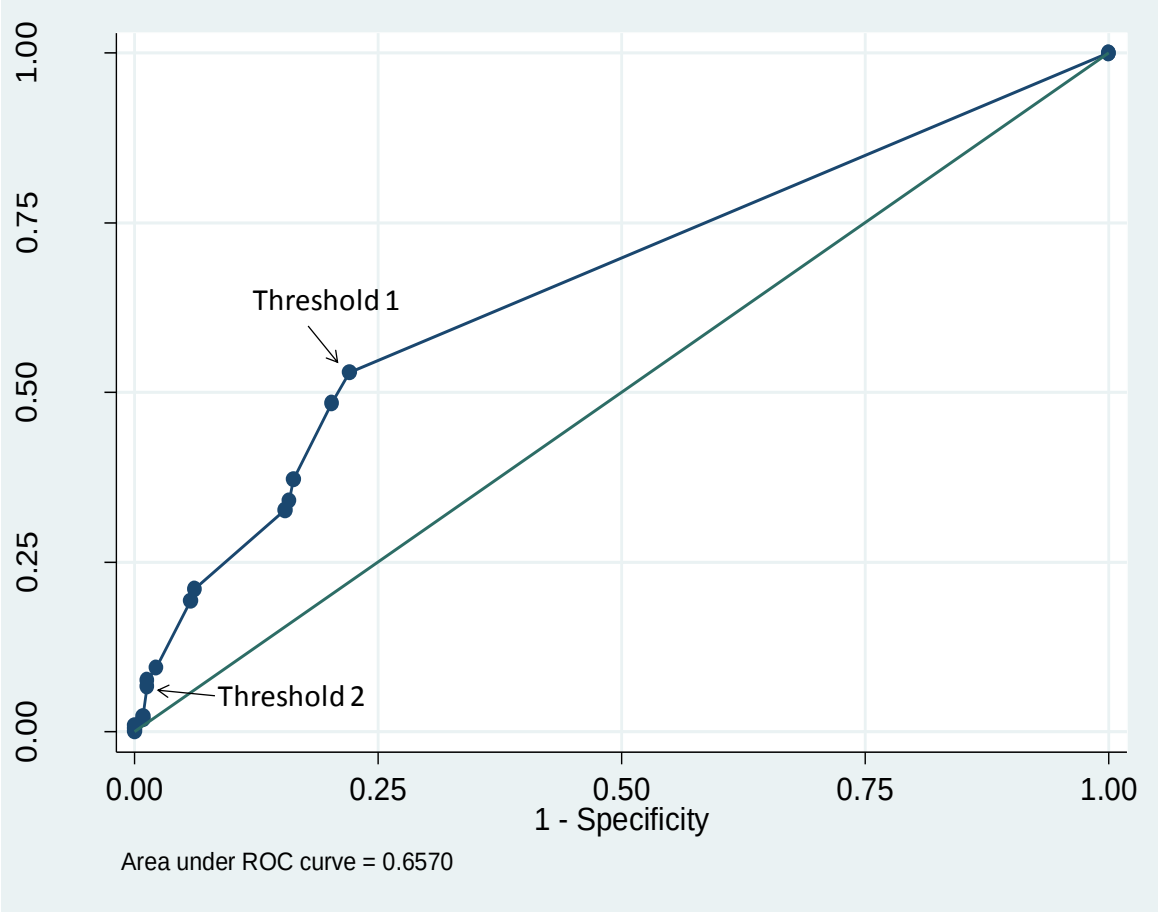
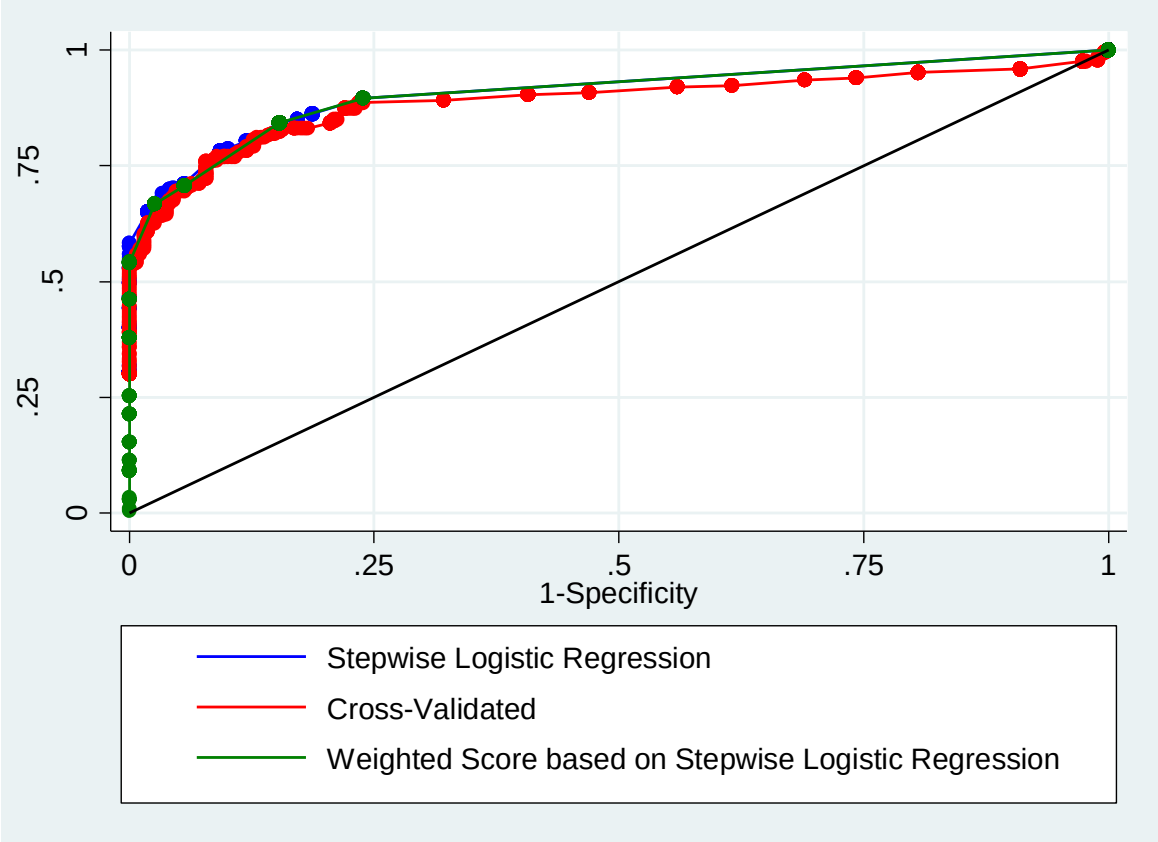


Figure 4-21 Receiver Operative Curves (ROC) for Symptoms Reported at Anytime Questionnaire Data



A 10-fold cross validation was performed on the model using questionnaire data for symptoms at anytime (i.e. the same data used to derive the original model). Figure 4-21 compares the resultant receiver operator curves (ROC) for the original model, the cross-validated version and the weighted score. Each ROC was similar showing that each of the models/score performed equally well.

Figure 4-22 Receiver Operative Curves (ROC) for Symptoms Reported Over 3-14 Months Before Reference Date on Questionnaire

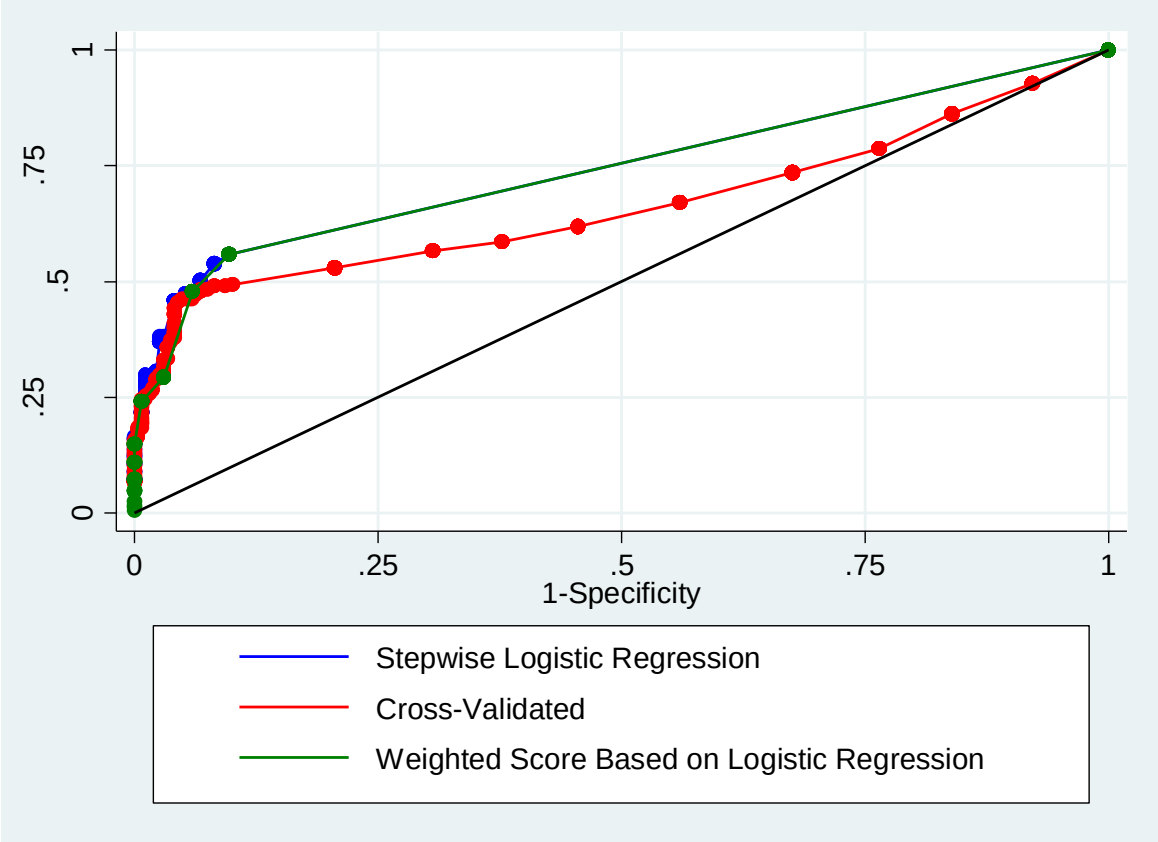


Figure 4-22 contains the same ROCs as above applied to questionnaire data for symptoms present over 3-14 months before reference date (1 year). At first glance, the cross-validated version performance appears to be poor in comparison to the other models. However, when sensitivity is less than 0.5, there is actually very little difference between the 3 models.

5 CHAPTER 5: Case-Control Discussion & Conclusions

5.1 Discussion

The findings of this chapter will be discussed in several different segments. This will include:

- Sensitivity of Prediagnostic Symptoms
- Specificity of Prediagnostic Symptoms
- Salient Symptoms
- Symptom Lead Time
- Exploring the Potential For a Symptoms Index
- Route of Diagnosis
- Delays in Diagnosis
- Early versus Late Stage Disease
- Potential for Using Symptoms as a Tool for 'Early' Detection
- Potential for 'Targeted Screening'
- Comparison of the Data Sources
- Severity & Frequency
- Missing Data
- Study Advantages
- Study Limitations

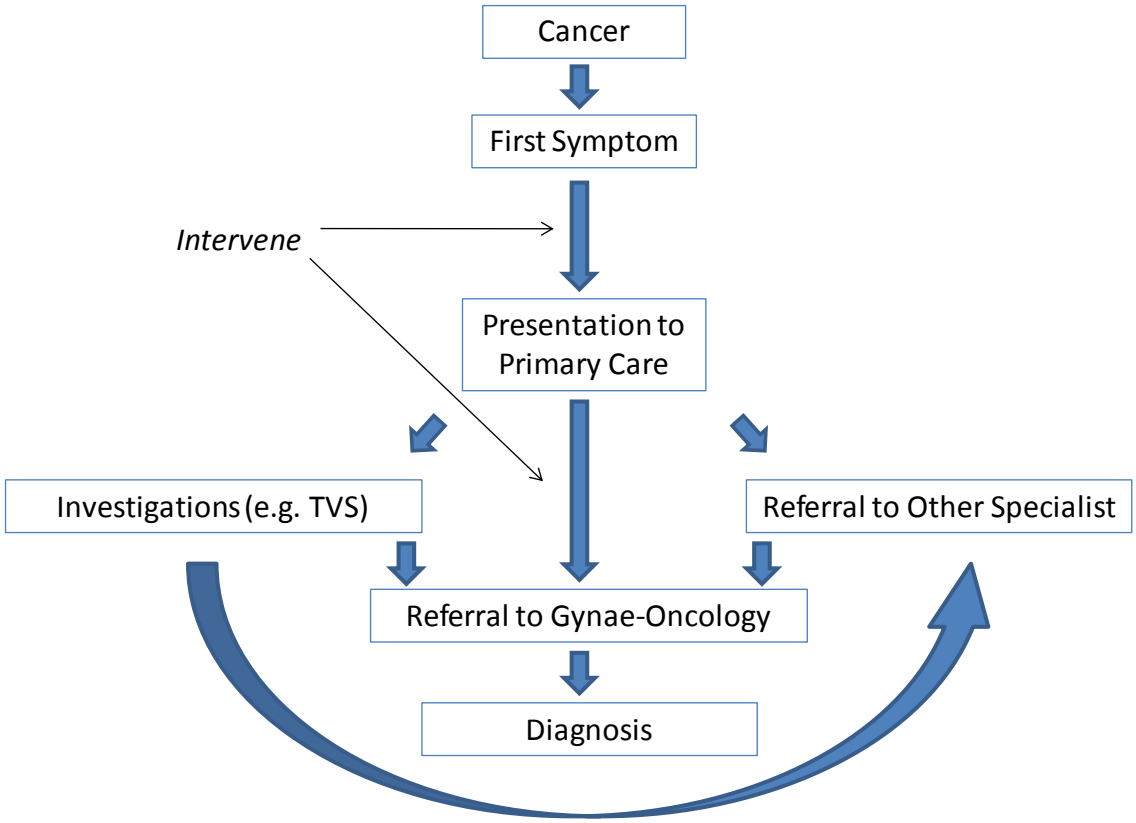
Sensitivity of Prediagnostic Symptoms

In the present study, there was little interest in the full spectrum of symptoms, but rather only the sensitive and specific ones. In corroboration with previous studies, abdominal and gastrointestinal symptoms predominated, while gynaecological symptoms were rare.[6-9, 12, 15, 16, 20, 21, 67-69, 75, 89, 104] Ninety-seven percent (questionnaire, [Q]) and 100% (interview, [I]) of cases had at least one relevant symptom at diagnosis (including longstanding symptoms). The most sensitive symptoms (excluding longstanding) in self-reported data were pelvic/abdominal pain (Q 53%, I 57%), increased abdominal size (Q 67%, I 77%), abdominal bloating (Q 69%, I 56%) and fatigue (Q 50%, I 55%). According to the GP note data, pelvic/abdominal pain or discomfort (58%) and increased abdominal size (27%) were of relatively high sensitivity, as were urinary other (37%) and nausea/vomiting (25%). Importantly, pelvic/abdominal pain or discomfort and increased abdominal size were the only

symptoms that were consistently amongst the top three most sensitive symptoms for all data sources. While these figures are not appropriate for direct comparison to the existing literature (due to the exclusion of longstanding symptoms), the underlying theme of abdominal pain, increased abdominal size, and bloating being the most common symptoms reported is in strong agreement with previous studies.[6-12, 15-17, 19-21, 67-69, 75, 89, 97, 99, 101, 104]

Much emphasis has been placed upon symptoms at diagnosis, however there is little to be gained by detecting symptoms that only develop within 2-3 months of diagnosis. This is because any symptoms presenting after this point offer little opportunity for intervention. In this study, the median time between first referral and diagnosis was 1.6 months (IQR 1.1, 2.5). Thus, it appears that according to current UK clinical practice, the time it takes to get from any referral to ovarian cancer diagnosis is typically around 2 months. Time between rapid referral G/O to diagnosis was not much shorter at 1.4 months (IQR 1.1-2.0). Critically, the interval we wish to have an impact on is the time between symptom onset and referral (as shown by the black arrows in Figure 5-1). Most of the current literature has failed to take this into account when evaluating their findings. In the UK, the period between referral and diagnosis is already being tightened as part of a Department of Health strategy to reduce cancer waiting times. The most recent waiting time targets were set at a maximum of 2 months (62 days) from urgent GP referral to first definitive treatment.[132]

Figure 5-1 Schematic Diagram Showing Typical Pathway from Oncogenesis to Ovarian Cancer Diagnosis



Unfortunately, there was a large reduction in individual symptom sensitivity 3-14 months before diagnosis compared to symptoms at diagnosis (~20-30% for the most sensitive symptoms). Crude odds ratios for individual symptoms were still large and significant at the $p<0.05$ level for many symptoms, however only in relatively small proportions of women. Nonetheless, 81%, 70% and 64% of cases still had at least one relevant symptom on interview, GP notes and questionnaire, respectively. A similar pattern of symptom sensitivity was maintained, however fatigue became comparatively more prominent on interview data while increased abdominal size and nausea/vomiting ceased to feature prominently on GP notes. Pelvic/abdominal pain or discomfort remained amongst the most sensitive symptoms (interview [36%], questionnaire [26%] and GP notes [26%]). Smith and colleagues also looked at symptom frequency for specific prediagnostic time periods using coded billing records.[20] They found that abdominal pain was the most frequently coded symptom at 1-3 months (31%) and 4-6 months (8.9%) prior to diagnosis. Generally, symptom incidence was low in the Smith study, particularly during 6 months or longer before diagnosis. Another study that used medical insurance claim forms, examined temporal symptoms patterns over the 9 months before diagnosis.[21] By 0-30 days prior to diagnosis abdominal pain had been recorded in 83% of cases. At 31-60 days prior, just 23% of cases had a recording of

abdominal pain. For both of these studies, symptom recording was restricted by coding rules, hence results probably underestimate actual symptom prevalence. Overall, individual symptom sensitivity in the present study was less than 50% for all symptoms and all sources over the 3-14 months before diagnosis. These data indicate that there is clustering of symptoms close to diagnosis.

Specificity of Prediagnostic Symptoms

Specificity is regarded as a crucial element of any symptoms tool for ovarian cancer diagnosis given the potentially serious physiological, psychological and financial downstream consequences of false positives. It is dictated by the background rate of relevant symptoms both in primary care (i.e. symptoms that elicit healthcare-seeking behaviour) and in the general non-consulting population (i.e. all symptoms experienced by women). Specificity was highest for loss of appetite, lump in abdomen and increased abdominal size for all three data sources (between 98-100% over one year [3-14 months before diagnosis]). Likewise, case-control differences over 3-14 months prior to reference date were also the greatest for these 3 symptoms, for all data sources (ORs all >10, see Table 4-8 to Table 4-10). Symptoms with significant case-control differences for GP notes also included pelvic/abdominal pain or discomfort, vaginal discharge, bloating, constipation, change in bowel habit, urinary frequency/urgency and urinary other.

Even though this study was not the ideal setting in which to examine symptom specificity (due to the potential for self-selection bias and a 'healthy volunteer effect'[133]), it still provided useful data in an area for which there is a paucity of evidence. The exclusion of pre-existing (longstanding) symptoms greatly improved specificity, with relatively low numbers of controls reporting 'new' symptoms on any data source. Importantly, there was only a modest loss in sensitivity for the symptoms that seemed to be the most affected by over-reporting (see Figure 4-2).

In general, controls reported more longstanding symptoms than cases. Reasons for this are unclear, but perhaps this reflects excess reporting of mild symptoms in controls, which cases failed to mention due to overshadowing by severe symptoms. Also, patients are known to over-report symptoms that correspond to current illness and underreport those that predate the illness.[116] Similarly, Vine *et al.*[7] found that limiting symptoms to those that started within 2 years of diagnosis reduced control symptom frequency from 52% to 38%, whereas case symptom frequencies remained the same. Collectively, these discrepancies imply that symptom studies which include longstanding existing symptoms may overestimate case-control differences, symptom

sensitivity, delays and symptom lead time. This incongruence in symptom reporting by controls will be discussed further in 'Comparison of the Data Sources'.

Salient Symptoms

The most salient ovarian cancer symptoms according to the exploratory analyses for symptoms indices (discussed later) were:

- Increased abdominal size (both scores)
- Loss of appetite (both scores)
- Pelvic/abdominal pain or discomfort (both scores)
- Lump in abdomen (both scores)
- Abdominal bloating (both scores)
- Weight loss (both scores)
- Constipation (both scores)
- Fatigue (both scores)
- Nausea or vomiting (unweighted score only)
- Change in bowel habit (weighted score only)
- Irregular vaginal bleeding (weighted score only)
- Postmenopausal bleeding (weighted score only)
- Vaginal discharge (weighted score only)
- Urinary other symptoms (weighted score only)

All of the symptoms in the Goff Index,[108] also featured in our list, except for difficulty eating or feeling full quickly which were not listed on our survey. Increased abdominal size, bloating, pelvic or abdominal pain or discomfort, weight loss, constipation, fatigue and loss of appetite were found to be salient in both of the approaches and are the most likely to be useful symptoms in any future ovarian cancer symptoms index. Back pain was found to be negatively associated with ovarian cancer in the multivariate analysis. Potentially, back pain could be used as a negative discriminator in a symptoms tool. Some of the symptoms were highly predictive due to a very low number of controls reporting the symptoms which produced large ORs (i.e. lump in abdomen, loss of appetite, vaginal discharge, change in bowel habit, IVB, PMB, nausea/vomiting). Conversely, the remaining symptoms on the list were relatively common in control women, but were still reported in excess by cases (i.e. fatigue, bloating).

Symptom Lead Time

Data from all three sources indicated that symptoms associated with disease develop 15 months prior to diagnosis. The (diagnostic) lead time of individual symptoms was quantified by calculating continuation odds ratios at different periods before diagnosis, using only symptoms that were 'new' in the last 2 years and assuming a constant rate of 'new' symptoms for controls. The continuation odds provided a more realistic estimate of (diagnostic) lead time by only including women who were still 'at risk' of developing the symptom at each time point relative to diagnosis. Also, the resultant confidence intervals were wider but more accurate. Any seasonal variations in control symptoms were expected to have smoothed over the two years worth of data. A priori, one might anticipate some increases in symptom incidence due to ageing but these should have been minimal given the short period examined (i.e. two years). Also, Goff and colleagues found that reporting of potentially related ovarian cancer symptoms in clinic controls (aged 15-90 years) decreased with age.[75]

The data from individual symptoms in this study suggested that there is a maximum diagnostic lead time of 15 (i.e. 14.99) months. However, after allowing 2 months for time between referral and diagnosis (median 1.6 months in this study), the real potential maximum referral lead time (i.e. actual lead time) is probably more in the region of 12 months. Odds ratios were substantially larger closer to diagnosis for all symptoms, but were close to 1 at 15-23 months before reference date. This indicated that currently ovarian cancer is symptomatic for about 15 months prior to diagnosis.

In general, odds ratios were larger for questionnaire data due to lower symptom reporting in controls. Certain symptoms (bloating, indigestion, back pain, fatigue) were recorded less frequently in the GP notes, potentially reflecting symptoms which are too common or mild to prompt presentation to primary care. On interview, controls reported higher rates of the same symptoms, which substantiated this theory. Conversely, recording bias is another possible explanation. Bankhead *et al.* found that many symptoms women claimed to have reported to the GP were not actually recorded by the GP.[94] Both are possible explanations for the lower sensitivity observed in GP notes. Loss of appetite and pelvic/abdominal pain or discomfort appeared to provide a maximum of 9-11 months (diagnostic) lead time across all data sources. The longest (diagnostic) lead time in GP notes was for urinary frequency/urgency, which was recorded in significantly more cases than controls 12-14 months before diagnosis (OR 2.9, 95%CI 1.2, 6.8).

Symptom lead time for individual symptoms differed between each data source. In general, lead time was longer for self-reported data compared to GP note data. GP

notes may be the most accurate since they are based on real dates, whereas self-reported dates are based on rounded dates (15th day of the month). However, as shown by the patient delay data, symptom onset dates in GP notes are likely to have been visit dates rather than true onset dates. Problem GP visits clearly started to diverge around 6 months before diagnosis (see Figure 4-11) which lends support to a more conservative diagnostic lead time of around 6 months in GP note data. Cumulative incidence of pelvic/abdominal pain for cases in GP notes sharply increased within 6 months of diagnosis. A study examining insurance claims also found that abdominal pain and GI symptoms displayed a similar pattern, albeit at the more latent time point of 90 days (~3 months) prior to diagnosis.[21]

Exploring the Potential for a Symptoms Index

When the potential for development of a symptoms index was considered, the idea was to take an exploratory approach and there was no desire to over-interpret the findings. A symptoms index is the most likely way in which 'targeted screening' would be applied since the positive predictive value (PPV) and sensitivity of individual symptoms is liable to be too small. It should be noted that none of the scores devised formally took into account the timing of symptoms. However, both scores were applied to data on symptoms that were present 3-14 months before diagnosis.

Specificity for Index 1 was extremely high, however sensitivity was poor and the largest proportion of cases that would have been detected was 47% (interview). The sensitivity and specificity for Index 2 was much more promising, hence will be discussed in greater detail.

Based on Index 2 for symptoms reported over 1 year, sensitivity was between 45% (GP notes) and 74% (interview) and specificity was between 74% (interview) and 87% (questionnaire). This is equivalent to between 13% (questionnaire) and 26% (interview) of healthy women that would require testing over one year. After time for referral to diagnosis was accounted for, the proportion of cases with an actual lead time of at least 3 months was between 26% (GP notes) and 56% (interview). Thus, there is potential to bring forward diagnosis in a reasonable proportion of cases. Median time from first Index 2 symptom to diagnosis was 6 months for GP notes and 8 months for interview, indicating a sizeable gain in diagnostic lead time for these women. After allowing for time from referral to diagnosis, this would be reduced to 4 and 8 months of actual lead time, respectively.

The performance of Index 2 (unweighted score) was very similar to that of the weighted score using the cut-off of 1. However, notably the specificity of the Index 2 was slightly

better (except on questionnaire). It is of interest to consider the difference in the performance of the weighted index using the different data sources. The sensitivity of 52.9% in GP notes is attainable since these are essentially prospectively recorded data. The fact that the specificity is only 78.0% which is considerably lower than on questionnaire, but similar to interview, suggests that many of these symptoms were transitory and/or self-limiting and therefore were not reported or recalled by controls at the time of recruitment. However, on interview these may have been reported or recalled in response to probing. Whether this means that there would be potential to improve the specificity of the index by requiring that the symptoms should be present for several weeks is unclear. Potentially, the triage should take into account not only the blood test result, but also whether the symptoms are still present by the time the result are available (typically likely to be about 2 weeks after the sample is taken). The greatest sensitivity for the thresholds chosen was for interview data at 69%, with a corresponding specificity of 80%. The greater sensitivity observed in the interview data suggests that it may be possible to impact on a greater proportion of ovarian cancers if women were to recognise symptoms and go to their GP earlier. The challenge would be to do that without dramatically reducing the specificity of the index. Another important finding was that in order to obtain over 90% specificity with the GP notes, the sensitivity falls to just 21.1%. Corresponding sensitivity on questionnaire and interview was 55.8% and 46.2%, respectively. It may be that women with high scores on the weighted index are already being referred by their GPs, so that the real gains for women with high scores are to be made by encouraging such women to consult with primary care sooner. On the other hand, it may simply be due to the fact that the weighted index was derived using questionnaire data (which is similar to the interview data).

As this process was hypothesis-generating, cross-validation was performed on the questionnaire data. The ROCs in Figure 4-21 showed that when the models and score were applied to symptoms reported at anytime, each approach performed equally well. However, when these were applied to symptoms present over 3-14 months before reference date (1 year), each of the approaches performed much worse (Figure 4-22). Overall, this demonstrated that timing of symptoms is crucial when developing a symptoms index. This fact is further exemplified by the results from the application of our data to the Goff index (discussed later). Furthermore, any potential model used to develop a symptoms index is likely to require validation and needs to carefully consider the source of the data.

As expected, symptom lead time (diagnostic and actual) was generally longer on self-reported data versus GP note data for individual and Index symptoms. This may reflect the period between symptom onset and health-seeking, in which case encouraging women to present to primary care earlier could increase the number of women with a sizeable lead time in GP notes. Another possible explanation for this apparent excess in lead time is that the delay in seeking healthcare is more accurately measured by self-reported data. Alternatively, recall bias or recall error may have played a role.

The longest (diagnostic) lead times for Index 1 and 2 were reported on interview, but again, it is unclear whether this is artificially long or represents 'real' lead time. Perhaps the timing of the interview created a bias since they were conducted after diagnosis which may have increased recall bias and recall error. However, anchor points were used on interview, which can aide recall, particularly with regard to dates. As such, symptom start dates on interview were expected to be more accurate than those on questionnaire. At the time of interview, most cases were already aware of their final diagnosis. Consequently, one might anticipate excess recall bias on interview versus questionnaire. Presumably, this disparity would be relatively modest as at the time of questionnaire completion cases are already aware that there is a high probability of cancer and patient anxiety will be high. Ultimately, it is unclear to what extent each of these contributed to the discrepancies in lead time. A more complex combined analysis using an algorithm may elucidate this. The possible reasons for discrepancies between data sources are discussed in further detail later.

Route of Diagnosis

First GP referral was difficult to extract for some women, and required subjective interpretation. The subjectivity arose from deciding which symptoms are attributable to disease. For women who were relatively symptom-free until a given point, route of diagnosis was easy to identify. For women who had a high background rate of symptoms, previous referrals often confused matters. For example, a woman may have been referred to secondary care one year before diagnosis for an acute episode of abdominal pain. This may be followed by a symptom-free period until an abdominal lump is palpated two months before diagnosis, for which she is sent to rapid referral G/O. In this situation, one might choose to assume that the referral for abdominal pain was unrelated and should not be counted, conversely, one might decide that this earlier referral for abdominal pain was the first GP referral. For the purposes of studying route of diagnosis in this study, the former assumption was made. A further issue was that some GP medical records were sparse. In the most extreme situations, there would be documentation of a referral, but no record of any visit or symptoms.

Almost all cases were diagnosed via symptoms, however only 40% were initially referred to a rapid referral to gynaecological-oncology clinic. This is a vast improvement on the 14% reported by a previous study based on data from 3 years after the introduction of rapid referral.[134] Referral to GI specialists continued to be common. One fifth were referred to gastrointestinal clinics, and 40% of these were urgent referrals which indicate that GPs suspected something serious, but erroneously assigned aetiology. This is not surprising given that most symptoms were abdominal or gastrointestinal.

Interestingly, the median time between referral and diagnosis was similar for rapid referral G/O versus all other referral types (1.4 months, IQR 1.1-2.0 and 1.8 months IQR 1.0-3.2, respectively). Both of these medians fall within the current government target of 62 days from urgent cancer referral to first definitive treatment[132] (date of diagnosis is usually date of surgery in ovarian cancer). Even when emergencies were excluded, the corresponding increase in median delay was less than one month (median 2.3 months, IQR 1.4-5.0).

Overall, the findings from the present study confirmed previous reports of convoluted diagnostic routes in ovarian cancer.[8, 11, 99] However, this did not appear to translate in to unacceptable delay (in referral to diagnosis) since the timelines are still coming within government targets. Also, convoluted diagnostic routes are not specific to ovarian cancer. Of all cancers diagnosed in the UK in 2005, just 42% were diagnosed via urgent cancer referrals and 5% through screening programmes.[135] Subsequently, more than 50% of cancers were diagnosed via other routes such as opportunistic diagnoses, A&E and non-urgent referrals.

Delays in Diagnosis

Delays in diagnosis are commonly divided into patient and provider delay. The former refers to the interval between symptom onset and healthcare-seeking, and the latter refers to interval required to make the diagnosis after first presentation.

Delay calculations are complicated by the subjectivity associated with symptom attribution. If no restrictions are placed on how far back symptom history is reviewed; delays can be artificially large. An extreme example of this is a Norwegian study which found that some women reported symptoms that started 10 years before diagnosis.[13] In our study, no differences in case-control symptom incidence were observed beyond 15 months prior to diagnosis. Hence, delay calculations were based on symptoms which started in the last 15 months before diagnosis. Restricting the symptoms to those new in the last 15 months has the drawback of reducing the apparent delays for

women who had symptoms associated with disease before this cut-off. However, judging by the data in the present study, this would be very rare. The advantage of using the cut-off is that it avoids the much greater problem of including spurious (pre-existing) symptoms which would artificially inflate delays.

Women with no symptoms according to our definition (i.e. none that started in the last 15 months and were 'possibly related' to ovarian cancer) were excluded from the analysis of delays. These women were considered to be ineligible for delay calculations, rather than being included as having a delay of 0. The rationale for this was that women with symptoms who sought healthcare promptly and had no delay were not the same as women who had no symptoms and therefore no delay. Therefore, all the delay calculations were based on symptomatic women (with eligible symptoms, as defined above). Women who were screen-detected or incidentally-diagnosed were also excluded. As these methods for calculating delays were novel, the delays results are not strictly comparable to those from other studies. Also, it should be noted that default dates of the 15th of the month were used in self-reported data, whereas the dates in GP note data were exact. This could have contributed to discrepancies between self-reported and GP note delays. Finally, delay data were missing for women with missing symptom onset date and/or first GP visit date, which may have affected the results.

Patient delay varied with each source. According to the questionnaire, patient delay was a median of 1 month (IQR 0-5). On interview, the median patient delay was 3 months (IQR 0-7) which was comparable to previous studies that also found patient delays of 2-3 months.[8, 18] A median of 0 months (IQR 0.0-0.3) patient delay was reported in GP notes, however it is unlikely that this is a true indication of patient delay. It is reasonable to assume that this is spurious and that either GPs did not accurately record symptom onset or women did not report it. Most of the GP notes received were from computerised records, and there has been evidence to suggest that documentation of symptom duration is better in paper versus electronic records.[121] A previous GP record study in the UK found that 78% of women with ovarian cancer consulted within one month of symptom onset, suggesting that patient delay in this country is small.[11]

Other reasons for differences in patient delay between data sources are less clear. As with symptom lead time data, it is possible that interview data contain more accurate dates due to the use of anchor points during elicitation. Conversely, perhaps questionnaire dates were more accurate as there was a shorter period of recall time. While a delay of 1 month (found on questionnaire) is probably of little consequence, a

delay of 3 months (found on interview) may be of greater concern. However, in the consideration of patient delay, it is important to allow a reasonable amount of time to get from symptom onset, to decision to consult, to making and executing the appointment (about 1-4 weeks depending on the symptom). Once this is accounted for, the remaining delay could reflect women failing to recognise symptoms as being important and dismissing them as normal body changes.[104, 107] Most women were symptomatic according to our definition for delay calculations (91% [Q], 99% [I], 97% [GP]). The interquartile range for patient delay for self-reported data was 0-5 months (Q) and 0-7 months (I), thus a quarter of symptomatic cases had a patient delay longer than 5 or 7 months, respectively. However, it should be noted that patient delays were only calculated on a proportion of symptomatic women due to missing data (see Table 4-17). Nevertheless, these data suggest that at least 25% of symptomatic women are not seeking health care within a reasonable time frame. A survey in the UK conducted by Dr Foster Intelligence found that only 16% of women knew that ovarian cancer existed, and just 34% of women were able to name a single symptom associated with ovarian cancer (*http://www.ovarian.org.uk/news/awareness_study.asp*). A lack of awareness of atypical symptoms has been linked to the risk of patient delay in other cancers.[136] However, in ovarian cancer there are many unanswered questions, and the fact remains that there is still no evidence to prove that getting women to present earlier will save lives. Hence, it would be prudent to carry out further research before promoting symptoms awareness in women, however, as already discussed in Part I symptoms awareness campaigns have been ongoing for some time.

Provider delays were shorter in self-reported versus GP note data (median of 2.1, 2.7 and 4.4 months for questionnaire, interview and GP notes, respectively). Again, note that provider delays were only calculated on a proportion of symptomatic women. GP notes are considered to be the most reliable data source for provider delay, and 4.4 months is a significant length of time. Moreover, these data indicate that 50% of symptomatic women had a provider delay of at least 4.4 months which is concerning. In addition, a quarter of symptomatic women had delays of at least 4.8 months (self-reported) or 10.0 months (GP notes). All three of the provider delays identified in this study were longer than the mean or median provider delays of ≤1 month reported in most of the literature,[6, 14, 68, 89] although a median of 3 months for early stage disease was reported in one study.[7] Delay calculations in this study were restricted to symptoms new in the 15 months prior to diagnosis. Thus, one would expect provider delays to be shorter than those reported by studies without symptom timing restrictions. It is possible that this excess in delay is specific to the UK, since most provider delays have been based on studies performed in other countries. However, Kirwan and

colleagues found that 73% of women were diagnosed within 3 months of first GP consultation in the UK.[11]

Overall, the provider delays identified in this study are of concern. Potentially, increasing ovarian cancer symptom awareness in GPs could help with this. However, as with patient delay, if GP education is to be considered, it should be done so with the knowledge that there is no evidence to show that it will be beneficial. Also, women may still be diagnosed with advanced disease, even if provider delays are eradicated.

As already inferred, provider delays are comprised of mandatory phases in the diagnostic work-up (e.g. time taken for tests and appointments), in addition to genuine delays. The sources of (genuine) provider delays are more obvious than those of patient delays. The vast majority almost certainly stem from misdiagnosis and incorrect referral pathways, while the rest are due to delays in healthcare system and diagnostic testing errors. These include waiting time for appointments or surgery and administrative errors (e.g. results not being reported, letters misplaced). Diagnostic testing errors can arise if abdominal rather than pelvic ultrasound is performed or if the ultrasound examiner is not experienced enough, and possibly if CA125 levels are not elevated. Patient characteristics such as co morbidities and frequency of GP visits could contribute to a prolonged diagnostic pathway. Diagnostic overshadowing occurs when doctors misattribute symptoms to pre-existing co-morbidities. Symptoms of frequent attenders may be taken less seriously by GPs compared to symptoms of women who rarely consult.

Uncertainty regarding the current diagnostic tools for ovarian cancer may play a role in provider delays. A study in the UK found that some GPs did not feel confident ordering CA125 tests, urgent ultrasound and fast-track referral due to insufficient knowledge of the criteria required and the criticism received when referral is deemed inappropriate.[85] Physician confusion over the role of CA125 was exemplified by a UK audit which found that 48 out of 799 CA125 tests ordered were for male patients (in whom there is no specific use).[137]

Women who presented to A&E before referral was made (or carried out) presumably had a sudden onset of severe symptom(s) or a sudden worsening of symptom(s). Or else, perhaps there was some sort of miscommunication between the women and their GPs, which culminated in the severity of symptoms being downplayed or underestimated. Importantly, the majority of women who presented to A&E had at least one prior visit to the GP for symptoms, which may indicate 'missed opportunities'. However, this is speculative and there were insufficient data to verify this. Potentially,

these women had aggressive cancers which developed rapidly producing severe symptoms over a short period of time. Notably, although emergency visits are usually associated with faster diagnosis, some women were not necessarily diagnosed as a (direct or indirect) result of the A&E visit, but instead followed a convoluted diagnostic route.

In countries where GPs have a gate-keeping role, there is often mounting pressure on scarce resources. Given that the average GP will see one case of ovarian cancer every five years,[30] using a 'targeted screening' approach could help by providing a quick assessment of women in whom there is low suspicion of disease. This would remove some of the 'guess-work' that is required early on in the diagnostic work-up, when GPs may be considering a number of possible and more plausible causes. Obviously, this would need to be accompanied by clear protocols to address issues such as intermediate CA125 results, and to ensure that women with disease are not further delayed by this process. Also, the proportion of women without disease who meet testing thresholds would need to be sufficiently low to avoid overloading GPs.

Total delay was the longest for interview data with a median of 8.4 months. In contrast, median delay was approximately 5 months for both questionnaire and GP note data. Once again, it is not clear if this excess on interview is real or due to recall error or recall bias. The combined total delay quantified the longest possible delay according to all of the data (although the maximum is 15 months due to the methods used). This was a median of 9.8 months, which is shorter than the median (total) delay of 12 months reported by another UK study.[99] A US study also reported a median total delay of 12 months.[7] However, these are all longer than the total delays reported in other countries (see Part I systematic review, median delay typically between 2-4 months).[6, 8, 13, 67, 78, 89, 112]

Time from first GP visit (for an eligible symptom) to referral was of particular interest. Data showed that the median interval was 1.7 months for any referral versus 2.7 months for an 'appropriate' referral (i.e. to gynaecology). Thus, it appears that health care providers (GP and specialists) take an extra month to make an 'appropriate' referral.

Despite misdiagnosis being cited as a cause of delays in diagnosis.[6-10, 12, 15-17, 19-21, 67-69, 75, 89, 97, 99, 101, 104] the difference in median time between referral and diagnosis was less than 2 weeks for rapid referral G/O versus any referral (1.4 and 1.8 months, respectively). Irritable bowel syndrome (IBS) and diverticular disease have been implicated in the misdiagnosis and delays in diagnosis of ovarian cancer. Our findings

did not support this, and the proportion of cases and controls with each condition was similar (see Table 4-1). Also, the proportion of cases with each condition was modest (<10%). Yet, less than half of the cases were initially given a rapid referral to G/O. These data are somewhat conflicting, but on the whole it seems that most women with ovarian cancer are not appropriately referred initially, but this may only have a small influence on time to diagnosis.

Delays in diagnosis are important as they have been shown to affect survival in other types of cancer.[138] In ovarian cancer, any association between delays in diagnosis and survival are yet to be proven. Crucially, our study demonstrated that delays in diagnosis do exist in a reasonable proportion of women. Furthermore, these delays appear to be longer than those reported in other countries. Most of the delay appears to be during the interval between first GP visit to diagnosis (provider delay) rather than between symptom onset and first GP visit (patient delay). Reasons for this are still unclear. So far social class has not been found to be a significant factor for overall delay in British women with ovarian cancer.[139] Insufficient data exist in delays in diagnosis and further work is needed. In particular, identifying which of the data sources provides the most accurate measure of diagnostic delay would be important, given the varying results from each data source.

Symptoms in Early versus Late Stage Disease

Detection of women while disease is in its early stages is the underlying goal of any ovarian cancer screening programme. Nowadays, the existence of prediagnostic symptoms in early stage disease is much more widely accepted. Almost all women with early stage disease reported symptoms in the present study (100% [GP], 96% [I], 95% [Q]). However, to date there has been little evidence to suggest that there is a chronological pattern to symptom development that would allow identification of women at early stages before they 'progress' to late stages. In fact in our study, the temporal pattern of symptoms for early versus late stage disease appeared to be extremely similar (see Figure 4-15 to Figure 4-17). This is at odds with the findings of Olson *et al.*[69] who found that symptom duration tended to be shorter for early versus late disease.

This failure to establish a temporal symptom pattern and the lack of evidence for 'progression' of symptoms has often been ascribed to women with early stages having biologically less aggressive disease than women with late stages, and vice versa.[9, 12] This hypothesis will be discussed further later on in this chapter.

The overall spectrum of symptoms was also comparable for the two groups with a few key exceptions. Early stage disease was associated with more urinary symptoms, IVB, PMB and findings of an abdominal lump, however these failed to reach significance even at diagnosis for any data source. These findings concur with previous symptom studies which also found that women with early stage disease tended to report more urinary symptoms, gynaecological symptoms[68, 94, 101] and an abdominal mass.[16, 68] This tendency for urinary symptoms to predominate in early disease could be due to tumour pressure effects since early stage tumours are known to be larger.[12, 140]

Women with advanced stage disease reported more indigestion, loss of appetite, abdominal bloating, increased abdominal size, nausea/vomiting and weight loss. These differences were statistically significant for increased abdominal size, indigestion and weight loss 3 months prior to diagnosis. These findings also correlate well with the literature. Several studies have found that abdominal swelling, fatigue and gastrointestinal symptoms[101] were cited more frequently by women with advanced stage disease.[9, 12, 16, 68, 69] Another feature of our data that corroborates well with the literature, is that women with late stage disease reported a significantly greater number of symptoms (see Table 4-14).[8, 12, 75, 104]

Systemic or constitutional symptoms (i.e. fatigue, weight loss, loss of appetite) are often assumed to be indicative of advanced stage disease. While this may be a valid presumption for cancers in general, it does not seem to hold true for ovarian cancer. Several studies have shown that women with benign, borderline and early stage disease also experience the same systemic symptoms.[6, 7, 9, 13, 67, 68, 75] In our study, systemic symptoms were also cited by women with early stage disease (see Figure 4-15 to Figure 4-17).

Potential for Using Symptoms as a Tool for 'Early' Detection in Ovarian Cancer

Early detection of cancer in general is an area of growing interest due to the favourable relationship between early diagnosis and lower mortality rates. The underlying hypothesis for using ovarian cancer symptoms as a tool for this is that acting on symptoms quickly will result in earlier stage diagnosis thereby reducing avoidable deaths and improving mortality. Although this hypothesis is a long way from being proven, the current study has shown that ovarian cancer clearly elicits symptoms in the months before diagnosis, even in early stage disease. The majority of these symptoms first appeared within 3 months of diagnosis, however in a small proportion of women; symptoms were detectable as much as 12-14 months before diagnosis.

Much emphasis has been placed upon the tendency for ovarian cancer symptoms to be persistent.[69, 75, 99] This apparent persistence is irrelevant in the approach proposed in this thesis, as testing would be offered the first time a woman presents with a symptom(s). However, one may wish to set a minimum duration for any potential symptoms screening tool.

Since ovarian cancer is relatively rare, and the downstream consequences of a positive screening test are serious, it has been suggested that a screening test (leading to surgery) for ovarian cancer needs to have a specificity of at least 99.6% to be acceptable.[51] More recently, positive screening may be managed by repeat screening or further non-invasive testing, so the specificity needs to be balanced against the consequences of a false positive at each 'testing' stage. The situation with ovarian cancer is also complicated by the possibility that the inherent biology of the disease may be such that some aggressive tumours develop so rapidly that 'early' detection is not viable.[66] Hence, slow growing tumours may have a bigger window of opportunity for detection. This concept challenges the notion that temporally early detection and treatment of cancer is always associated with a favourable outcome. However, this may only be an issue for a minority of cases and there still could be a sizeable proportion of women in whom diagnosis could be expedited. As mentioned earlier, symptoms that started within 3 months of diagnosis are unlikely to provide any valuable lead time since much of this period is already needed for referral to diagnosis.

Unfortunately, there are two critical unknowns in the oncogenesis pathway for ovarian cancer. The first is the length of time between the onset of malignancy to the appearance of symptoms. The second is the period of advancement required to alter the natural history of disease. There are virtually no data on the growth rate of ovarian cancer tumours, however hysterectomy studies have shown that disease can develop in as little as one year.[141] Optimal surgical cytoreduction (i.e. <1 cm residual disease) is one of the most significant prognostic factors in ovarian cancer.[142-144] The probability of attaining optimal surgical cytoreduction can be thought of as a function of time, the details of which are currently unknown. It is possible that even a short (actual) symptom lead time of 1-2 months could be critical. Clearly, this is speculative and further research is required to prove this. Nevertheless, these short gains in lead time may be worthwhile pursuing if there is potential for psychological and physiological benefit, providing the 'harm versus help' issue is addressed.

Positive predictive value (PPV) of symptoms was not calculated since accurate computation requires data on the prevalence of women aged ≥45 with occult ovarian cancer in the general population, and the proportion of such women who have

symptoms. These data are currently unknown. Estimates should be available from the UKCTOCS study in the future. It is worth noting that in previous studies, PPV has been calculated based on the assumption that symptom sensitivity is constant before diagnosis. However, prediagnostic symptom sensitivity (and therefore PPV) is likely to fluctuate over time, with greater sensitivity expected closer to diagnosis. This was demonstrated by the change in OR of symptoms at different prediagnostic periods. Questionnaire sensitivity (i.e. cases) for bloating at diagnosis was 71%, but only 30% 3-14 months before diagnosis. Therefore, any future PPV calculations should take this into account.

The concept of using symptom recognition as the first step in a multi-step screening programme is not new. A UK group performed a pilot study in primary care practices in East London to investigate the feasibility of implementing a 'targeted screening' programme.[111] However, GP compliance was hugely variable, and no cases of ovarian cancer were identified. Hence, strong conclusions regarding the potential effectiveness of 'targeted screening' could not be drawn. Since the initiation of this thesis, Goff *et al.*[108] have taken this to the next level by developing a symptoms index.

The symptoms index developed by the Goff group aims to discriminate between women with undiagnosed ovarian cancer and healthy women.[108] The index was evaluated retrospectively on a set of cases and controls (high risk women participating in a screening study), working with the assumption that women with a positive index would be sent for CA125 testing and TVS.[103] Using this approach, they hoped to be able to identify women at the primary care level who do not have elevated CA125 but might benefit from immediate TVS. This strategy removes the disadvantage of the low sensitivity of CA125 which is just 50% for stage I cancers,[145] but the lack of specificity of TVS could be a major issue.[63] The authors found that half (11%) of the cases who did not test positive for CA125 tested positive for the index. For the index in combination with CA125 as a first-line test, there were 53% of cases who were positive for both. Sensitivity for early stage disease was relatively high for the index and CA125 combined with 81% of cases testing positive for at least one of the index or CA125. The corresponding proportion of women with late stage disease was 95%. While these figures are promising, specificity could be a problem since 17% (combined tests) and 12% (index) of their control group also tested positive. Importantly, timing of symptoms was not taken into account in these results.

Application of the Goff index to our data produced strikingly similar results to those reported by the group. Notably, our symptoms survey did not specifically ask about difficulty eating or feeling full quickly, which may explain some of the lower sensitivity

observed in comparison to Goff's findings. Also, the frequency of 16-31 days per month was applied to our data instead of >12 times per month. Otherwise our data were largely comparable. Critically, there was a large reduction in sensitivity for symptoms reported at anytime versus symptoms restricted to 3-14 months (1 year) before diagnosis. Sensitivity was reduced to just 27% for women aged ≥50 on questionnaire data. Again, this highlights the importance of accounting for timing of symptoms when developing a symptoms index.

The sensitivity and specificity of the secondary test is critical for any potential symptoms-based diagnostic tool. Currently, this would be a CA125 test and/or a transvaginal ultrasound (TVS), however in the future there may be a better test. Recent data suggest that the specificity of TVS would be too low for this purpose.[63] Perhaps a better approach would be to use both symptoms and CA125 as a first line test. If CA125 results are intermediate, a follow up test with TVS and/or repeat CA125 could be done, after 12 weeks. The risk of Ovarian Cancer (ROC) algorithm that is employed in UKCTOCS could also be tied in with the serial CA125 testing. Obviously the acceptability and feasibility of such a proposal would need to be fully assessed. Results from an unpublished study indicate that much work would need to be done to get GPs on board with such an approach (Megan Goldsmith *et al.*, Primary Care Group Oxford, unpublished data).

Potential for 'Targeted Screening'

'Targeted screening' would be more akin to routine referral in that the index of suspicion would be substantially less than that required for rapid referral (due to the low predictive value of symptoms). For most symptoms, the duration would need to be long enough for women to consider making a GP appointment (e.g. loss of appetite for one day is unlikely prompt health-seeking behaviour). Once the woman has presented to the GP, the decision as to whether testing would be offered immediately or whether it would be more sensible to 'watch and wait' would probably be at GP discretion. Any 'targeted screening' programme would probably require 1-2 weeks to acquire the blood sample and process results. After allowing for these factors, one would need to allow a minimum of 2-3 months from symptom onset to diagnosis for 'targeted screening'.

In order for 'targeted screening' to be useful, there are several important factors to consider. Firstly, if population screening is unable to offer any mortality benefit, it is extremely unlikely (if not impossible) that symptoms will have an impact on mortality. The efficacy of population screening for ovarian cancer is still unknown (results are awaited from the UKCTOCS study due in 2015). Currently, population ovarian screening in postmenopausal women is thought to provide about 1.9 years of mean

lead time.[146] Results from the present study indicate that the longest (diagnostic) lead time given by symptoms is 14 months (or 12 months actual lead time), but only in a minority of women. Data from Index 2 suggest that 45%-74% of women would have a diagnostic lead time of at least 3 months. Secondly, even if population screening is shown to favourably impact survival, the amount of lead time required to achieve this remains unknown. It may be that less than 1.9 years is required to see a stage shift. If the requisite length of time was say, 6-9 months, there is a possibility that 'targeted screening' would still have a competitive impact. Also, some of the 1.9 years might be due to 'overdiagnosis' of indolent cancers by population screening. In addition, if the time to get from a first-line test to diagnosis (in population screening) takes several months, the estimated lead time would be shorter than 1.9 years. Thirdly, the proportion of women with symptoms that would have a positive screening test is unknown (assuming screening test is CA125 and/or TVS). Under the present circumstances, if women with symptoms do not test positively with CA125 and/or TVS, 'targeted screening' would be futile. Currently, the relationship between CA125 levels and symptoms is poorly defined, and marked variability was found in a small set of case studies.[147] Finally, if the cost-benefit ratio was more favourable for 'targeted screening' than population screening, 'targeted screening' may be the more viable option.

Comparison of the Data Sources

At the study outset, the duration of telephone interviews, symptom data quality and response-rate to GP note request and interview were unknown. In addition, this study aimed to collect a large volume of complex data from a difficult patient group. Thus, it was anticipated that several aspects would change as the study evolved. The decision as to what proportion of each of the 3 data sources would be obtained for each subject was altered as the study progressed based on feasibility, experience and response-rates.

Perceived benefits of using three modalities to collect the same data were as follows. The main advantage of the questionnaire was that it allowed for baseline symptom data to be obtained rapidly from all women in the study. This was prudent given the difficulties in studying acutely unwell women with high mortality. It was anticipated that symptom data would be complex to collect and would require clarification. Telephone interviews provided an opportunity to explore symptoms in greater detail by probing for additional information and to detect subtleties which were likely to be excluded from a paper questionnaire. This was important since a previous study had shown that symptom reporting can vary depending on semantic interpretation and the women's

perception of symptom aetiology.[94, 99] Interviews also have the advantage of having a motivating effect on the respondent, by providing full, clear definitions and provide the opportunity for probing ambiguous responses, or querying inconsistent answers. GP notes have the crucial advantage of providing objective data for symptoms and symptom lead time without recall bias. Of the three data sources, GP notes were viewed as providing the most robust data on lead time. As such, questionnaire and interview data are generally viewed as being more sensitive but less specific than GP note data.

As outlined in Chapter 2, there are various types of bias and errors inherently associated with each data source. For self-reported data the main types are recall bias and recall error. For GP note data, these are recording bias and recording error. Differences between the data sources were not formally analysed, however there were some clear distinctions.

Self-reported data were fairly comparable in terms of sensitivity for symptoms that were specified on the checklist, however spontaneously reported symptoms were volunteered more readily on interview. This could be due to the questionnaire format which may have been inadequately designed to capture such complex data, especially given that it was nested within a larger questionnaire with other demanding questions (e.g. “what was your skirt size in your early twenties?”). Also, asking women about specific symptoms is different from asking women to spontaneously report symptoms. This was exemplified on interview by the observation that only women who were specifically asked about night sweats reported them.

Symptom reporting was higher on interview than questionnaire for both cases and controls. Also, most women who reported having no symptoms on questionnaire, subsequently reported symptoms on interview or in GP notes. Recall bias (on interview) is an unlikely explanation since symptoms were often recorded in GP notes. These discrepancies may reflect semantic issues or recall error. They could be related to women not perceiving the symptoms as being associated with cancer, as highlighted by previous studies.[17, 107] During the interview, women were reminded that symptoms should be reported regardless of any perceived causes, whereas for questionnaire it is unclear how frequently this was done, if at all. Perhaps the timing of administration played a role. Cases were recruited at a stressful time when there would have been a constant barrage of forms to sign and motivation may have been low. Likewise, controls were recruited in busy clinics and may have rushed to complete their questionnaires. In contrast, interviews were performed in a quieter environment with fewer distractions (subjects were usually at home) and there was an opportunity to

probe and discuss symptoms more freely. On interview, loss of appetite was asked as 'loss of appetite or feeling full quickly'. These two symptoms are distinct and should not have been combined, but unfortunately this error was only noticed later on in the study. Nonetheless, this oversight did not appear to have any major effect on the data. The magnitude of difference in incidence for loss of appetite between questionnaire and interview was comparable to those observed for other symptoms.

Overall, controls reported a relatively constant rate of new symptoms. However, there was some tendency for increased symptom reporting closer to the reference date (see Figure 4-3). As expected, this was not observed in the GP note data. A possible explanation is 'forward telescoping', a phenomenon in which subjects have no clear date or event to bind their recall.[148] As a consequence, remote events (outside of the reference period) tend to be brought forward in time and are reported as occurring more recently. The risk of telescoping is known to increase in proportion to the length of time subjects are asked to recall over[116] and 12 months is a fairly long period of time.

Another disparity between questionnaire and interview results was for severity rating. On interview, symptoms were generally reported as less severe. Perhaps, asking details for when the symptom first started may have been too difficult or detailed to extract on a questionnaire format. However, this trend was not observed for frequency data.

One of the largest differences in symptom reporting was observed in control women. Controls reported considerably more symptoms on interview than questionnaire (see Figure 4-7 and Figure 4-8). This could be due to under-reporting on the questionnaire or over-reporting on interview. Alternatively, this may be explained by qualitative symptom properties.

Although this is an epidemiological study focussed on quantitative endpoints, there were various qualitative issues that were highlighted via the interview process. These issues shed light on some of the difficulties in obtaining symptom data such as the inability of the chosen symptom measures to accommodate for the variable nature of the symptoms investigated, and may in part explain some of the missing data. Specifically, when symptoms occur in extremely variable patterns it is confusing to try and assign frequency in 'days per month'. For example, a woman may have an acute episode of abdominal pain that lasts for one day, followed by several pain-free months, and then have the pain again lasting for several weeks. In these situations the symptom onset date and ongoing status also become confusing to identify.

Symptoms that were activity-dependent were also problematic as they vary with the amount or type of activity. This often occurred with fatigue, indigestion and back pain. Hence, onset dates, frequency and severity are difficult to assign for activity-dependent symptoms. Indigestion, bloating, fatigue, constipation, back pain and urinary frequency/urgency were reported in particularly higher proportions during interview, especially for controls. Anecdotally, this seemed to be due to a reluctance (cases and controls) to report symptoms that were mild, transient, activity-dependent, longstanding or perceived as being 'normal' for their age. All of these symptoms were recorded on interview but may have been omitted from the questionnaire since women often did not think they were relevant when any of the aforementioned vague features were present.

Some women (particularly controls) were reluctant to report symptoms on interview for which they had already assigned a cause. One control attributed urinary frequency and fatigue to ageing and said "all my friends have the same changes". Another control reported fatigue then added "but it's normal for my age". Likewise, there seemed to be an unwillingness to report symptoms that were diet-related since onset was typically sporadic and causes were 'known'. These findings echoed the reports of Bankhead *et al.*[99] who found that mislabelling and misinterpretation of symptoms can affect symptom measures.

Related to this, is the issue with symptom definition and how this can vary between clinical definitions, what is normal for the patient, and what is perceived by the patient. For example, most women on interview considered constipation to mean having less than daily bowel motions. The term 'increase in abdominal size' was chosen over 'abdominal swelling' and 'abdominal distension' since previous research has shown that most women have a poor understanding of the latter terms.[94] However, during interview it became apparent that 'increase in abdominal size' was reported if a woman felt that she had gained weight, particularly with controls. Abdominal distension in women with ovarian cancer is often associated with weight gain due to collection of ascitic fluid; however this is clearly clinically distinct from weight gain due to increased body mass. As a corollary, increased abdominal size may be over-reported by controls in self-reported data, and is unlikely to represent true abdominal distension. Indeed, 50% (17/34) of controls who reported increased abdominal size also reported weight gain on interview, compared to just 10% (11/115) of cases. This is also reflected by the low incidence of increased abdominal size in GP notes. All of these factors can contribute to variable responses in self-reported symptom data and should be taken into consideration in any further studies involving collecting symptom data directly from women.

GP note data are distinctive in several important ways. Recording of symptoms in GP notes are diagnosis-driven[24] and therefore may not accurately reflect the full spectrum of symptoms experienced by women (reflecting recording bias). As such, symptom sensitivity was expected to be lower compared with self-reported data. This was typically around the magnitude of 10-40%. Despite this, GP notes are still considered to be the 'hardest' and least (recall) biased data. Another advantage of the GP notes collected for the study was that records were from various geographical regions. This eliminated any potential recording bias arising from using any one particular surgery or region. In addition to recording bias, women have been found to selectively report symptoms to GPs,[94] which may also contribute to lower symptom sensitivity in GP notes. Finally, GP note data do not contain accurate symptom duration data, therefore some of the symptoms included as incident symptoms in our study may have in fact been prevalent or longstanding.

Overall, the symptoms recorded in the GP notes for cases were encouragingly similar to those recorded on self-reported data (see Table 4-8, Table 4-9, Table 4-10). Any disparities observed were mainly for symptoms that were not specifically listed on the questionnaire/interview checklist (such as vaginal discharge, other urinary symptoms etc.). One of the key expected differences between GP notes and self-reported data was that the latter would detect symptoms that are not reported to the GP. The four most sensitive symptoms according to self-reported data were increased abdominal size, pelvic/abdominal pain or discomfort, abdominal bloating, and fatigue. Conversely, on GP notes the latter two were replaced by 'urinary other' and nausea/vomiting, both of which were reported in negligible amounts on self-reported sources. However, nausea or vomiting ceased to feature in the top four most common symptoms once symptoms were restricted to those present over 3-14 months before diagnosis. This incongruence could reflect real differences in what symptoms women report to their GP.

Alternatively, some of the symptoms recorded in the GP notes may reflect signs as opposed to symptoms for which a woman sought medical advice. For example, a lump in the abdomen that was palpated by the GP on examination but was not noticed by the woman. There was no real desire to distinguish between the two in the present study, however this could have contributed to differences in symptom frequency between GP notes and self-reported data.

Urinary symptoms featured consistently more prominently in cases and controls, on GP notes in comparison to self-reported data. 'Urinary other' comprised a group of symptoms that were not necessarily related (i.e. dysuria, incontinence, haematuria,

retention, poor stream, changes in urine colour or smell). The comparatively lower reporting of 'urinary other' symptoms on self-reported data is probably because they were not specifically asked for in the survey. However, this does not explain the case-control differences observed. Moreover, urinary frequency or urgency were specifically asked for on self-reported data, and they were also recorded noticeably more often on GP notes. Further research is required to elucidate the reasons for this discrepancy in urinary symptom reporting between data sources.

Severity & Frequency

Severity and frequency are not necessarily independent. One would expect that as frequency increases, perceived severity also increases. Information regarding severity and frequency of symptoms may only be useful if they can be used as a discriminatory tool in a symptoms index. Both parameters are expected to vary with disease progression and possibly activity, so ratings were requested for when the symptom first started. The two studies which have examined severity and frequency did not set any rating criteria.[69, 75] In the absence of any instructions for the timing of when severity and frequency should be rated, a reasonable assumption would be that ratings are given arbitrarily or with a bias towards the most extreme.

The trend towards lower severity ratings on interview could indicate that women did not follow or read this timing instruction (i.e. severity/frequency when symptoms first started) on questionnaire. Women who reported severe and more frequent symptoms were more likely to be cases, as were women who had symptoms that worsened over time. In contrast, women who reported symptoms that stayed the same over time were more likely to be controls.

Further analyses are planned to address severity and frequency more formally. Given the propensity for severity and frequency to change over time, there may be potential to incorporate increases in severity and/or frequency (i.e. worsening) into a decision rule. For example, if CA125 results are subthreshold one might decide to repeat CA125 in 12 weeks if symptoms are worsening, conversely a 'watch and wait' approach may be more appropriate if symptoms are the same or better.

Missing Data

Missing data were probably attributable to several factors. Symptoms can be difficult to accurately recall, especially if they are acute and sporadic. Women may have been simply unable to recall start dates or symptom characteristics (especially if they started a long time ago). As mentioned above, when symptoms occur in variable patterns (e.g. once every 3-5 months for 1 week) it can be difficult to identify when the symptom

started and to assign frequency. Severity is also challenging to assess if there is variation over time. Control questionnaire data quality appeared to be lower than that of the cases and contained more missing data. This may have been attributable to a lack of motivation, since one would expect cases to have more of a vested interest in contributing to the research and having their symptoms analysed.

Study Advantages

Previous studies have used *either* medical records *or* self-reported (symptom checklists or open-ended questions) to obtain symptom data. This is the first study we are aware of that has used three different sources to collect the same data. The advantage of this approach was that the low sensitivity of medical records could be balanced by the low specificity of self-reported data. In order to truly maximise on this, a combined analysis is required. Given that the data sources vary with regard to several different aspects, a complex algorithm needs to be devised, and this is currently planned for a future analysis.

One of the key advantages to this study was the novel approach of using continuation odds ratios to examine symptom incidence over the period leading up to diagnosis. Most studies have collected data retrospectively, and in their interpretation of results they have failed to account for the way in which symptoms develop when they occur prospectively in real time. As demonstrated in this study, much of what is observed at diagnosis (in terms of symptom severity, frequency and incidence) may dramatically differ in the preceding months.

Also, the systematic accrual of newly diagnosed cases in participating hospitals meant that women consented to participate without regardless of symptom history. As a corollary, we expect less selection bias. Likewise, controls were consecutively recruited at attendance for annual screening, and were not selected based on symptoms.

Study Limitations

Possible limitations of this study include that the control group were derived from a screening study. Screening studies are known to favour subjects who tend to be healthier and of higher sociodemographic background than the general population. Furthermore, UKCTOCS-enrolled women who have significant symptoms or who are unwell may further self-select by choosing not to attend screening. This would not affect the main UKOPS study biological endpoints, however this could affect symptom data collection. The potential for self-selection bias is one of the reasons why the GP workload study was undertaken (Part III). The analysis of GP notes from an

unselected population should highlight any effects of self-selection on the symptom specificity observed in the present study. There may have been some bias against women with particularly advanced disease since these women may have felt too unwell to participate. Indeed, the study had quite a large proportion of women with early stage disease (40%). Unfortunately, information on what proportion of women with ovarian cancer were not approached or refused to participate at each centre was not available.

The fact that we treated symptoms with missing onset dates as if they were longstanding symptoms would have underestimated symptom sensitivity for cases, but it was felt that this conservative approach was better.

Another potential drawback is that cases and controls came from slightly different age groups and have different menopausal status. Cases were a mix of pre-, peri- and post- menopausal women aged 45 and above, whereas controls were all post-menopausal, aged 50-74 at UKCTOCS entry (i.e. a maximum of 81 years in our study since UKCTOCS started 7 years before UKOPS). However, 88% of cases were aged 50 and above. Possible implications of this are that symptom sensitivity and odds ratios may be inaccurate for symptoms that are common (or rare) in the peri-menopause or due to menses. Some of the main differences would be expected in vaginal bleeding symptoms, however these are not highly reported by women with ovarian cancer. In addition, the proportion of women with IVB or PMB in our study was very similar to those reported by previous studies. Similarly, the inclusion of cases aged over 74 (despite having limited controls in this age group [16%]) in the final analysis was justified by the fact that their symptom profiles were comparable to those in the original target age group of 45-74.

The fact that all of the interviews and the data extraction from GP notes were performed by the same person who was not blinded to case-control status could have introduced some bias. However, all symptoms were extracted using a systematic coding frame and a standardised interview script were used to try to minimise any such bias. Equally, the questionnaire was administered by different nurses at each centre and each nurse may have influenced questionnaire completion differently, or offered varying levels of help. At some centres, the nurse always completed the questionnaire with the women. Certainly, there were less missing data from these centres. Also, in retrospect, the questionnaire design was too complicated for the context of this study. For example, symptom duration was meant to be captured by the question ‘how long have you had this?’ During the data cleaning process it was clear that many women had recorded this incorrectly since duration did not match up with the start dates reported and time of questionnaire completion. In the data cleaning process, duration

was recalculated based on the symptom onset date. Questions requiring the subject to perform arithmetic (such as duration) should have been avoided.

In development of the symptom indices, incidentally-diagnosed women were included since excluding them would unduly make the indices appear more favourable. Screen-detected women were also included since their self-reported symptom profiles appeared to be similar to women who were detected via symptoms. However, in retrospect it would have been more appropriate to exclude screen-detected women as any length of symptom duration would be censored by the screening.

A further possible criticism is that questionnaire was not validated, which gives rise to uncertainty over whether the questionnaire detected what it was designed to detect. However, the results indicated that there was much internal validity. The evidence for this included:

- No case-control differences were observed for any symptom beyond 15 months prior to reference date suggesting that recall bias (if present) is limited to the first 15 months).
- Control symptoms over time were uniform suggesting that there is no increase in recall error in controls (although minor increases were observed close to date of consent).
- Odds ratios close to diagnosis were extremely large (i.e. OR>20), showing that the questionnaire was able to identify features that are present in the weeks preceding diagnosis of ovarian cancer.
- Case-control comparisons for unrelated symptoms were similar suggesting that the survey was distinguishing ovarian cancer symptoms from other symptoms.
- Symptoms recorded in GP notes were similar to self-reported data also suggesting that the survey was detecting ovarian cancer symptoms.

Collectively, these provide confidence to assume that the findings were real, despite use of an unvalidated questionnaire.

Many assumptions had to be made when databasing and cleaning the data, hence databasing rules were devised to ensure consistency. In general, these rules favoured a conservative approach which underestimated case-control differences. For example, symptoms that had missing onset dates were treated as longstanding symptoms and therefore were not included in odds ratios and symptom sensitivity. Also, the sensitivity of 1 year's worth of symptom incidence was examined with the 3 months prior to diagnosis excluded. One of the potentially controversial databasing rules was to assume that symptoms were ongoing if ongoing status was missing. However, our

approach was not dependent on symptom duration, but rather on the first time the symptom appeared.

Due to poor recruitment, several women who did not meet the original eligibility criteria were included. A restricted analysis excluding these women was performed, leaving 155 cases and 244 controls (see 3.3.11 Data Analysis). Results did not imply that the inclusion of the "ineligible" women had any dramatic effects on the study outcomes. According to the questionnaire, crude ORs (95%CI) for abdominal/pelvic pain or discomfort 3-14 months before diagnosis were 12.8 (5.4, 30.2) in the restricted analysis versus 11.5 (5.4, 24.5) in the main analysis. On interview, abdominal/pelvic pain or discomfort at diagnosis was cited by 54% (95% CI: 41, 67) versus 63% (95% CI: 54, 71) of cases for the restricted and main analysis, respectively.

5.2 Conclusions

Until a viable screening tool is identified, symptoms remain the most common diagnostic route for ovarian cancer. The present study has shown that the diagnostic process could be initiated at least 3 months earlier than the current date of diagnosis, in 45% (GP notes) to 74% (interview) of cases. However, it is unrealistic to assume that diagnosis could be bought forward by the full amount of symptom lead time since testing, referral and histological confirmation (i.e. final diagnosis) requires several weeks. Given that between 22% (GP notes) to 26% (interview) of healthy women would also require testing in one year, it would not be reasonable to refer any woman with an Index symptom for urgent assessment. Nevertheless, after allowing 2 months for referral to diagnosis, 26% (GP notes) to 56% (interview) of cases would still have their actual diagnosis bought forward by at least 3 months. Since residual disease is a strong predictor of survival, this small window of opportunity could potentially have an impact.

Problem GP visits in cases start to increase about 6 months prior to diagnosis providing further evidence that there is potential to bring forward diagnosis. Symptoms associated with ovarian cancer appear to develop within 15 months of diagnosis. It may be important to exclude symptoms which started earlier than this in any future study of symptoms in women with ovarian cancer.

It is crucial to take the timing of symptoms into account when assessing any potential symptoms-based diagnostic tool or performing studies on symptoms. Symptom incidence for cases decreases with increasing time before diagnosis. Even from as little as 3 months prior to diagnosis, large reductions in symptom incidence were

observed. Other symptom characteristics such as frequency and severity are also likely to increase over time in cases.

Methods of symptom data elicitation are important determinants of symptom reporting. Further analysis to combine the data from each of the three data sources will help to elucidate these differences.

Abdominal and gastrointestinal symptoms were commonly cited by cases which may explain the convoluted diagnostic routes observed, with less than 50% of women being initially referred to rapid referral gynaecological-oncology. However, referral type did not seem to translate into a delay in diagnosis.

Crucially, a worrying proportion (approximately 25%) of symptomatic women had patient and provider delays of longer than 4 months (Table 4-17). There are many unanswered questions relating to increasing symptoms awareness and its repercussions (negative and positive). Ideally, further research should be carried out before promoting symptoms awareness in women or GPs.

The total and provider delays identified in this study were slightly longer than most that have been reported in the literature, despite using methods that would tend to favour shorter delays. This could indicate that delays in the UK are longer, which would be consistent with what some researchers have suspected to be the cause of the poorer survival rates observed in comparison to other European countries. However, given the variation in delays with each data source, and in the literature, further research is needed to confirm this.

PART III: PUBLIC HEALTH IMPLICATIONS

Part III will explore the potential impact of ‘targeted’ ovarian cancer screening on public health resources. This includes a pilot study using electronic patient records to estimate the GP workload associated with offering ‘targeted screening’, and a general discussion on the findings of this thesis, ending with suggested future directions.

6 CHAPTER 6: Primary Care Workload: Implications of 'Targeted' Ovarian Cancer Screening

6.1 Introduction

This was a pilot study designed to provide a crude extrapolation of the general practitioner (GP) workload associated with offering 'targeted' ovarian cancer screening. Consideration of workload impact is crucial given the pressure on primary care clinics.[149] In particular, we wished to estimate the proportion of women who would be offered 'targeted screening' over the course of 1 year. In order to do this, the proportion of women aged 45-74 presenting to primary care in one week who reported vague symptoms possibly related to ovarian cancer was used to estimate the proportion who would present in one year. This study also provided the opportunity to gain perspective on the case-control data from the previous section, given that the controls were from a screening study, and a 'healthy volunteer effect' (i.e. women who volunteer for screening studies tend to be healthier on average with a better chance of survival) was anticipated.

The most sensible approach for collecting the requisite data was via an 'audit-type' study examining GP medical records. Most GP practices in the UK have switched from handwritten to electronic patient records (EPR). By using practices with EPR systems in place, it was possible to obtain large amounts of data rapidly with a simple download. Anonymous individual-level records of women aged 45-74 consulting during a single week were extracted from two urban and two non-urban GP practices based in London, Newbury, Bracknell and Wokingham.

6.2 Aims

6.2.1 Primary Objective

- To determine the proportion of women aged 45-74 and 55-74 years, presenting to primary care clinics with vague symptoms common to ovarian cancer in one week.

6.2.2 Secondary Objectives

- To identify the proportion of symptom(s) that:
    - Can be attributed *a priori* to any particular disease or condition.
    - Cannot be attributed *a priori* to any particular disease or condition

- To estimate the age-specific incidence of all vague ovarian cancer symptoms and of unattributable vague ovarian cancer symptoms

6.3 Methods

6.3.1 Study Population

Inclusion Criteria

- Women aged 45-74 years presenting to primary care during the study period (one week)
- Permanently registered patients

Exclusion Criteria

- History of ovarian cancer
- Women receiving active treatment for other cancers (except for hormonal treatment only)

Restriction to postmenopausal women was desirable; however menopausal status is not routinely documented in GP notes so this was not possible. Ascertainment of menopausal status requires specific information such as length of time since last menstrual period, luteinising hormone levels, length of hormone replacement therapy (HRT) treatment and previous hysterectomy. Since the average age of menopause is 50;[150] the range of 45-74 years was applied. The upper age limit was in put place because it is unlikely that 'screening' would be offered to women older than 74. Only permanently registered patients were included since testing is unlikely to be offered to women who are difficult to follow-up.

Initially, history of bilateral oophorectomy was also an exclusion criterion however this had to be dropped due to its poor and often contradictory documentation in the notes (e.g. some women had a unilateral oophorectomy recorded years after the date of a bilateral oophorectomy).

6.3.2 Selection of GP Practices

GP surgeries were selected based on list size, EPR system and location. Each GP surgery had to meet the following criteria:

- List size of at least 5500
- EMIS LV (see details below) EPR system in place for more than one year with no handwritten notes.
- At least one female GP

- No unusual demographic composition

The threshold for list size was based on the average of 5891 patients for England set by the 2004 General Medical Services contract.[151] Although several EPR systems are available in the UK; EMIS (Egton Medical Information Systems) LV was chosen, primarily because of its use in the Clinical Effectiveness Group (CEG) at Queen Mary University (QMUL). The CEG provided access and support which were critical for developing the search strategy. In addition, EMIS is the most common EPR in primary care with around 55% of GPs using it daily.[152] There are three different versions available including EMIS Primary Care System (PCS), EMIS Limited Version (LV) and EMIS Web. The requirement for at least one year of EMIS-use was applied in attempt to gain high quality data since EPR data quality is likely to vary with the user-friendliness of each software package and the skill of individual users. The exclusion of handwritten notes was to obviate the need to decipher handwriting and manually anonymise data. At least one female GP was required at each practice in case women are reluctant to report gynaecological symptoms to male GPs, or pelvic vaginal examinations are less likely to be performed by male GPs. Surgeries with unusual demographic compositions (such as a disproportionate number of young women) were excluded in order to gain a representative sample.

6.3.3 Study Design

A cross-sectional analysis of GP notes was performed at the following practices:

- Chrisp Street Practice – an urban surgery in London (week examined 29.10.07 to 05.11.07)
- Northcroft Surgery – an urban surgery in Newbury (week examined 15.05.08 to 22.05.08)
- Woosehill Surgery – a non-urban surgery in Wokingham (week examined 15.05.08 to 22.05.08)
- Boundary House Surgery – a non-urban surgery in Bracknell (week examined 25.06.08 – 07.07.08)

As mentioned above, an electronic search was performed at each practice. The search strategy was developed and tested with dummy patient data using the 'EMIS Aspects' search function at the CEG. Piloting of the search strategy and the EMIS anonymiser (an executable program specially written at the Wolfson Institute to anonymise the data) was performed at Chrisp Street practice. After the first successful download, the data were cleaned and extracted, and a brief analysis was performed to

ensure that the main study endpoints were achievable before the remaining downloads were performed.

The week examined preceded the date of data download by one day in order to allow time for data entry for visits outside of the surgery. This produced a more complete dataset. There was no requirement for the download to occur during the same week in each practice. The main steps in the data collection process involved:

1) Defining the search population – i.e. women aged 45 to 74 who visited between the census dates.

2) Using the EMIS patient summary function to create a text file containing the relevant (non-anonymous) information.

3) Saving the patient summary file to the computer desktop.

4) Running the EMIS anonymiser – this produced three comma separated (csv) files which contained the anonymised data linked by unique patient identification numbers.

The anonymous csv data files were then taken from the surgery to be cleaned and analysed.

The patient summary function produces standardised content with options to add extra information. For the present study, additional data on the last 10 'consultations' (including visits in the census period), significant past and active events/diseases (i.e. medical history), and ethnicity were requested. 'Consultations' in EMIS LV includes events that are not actual consultations such as letters (incoming and outgoing), and lab results. Therefore, each dataset required cleaning to obtain the true proportion of women with a consultation in the census week. Consultation data contained details of visit dates, symptoms (free text and coded), tests, prescriptions, and referrals. By looking at the last 10 'consultations' before the date of download, information on symptom development was gleaned.

Consultations for this study were defined as surgery visits, home visits and telephone calls with a GP or nurse. These included after hours visits. Telephone calls were only included if symptoms were mentioned or medical advice was discussed. Certain visits, such as visits with a phlebotomist, were excluded. Visits for smears and inoculations were included (even though the reason for visit was not an active health problem) on the premise that women may still report symptoms. Eligibility criteria for consultations are listed below in Table 6-1. Women had to have at least one eligible consultation in

the study period to be included in the analysis. If more than one consultation per day occurred, each was counted individually as a single consultation. However, most data are presented as the number of women who consulted during the period, not the number of consultations.

Table 6-1 Consultations versus Non-Consultations

Consultation	Non-Consultation
Any surgery visit, home visit or telephone call with a GP or nurse (telephone calls must have involved medical advice or symptom reporting)	Blood tests with phlebotomist Visits with a counsellor Smoking clinics

The following data were extracted:

- Consultation dates
- Location of visits (e.g. home visit or GP surgery)
- GP or nurse consulted
- Symptoms recorded
- Co-morbidities
- Age of women (years)
- Tests or referrals ordered (including reason for referral if data permitted)

Other general data taken from each surgery included:

- Details of total list size age and sex breakdown
- Proportion of women aged 45-74 and 55-74 in the practice
- Number of GPs
- Full-time or part-time work status of each GP
- Total number of women registered with the practice

All consultations were manually examined for symptoms since the majority of symptoms were recorded as free text. Extracted symptoms were binary coded as relevant or not relevant to ovarian cancer according to a pre-defined list of 'key' symptoms (see Table 6-2). This 'key' symptom list was based on symptoms commonly reported in the literature by women with ovarian cancer.[6-8, 12, 16, 17, 67, 69, 75, 89] Frequency, urgency and retention were the only urinary symptoms considered to be salient. The terms 'UTI', 'cystitis' and 'urinary symptoms' frequently appeared in the notes without any specific symptom details. These were included as 'key' symptoms since urgency or frequency were likely to be present. Although back pain has frequently been cited in other studies it was excluded in the present study since no case-control differences were observed in the study in Part II. Likewise, back pain was not found to be a discriminatory factor in Goff's case-control study.[75] Irregular menstrual bleeding was also excluded since 'targeted screening' would most likely be limited to

postmenopausal women. As in the case-control study (Part II), 'key' symptoms were open to interpretation. The symptom definitions devised in the Part II were applied to the data (e.g. 'feels weak' was considered to be fatigue). The full list of symptoms included from each practice is found in Appendix X.

Table 6-2 'Key' Ovarian Cancer Symptoms

Pelvic or abdominal pain or discomfort
Abdominal bloating
Increased abdominal size/abdominal distension
Abdominal lump
Indigestion or heartburn
Constipation
Diarrhoea
Nausea or vomiting
Other GI symptoms (e.g. wind, change in bowel habit)
Loss of appetite
Difficulty eating
Weight loss
Fatigue
Urinary frequency/urgency/retention (including terms 'UTI',' cystitis')
Postmenopausal bleeding
Post-coital bleeding
Pain during intercourse
Vaginal discharge

In attempt to refine the estimation of *true* potential workload; 'key' symptoms were examined in more detail. Information from previous visits and medical history was used to binary code each 'key' symptom according to the following:

1) 'New' – in the last 3 months

2) 'Worsened' (based on information in free text)

3) 'Attributable'

These categories allowed for formulation of different workload estimates for different clinical scenarios. All 'key' symptoms were included regardless of whether they were mentioned at a nurse or GP consultation.

The 3 month threshold for the definition of a 'new' symptom was based on the median time between census week visit and earliest real consultation obtained by downloading the last 10 before download date (as per search strategy). The lowest median from each of the four practices was 3.5 months (IQR 1.4-7.8). Thus, 3 months was felt to be a reasonable time period over which to assume consultation details would be available for most women. Any symptom that was not recorded in the 3 months prior to reference date was considered to be 'new'. The labelling of 'new' took in to account symptom groupings as defined in Part II (e.g. loose bowels and diarrhoea were both

considered to be diarrhoea). It was not possible to know about symptoms that were present but not reported or recorded, so this classification served only as a rough approximation. Notably, if a woman presented for the first time with a symptom that started 6 months ago, this would still be classified as 'new'. Also, if a woman consulted for the same symptom twice within 7 days, but the first visit was *before* the census week and the second visit was *in* the census week, the latter would be classified as 'old' even though in clinical practice it would still be considered 'new'. This because our interest was in reports of symptoms that might have led to a 'screening test' if 'targeted' screening is introduced. Clearly, if testing was performed the first time the symptom was reported, it would not be repeated 3 days later.

An attempt was made to separate symptoms that were attributable *a priori* to pre-existing or new disease (e.g. acute lower back pain after injury) from symptoms that could not be attributed to any particular disease or condition. This was done because testing is unlikely to be offered to women with longstanding symptoms from pre-existing conditions (assuming no worsening or change). Equally, GPs may not wish to offer testing when other causes are more likely (e.g. fatigue associated with cold or flu symptoms). The decision over which symptoms were attributable was arbitrary as there are many possible causes for any given symptom and many possible symptoms for any given disease. For example, chronic gastritis alone could elicit dyspepsia, epigastric pain, bloating, nausea, flatulence and loss of appetite, although one may not wish to consider the entire list of potentially related symptoms to be attributable. Medical history and consultation data (from previous and census week visits) were used in the classification process. In general, the designation of 'attributable' status was conservative and in reality a GP may decide not to offer testing in many more scenarios than those selected for this study. Mainly sensible and straightforward causes were permitted such as fatigue after major surgery, nausea after whiplash injury, vomiting with alcohol withdrawal. The most common reasons for 'attributable' labelling included:

- GI symptoms and history of IBS or diverticulitis/diverticular disease
- Indigestion/acid reflux/epigastric pain and history of these, or, gastritis or recently documented *Helicobacter pylori* infection
- Fatigue and symptoms typical of viral infections, common cold or chest infections
- Diarrhoea/nausea/vomiting and viral symptoms (e.g. muscle ache, runny nose etc.)

Since this was a pilot study these relatively loose definitions of 'attributable' were felt to be acceptable.

'Key' symptoms were coded into major symptom groups for each practice including abdominal, gastrointestinal (GI), gynaecological, urinary and general (details listed in Appendix X). This provided an overall impression of what types of symptoms women in the target age group visit primary care with. Coding was kept closely in line with the case-control study rules from Part II in this thesis. Symptoms listed in the US ovarian cancer consensus statement[153] were flagged, however the criterion of daily symptoms for more than a few weeks was not possible to apply with the limited information available.

6.3.4 Sample Size

In the UK, women aged 45-74 have an average of 1600 consultations per GP per year.[154] This is equivalent to approximately 32 visits per GP per week. With four average sized GP practices a total of 15-20 GPs would be included in the study, therefore we expected to get around 550 visits made by women aged 45-74 in one week.

Based on the assumption that 550 visits were expected during the study, Table 6-3 shows the various 95% confidence intervals associated with different proportions of these visits being eligible for testing during one week (i.e. estimate of workload).

Table 6-3 Table of 95%CI associated with different proportions of visits that trigger 'targeted screening' for women aged 45-74 (based on 550 visits)

Proportion	95% CI
20%	16.7-23.6
10%	7.6-12.8
6%	4.2-8.3
2%	1.0-3.5

6.3.5 Analysis

Analysis of study data was performed using two age bands:

1) Women aged 45-54 years (mixture of pre-, peri- and post-menopausal)

2) Women aged 55-74 years (predominantly postmenopausal)

The smaller age band of 55-74 was intended to provide data that would be more representative of postmenopausal women. The two groups were compared in terms of symptom prevalence.

Symptoms were analysed on an individual and collective basis. The proportion of women aged 45-74 and 55-74 who presented during the census week with symptoms

was calculated with 95% confidence intervals for proportions. This was done according to different symptom categories including:

- Any ‘key’ symptom
- Any ‘new’ ‘key’ symptom
- Any ‘key’ symptom unattributable to other diseases
- Any ‘key’ symptoms for which testing would realistically be offered (as described above)
- ≥3 key symptoms
- ≥3 key symptoms with at least one ‘new’ or ‘worsened’
- Any one of each grouped symptom (i.e. abdominal, GI, gynaecological, urinary, general)
- Any consensus symptom

If the same ‘key’ symptom was reported more than once in the census week by the same woman, it was counted as one ‘key’ symptom.

In reality, ‘targeted screening’ would not be performed in all women presenting with ‘key’ symptoms, such as women who are already awaiting ultrasound or who have recently been tested. Hence, the consultation data for each woman with a ‘key’ symptom were closely examined and women who would realistically be offered testing were identified. The following women were excluded:

- Women who already met the criteria for testing in a visit in the 3 months prior to census week – one would not offer testing twice within this period.
- Women who had a pelvic ultrasound scan in the past 6 months (since second level screening is likely to include pelvic ultrasound).
- Women who already had a rapid referral to gynaecological oncology in the past 6 months.
- Women in palliative care (testing would not be offered to women with a short life expectancy).
- Women referred, or undergoing investigations for, a possible malignancy (including ultrasound).
- Women who only mentioned ‘key’ symptoms that had already resolved.
- Women who only had ‘key’ symptoms that were classified as ‘attributable’.

In the consideration of ≥3 key symptoms, calculations were performed separately for ≥3 reported in the census week (i.e. calculated over the entire week’s worth of visits) and ≥3 reported in one census day. However, these were identical so results are reported as ≥3 key symptoms in one week.

In order to make the data more comparable with the case-control study, the symptom index (Index 2) developed in Part II, was applied to the data. Index symptoms included:

- Increased abdominal size
- Lump in abdomen
- Loss of appetite
- Pelvic/abdominal pain or discomfort
- Constipation
- Fatigue
- Bloating
- Nausea or vomiting
- Weight loss.

Any women with at least one index symptom during the reference week were identified. A second definition of 'new' versus 'old' was devised. Symptoms were considered to be 'new' if no other index symptoms appeared in the previous 10 'consultations'. Symptoms recorded more than one year before the reference week were excluded from this criterion.

A crude extrapolation to estimate the proportion of women in each age group and each symptom category that would be offered testing over one year was performed using the equation:

$$1 - (1 - p)^{52}$$

if *p* is the proportion of women to be offered testing in a single week.

6.4 Results

A total of seven practices were approached and five agreed to participate. One urban London practice was excluded as the list size was too small. The original plan was to collect data from one urban and one non-urban practice from London and Oxfordshire each. Unfortunately, only one suitable practice from London agreed to participate in time to take part (Chrisp Street Practice). Thus, an extra surgery was obtained from Oxfordshire. Only anonymised data were used in the current study, however because the data collection process was novel, initially there was some confusion over which approvals were required. Advice regarding this was sought from the Patient Information Advisory Group (PIAG), Barts and the London Research & Development Office and the East London and the City Ethics Committee. PIAG approval was

deemed unnecessary so long as the researcher did not see any identifiable data and the download was performed by a member of the surgery staff. Research Ethics Committee approval was also not obligatory. The requirement for Primary Care Trust (PCT) approval was less clear and the two PCTs that were approached responded differently. The North East London Consortium for Research and Development (NELCRAD) could not decide whether the proposed work constituted a formal research study or an audit. Ultimately, NELCRAD suggested that the study could be taken on as a GP-run study in which case PCT approval would not be required. This is how the study was run at Chrisp Street Practice. For the three remaining GP practices examined, PCT approval was provided by the Oxford Primary Care Trust.

Table 6-5 contains GP surgery details and the number of women consulting during the census week for each individual practice and overall. The combined list size was 38,921, including 5737 women aged 45-74 and 3031 women aged 55-74. A total of 479 women aged 45-74 and 287 women aged 55-74, consulted during one week. Of the women aged 45-74 who consulted during the week, two had a previous history of ovarian cancer and two were receiving treatment for an active cancer (see Table 6-6). As mentioned in 6.3.1 Study Population, these women were excluded from the analysis since testing would not be offered to such women in clinical practice. This left a total of 475 women aged 45-74 and 285 women aged 55-74 for analysis.

Females comprised approximately 50% of the practice population for all surgeries. The proportion of registered females that were aged 45-74 and 55-74 at each practice was also comparable (combined results 15% and 8%, respectively). In contrast, the proportion of women in each age group who consulted during the week varied between practices. At Chrisp Street (14%) and Northcroft (11%) practices, the proportion of women consulting was more than twice that of Boundary House (5%) and Woosehill surgery (5%). However, of the women consulting in each age group, the proportion that reported symptoms was similar for all four practices. Likewise, the proportion of consulting women who reported 'key' symptoms was also comparable. Interestingly, symptom reporting did not differ between the smaller predominantly postmenopausal group and the larger mixed peri- and post- menopausal group.

Overall, 8% of all women aged 45-74 consulted during the census week. Of the consulters, 62% (293/475) reported at least one symptom and 17% (80/475) had at least one 'key' symptom. Corresponding figures for women aged 55-74 were similar with 9% presenting in the census week, 57% with any symptom and 15% with any 'key' symptom.

Approximately 60% of consultations made by women in the broader target age group (45-74) were with the GP. Most of the consultations took place in person at the surgery (see Table 6-4). A similar rate of telephone consultations took place at each surgery except for Northcroft surgery, which had noticeably fewer. Some women consulted more than once in the week and although 475 women consulted, there were 566 consultations.

Symptom reporting was higher at GP versus nurse consultations. Of the women aged 45-74 who reported 'key' symptoms, 86% (69/80) were during a GP visit and 14% (11/80) were during a nurse visit.

Table 6-4 Consultation Location

	Number of Visits (%)				
	Chrisp St	Boundary	Northcroft	Woosehill	**Total**
Women aged 45-74					
Surgery	160 (77%)	60 (83%)	170 (93%)	86 (83%)	**476**
Telephone	43 (21%)	12 (17%)	6 (3%)	17 (16%)	**78**
Home	4 (2%)	-	7 (4%)	1 (1%)	**12**
				TOTAL	**566**
Women aged 55-74					
Surgery	92 (77%)	37 (84%)	108 (93%)	49 (80%)	**286**
Telephone	24 (20%)	7 (16%)	4 (3%)	11 (18%)	**46**
Home	3 (3%)	-	4 (3%)	1 (2%)	**8**
				TOTAL	**340**

Note that women may have more than one consultation during the census week

Table 6-5 Details of GP Surgery & Women Aged 45-74 (55-74) Consulting During Census Week

	Chrisp St	Woosehill	Northcroft	Boundary	**TOTAL**
List size	11504	10652	8928	7837	**38921**
Total number females	5645 (49%)	5388 (51%)	4503 (50%)	4022 (51%)	**19558 (50%)**
Total number females aged 45-74	1210 (11%)	1809 (17%)	1457 (16%)	1261 (16%)	**5737 (15%)**
Total number females aged 55-74	658 (6%)	861 (8%)	877 (10%)	635 (8%)	**3031 (8%)**
GPs at practice					
Full-time	1	3	3	3	**10**
Part-time	11	2	2	1	**16**
Total	12	5	5	4	**26**
Women aged **45-74** consulting during week					
with GP	133	56	124	38	**351**
with Nurse	50	36	45	24	**155**
Total (GP or nurse)	168*	85	163**	59	**475**
Proportion consulting of all women aged 45-74	14%	5%	11%	5%	**8%**
Women with ≥1 symptom recorded	112/168 (67%)	50/85 (59%)	93/163 (57%)	38/59 (68%)	**293/475 (62%)**
Women with ≥1 'key' symptom	27/168 (16%)	14/85 (16%)	28/163 (17%)	11/59 (19%)	**80/475 (17%)**
Proportion of women in age group with ≥1 'key' symptom	27/1210 (2.2%)	14/1809 (0.8%)	28/1457 (1.9%)	11/1261 (0.9%)	**80/5737 (1.4%)**
Women aged **55-74** consulting during week					
with GP	73	36	77	22	**208**
with Nurse	31	17	30	16	**94**
Total (GP or nurse)	96	51	104*	35	**286**
Proportion consulting of all women aged 55-74	15%	6%	12%	6%	**9%**
Women with ≥1 symptom recorded	56/96 (58%)	31/51 (61%)	54/104 (52%)	21/35 (60%)	**162/286 (57%)**
Women with ≥1 'key' symptom	12/96 (13%)	8/61 (13%)	18/104 (17%)	5/35 (14%)	**43/286 (15%)**
Proportion of women in age group with≥1 'key' symptom	12/658 (1.8%)	8/861 (0.9%)	18/877 (2.1%)	5/635 (0.8%)	**43/3031 (1.4%)**

*Excludes one woman with ovarian cancer

**Excludes one woman with ovarian cancer and two women with active cancer

Note that combined total numbers of women consulting during census week exclude women with ovarian or other active cancers. Women aged 45-74 (n=4), women aged 55-74 (n=1)

Table 6-6 contains demographic details of women who consulted in each age group. Ethnicity data were poorly documented at the most of the surgeries examined, however the majority of women are likely to be British or mixed British according to the surgery catchment areas. The proportion of consulting women with IBS or diverticular problems recorded was similar across all practices.

Table 6-6 Demographic Details of Women with a Consultation in the Census Week

	Chrisp St	Woosehill	Boundary	Northcroft	**TOTAL**
Women aged 45-74	n=168	n=85	n=59	n=163	**475**
Median Age (IQR)	58 (49-67)	56 (51-63)	56 (49-67)	60 (51-65)	**58 (50-66)**
Ethnicity					
British/Mixed British	110 (65%)	24 (28%)	-	1 (1%)	**133 (28%)**
Black*	9 (5%)	-	-	-	**10 (2%)**
Unknown	50 (30%)	61 (72%)	59 (100%)	162 (99%)	**332 (70%)**
Relevant Diseases					
IBS	9 (5%)	11 (13%)	4 (7%)	15 (9%)	**39 (8%)**
Diverticular problem**	8 (5%)	5 (6%)	4 (7%)	11 (7%)	**28 (6%)**
Women aged 55-74	n=96	n=51	n=35	n=104	**286**
Median Age (IQR)	66 (59-72)	62 (58-67)	64 (59-71)	63 (60-68)	**63 (60-69)**
Ethnicity					
British/Mixed British	64 (67%)	14 (27%)	-	1 (1%)	**79 (28%)**
Black*	9 (10%)	-	-	-	**9 (3%)**
Unknown	23 (24%)	37 (73%)	35 (100%)	103 (99%)	**198 (69%)**
Relevant Diseases					
IBS	6 (6%)	5 (10%)	3 (9%)	12 (12%)	**26 (9%)**
Diverticular problem**	-	2 (4%)	3 (9%)	10 (9%)	**15 (5%)**

*Includes Caribbean, Black British
**Includes diverticular disease, diverticulosis, diverticulitis

Grouped symptom proportions were more variable between practices due to smaller numbers (Table 6-7). GI symptoms were the most commonly reported and gynaecological symptoms were the least common. A greater proportion of women reported GI symptoms at Northcroft and Boundary surgeries, however this could be due to random variation. Overall, symptom reporting was comparable for all grouped symptom categories.

Table 6-7 Proportion of Women in each Age Group Presenting During the Week with Grouped Symptoms (95%CI) at each Surgery

	Percentage of Women with Symptom (95%CI)			
	Chrisp Street	Woosehill	Northcroft	Boundary
Women aged 45-74	n=168	n=85	n=59	n=163
Abdominal	4.2 (1.7, 8.4)	3.5 (0.7, 10.0)	3.1 (1.0, 7.0)	5.1 (1.1, 14.1)
GI	7.1 (3.7, 12.1)	4.7 (1.3, 11.6)	11.0 (6.7, 16.9)	11.9 (4.9, 22.9)
Urinary	1.8 (0.4, 5.1)	4.7 (1.3, 11.6)	1.8 (0.4, 5.3)	-
Gynaecological	0.6 (0.0, 3.3)	-	0.6 (0.0, 3.4)	-
General*	6.5 (3.3, 11.4)	4.7 (1.3, 11.6)	6.1 (3.0, 11.0)	6.8 (1.9, 16.5)
Women aged 55-74	n=96	n=51	n=35	n=104
Abdominal	3.1 (0.6, 8.9)	3.9 (0.5, 13.4)	1.9 (0.2, 6.8)	2.9 (0.1, 14.9)
GI	7.3 (3.0, 14.4)	5.9 (1.2, 16.2)	11.5 (6.1, 19.3)	8.6 (1.8, 23.1)
Urinary	2.1 (0.3, 7.3)	2.0 (0.0, 10.4)	2.9 (0.6, 8.2)	-
Gynaecological	-	-	1.0 (0.0, 5.2)	-
General*	4.2 (1.1, 10.3)	5.9 (1.2, 16.2)	2.9 (0.6, 8.2)	5.7 (0.7, 19.2)

*Includes weight loss, loss of appetite, fatigue

Table 6-8 shows the estimated proportion of women who presented at each practice in each age range that would be offered testing in one week based on different criteria. According to the basic testing threshold of any 'key' symptom, 80 women (i.e. 16.8% of consulting women) aged 45-74 would be eligible, which is equivalent to 1.4% (95%CI 0.2%, 1.7%) of all women aged 45-74 (combined list size). The proportion of 'key' symptoms that were considered to be 'unattributable' to other diseases varied between surgeries.

Of women aged 45-74 who had a 'key' symptom, 57 out of 80 (71%) reported a 'new' symptom. For women aged 55-74, this was 28 out of 43 (65%). Hence, about two thirds of women with 'key' symptoms would require testing if the threshold was based on 'new' symptoms only.

In order to get a more clinically relevant estimate, women who would not realistically be offered testing were excluded (e.g. women with 'attributable' symptoms only, women in palliative care or already referred for a malignancy etc.). The proportion of women who required testing was markedly reduced across all practices. The overall figures were 29 women aged 45-74 (i.e. 6.1% of consulting women) and 12 women aged 55-74 (i.e. 4.2% of consulting women).

All women who had ≥3 'key' symptoms also had at least one 'new' or worsening symptom, thus only data for women with ≥3 'key' symptoms are presented in Table 6-8.

Very few of the consulting women met this testing threshold; just 3.0% at Chrisp Street, 1.8% at Northcroft surgery and none at Woosehill and Boundary surgeries.

Reporting of consensus symptoms varied slightly at each practice but overall was low (5.9% of women aged 45-74). The proportion of women that met each testing threshold was comparable for each age group.

The symptom index (Index 2) that was developed in Part II, was applied to the data (see Table 6-9). Again, results were similar for each age group, as few as 0.8% of all eligible women reported at least one new index symptom over the week.

Table 6-10 shows the estimated proportion of women in each age group that would require testing in one year. An estimated 51.8% (95%CI 44.0%, 59.7% for 45-74 group) of women would present with a 'key' symptom in each age group. For any new index symptom, 35.4% (95%CI 27.5%, 44.0%) and 32.7% (95%CI 22.2%, 44.8%) of women aged 45-74 and 55-74, respectively would be tested in one year. The proportion of women who were estimated to report 3 or more 'key' symptoms was very low, just 7.0% (95%CI 3.1%, 13.3%) of women aged 45-74.

Table 6-8 Estimated Proportion of Women to Offer Testing to in One Week According to Different Testing Thresholds (GP or Nurse)

	Number (percentage) with 95% Confidence Intervals								
	Chrisp Street		Woosehill		Northcroft		Boundary		**TOTAL**
	n (%)	95%CI	n (%)	95%CI	n (%)	95%CI	n (%)	95%CI	**n (%)**
Women Aged 45-74	n=168		n=85		n=163		n=59		**n=475**
Key Symptoms									
Any	27 (16.1)	10.9, 22.5	14 (16.5)	9.3, 26.1	28 (17.2)	11.7, 23.9	11 (18.6)	9.7, 30.9	**80 (16.8)**
'New'	20 (11.9)	7.4, 17.8	10 (11.8)	5.8, 20.6	20 (12.3)	7.7, 18.3	7 (11.9)	4.9, 22.9	**57 (12.0)**
'Unattributable'*	18 (10.7)	6.5, 16.4	7 (8.2)	3.4, 16.2	15 (9.2)	5.2, 14.7	8 (13.6)	6.0, 25.0	**48 (10.1)**
Actually Eligible for Test**	13 (7.7)	4.2, 12.9	2 (2.4)	0.3, 8.2	10 (6.1)	3.0, 11.0	4 (6.8)	1.9, 16.5	**29 (6.1)**
≥3 Key Symptoms	5 (3.0)	1.0, 6.8	-	-	3 (1.8)	0.4, 5.3	-	-	**8 (1.7)**
Any Consensus Symptom	10 (6.0)	2.9, 10.7	7 (8.2)	3.4, 16.2	8 (4.9)	2.1, 9.4	3 (5.1)	1.1, 14.1	**28 (5.9)**
Women Aged 55-74	n=96		n=51		n=104		n=35		**n=286**
Key Symptom									
Any	12 (12.5)	6.6, 20.8	8 (15.7)	7.0, 28.6	18 (17.3)	10.6, 26.0	5 (14.3)	4.8, 30.3	**43 (15.0)**
'New'	7 (7.3)	3.0, 14.4	6 (11.8)	4.4, 23.9	13 (12.5)	6.8, 20.4	2 (5.7)	0.7, 19.2	**28 (9.8)**
'Unattributable'*	7 (7.3)	3.0, 14.4	2 (3.9)	0.5, 13.5	9 (8.7)	4.0, 15.8	4 (11.4)	3.2, 26.7	**22 (7.7)**
Actually Eligible for Test**	5 (5.2)	1.7, 11.7	-	-	6 (5.8)	2.1, 12.1	1 (2.9)	0.1, 14.9	**12 (4.2)**
≥3 Key Symptoms	3 (3.1)	0.6, 8.9	-	-	2 (1.9)	0.2, 6.8	-	-	**5 (1.7)**
Any Consensus Symptom	5 (5.2)	1.7, 11.7	3 (5.9)	1.2, 16.2	5 (4.8)	1.6, 10.9	1 (2.9)	0.1, 14.9	**14 (4.9)**

*Symptoms unattributable to other diseases

**Excludes women who would not be offered testing because symptoms attributable to other diseases, pre-existing symptom, referred to ultrasound or other specialists, in palliative care, already would have been tested at previous visits etc.

Table 6-9 Number (%) of Women with an Index 2 Symptom – All Consulting and All in Age Group

	Chrisp		Woosehill		Northcroft		Boundary		**TOTAL**
Women aged 45-74*									
Consulting Women	n=168	95%CI	n=85	95%CI	n=163	95%CI	n=59	95%CI	**n=475**
	18 (10.7%)	6.5%, 16.4%	9 (10.6%)	5.0%, 19.2%	18 (11.0%)	6.7%, 16.9%	3 (5.1%)	1.1%, 14.1%	**48 (10.1%)**
All Women in Age Gp.	n=1210	95%CI	n=1809	95%CI	n=1457	95%CI	n=1261	95%CI	**n=5737**
	18 (1.5%)	0.9%, 2.3%	9 (0.5%)	0.2%, 0.9%	18 (1.2%)	0.7%, 1.9%	3 (0.2%)	0.0%, 2.2%	**48 (0.8%)**
Women aged 55-74**									
Consulting Women	n=96	95%CI	n=51	95%CI	n=104	95%CI	n=35	95%CI	**n=286**
	7 (7.3%)	3.0%, 14.4%	6 (11.8%)	4.4%, 23.9%	10 (9.6%)	4.7%, 17.0%	0	-	**23 (8.0%)**
All Women in Age Gp.	n=658	95%CI	n=861	95%CI	n=877	95%CI	n=635	95%CI	**n=3031**
	7 (1.1%)	0.4%, 2.2%	6 (0.7%)	0.3%, 1.5%	10 (1.1%)	0.5%, 2.1%	0	-	**23 (0.8%)**

*Excludes one woman with recent colectomy for suspected colon cancer and one woman in palliative care
**Excludes one woman with recent colectomy for suspected colon cancer
Note: Index 2 symptoms include increased abdominal size, lump in abdomen, loss of appetite, pelvic/abdominal pain or discomfort, constipation, fatigue, bloating, nausea or vomiting and weight loss.
Abbreviations: Gp. Is “Group”

Table 6-10 Proportion of Women to be Screened Extrapolated Estimate Over 1 Year with 95% Confidence Intervals

	Percentage	95% CI
Women Aged 45-74		
Key Symptom		
Any	51.8%	44.0%, 59.7%
'New'	40.5%	32.5%, 49.0%
'Unattributable'*	35.4%	27.5%, 44.0%
Actually Eligible for Test**	23.2%	16.2%, 31.5%
≥3 Key Symptoms	7.0%	3.1%, 13.3%
Any Consensus Symptom	22.5%	15.6%, 30.8%
Any New Index 2 Symptom	35.4%	27.5%, 44.0%
Women Aged 55-74		
Key Symptom		
Any	52.4%	41.6%, 63.2%
'New'	38.3%	27.4%, 50.2%
'Unattributable'*	31.5%	21.1%, 43.6%
Actually Eligible for Test**	18.6%	10.1%, 30.3%
≥3 Key Symptoms	8.2%	2.7%, 18.2%
Any Consensus Symptom	21.4%	12.3%, 33.2%
Any New Index 2 Symptom	32.7%	22.2%, 44.8%
Actually Eligible for Test**	14.3%	8.6%, 21.9%

*Symptoms unattributable to other diseases
**Excludes women who would not be offered testing because symptoms attributable to other diseases, pre-existing symptom, referred to ultrasound or other specialists, in palliative care, already would have been tested at previous visits etc.

A total of 15 blood tests and 3 ultrasounds were ordered during the week for women aged 45-74 presenting with a 'key' symptom. However, only 10 of the blood tests were ordered for investigation of symptoms and the rest were routine. This is equivalent to 19% (15/80) of women in the age group consulting during the week with 'key' symptoms receiving a blood test. Six women were referred for further care.

6.5 Discussion

GP workload is an important consideration in the feasibility assessment for 'targeted' ovarian cancer screening. Consultation rates are known to be high in the very young and very old, and appear to be on the increase for older age groups in the UK (over 65s).[155] The total number of women visiting in one week was 479, slightly less than the 550 visits expected based on the estimates from the Office of Health Economics. [154] However, some women visited more than once in the week and the total number of consultations was actually 560, which is very close to our projection. Unfortunately, multiple visits were not accounted for in our calculations, thus our sample size was slightly smaller than desired.

'Targeted screening' might only be offered to women with more than one relevant symptom. Many of the individual symptoms that might prompt testing alone are already likely to prompt a referral to secondary care, for example PMB, lump in abdomen, change in bowel habit or abdominal distension (with suspected ascites). Under these circumstances, the index of suspicion for malignancy would be too high for 'targeted screening'. Conversely, testing for urinary frequency or fatigue alone could result in a surplus of unnecessary testing in women with simple urinary tract or viral infections. In order to address these issues, theoretical testing was assessed on several different levels in this study.

One of the main arguments against the promotion of using symptoms as a tool for early detection of ovarian cancer is the lack of specificity of symptoms which would result in many women requiring unnecessary investigation. More than 50% of postmenopausal women suffer from at least one of vaginal discomfort, dysuria, dyspareunia, recurrent lower UTI and urinary incontinence.[156] This pilot study has indicated that only a small proportion of women aged 45-74 present to primary care with symptoms potentially related to ovarian cancer per week. However, extrapolation over one year estimated that 'targeted screening' would potentially be offered to about half of all women aged 45-74 (51.8%) if no restrictions were set. Although this is a large proportion of women, the extrapolations over one year in this study are likely to be slightly overestimated. This was exemplified by the substantial reduction in testing when only 'new' and 'unattributable' symptoms were included. The proportion of women estimated to require testing over one year for a 'new' symptom was 40.5% (95%CI 32.5%, 49.0%), and for an 'unattributable' symptom this was 35.4% (95%CI 27.5%, 44.0%). The criteria for defining 'new' were limited to the data available. In the present study, 'new' index symptoms were supposed to be new in the last year. However, it is likely that we

had insufficient consultation data to cover this period and some of the symptoms categorised as 'new' were in fact 'old'. Furthermore, menopausal status was unknown in this study, and 'targeted' screening is likely to be restricted to postmenopausal women. Therefore, in reality the number of women to be offered testing would be slightly lower.

Symptom reporting in the 45-74 year age group was relatively high (approximately 2/3 of women consulting), however 'key' symptom reporting was not. Any testing threshold for 'targeted screening' is likely to include some sort of symptom duration restriction which was unaccounted for in this analysis. For example, it would not be sensible to test after one day of loss of appetite. The symptom index produced by Goff *et al.* requires symptoms to have been present for less than one year but more than 12 times per month.[108] The application of similar criteria should further reduce the number of women who require testing, as would the restriction to only certain 'key' symptoms as in Goff's index (increased abdominal size or bloating, pelvic or abdominal pain, difficulty eating or feeling full quickly). Furthermore, testing would not be offered to women who have already been recently tested (say the last 6-12 months) which would also decrease testing numbers. According to Index 2 (developed in Part II), only an estimated 35.4% (95%CI 27.5%, 44.0%) of women aged 45-74 would require testing in one year.

The proportion of women who were actually eligible for 'testing' was intended to provide a more clinically realistic estimate. Although the approach used was crude, this almost halved the proportion of women who would require 'testing' for any 'key' symptom (51.8% [95%CI 44.0%, 59.7%] was reduced to 23.2% [95% CI 16.2%, 31.5%]) and any 'new' Index 2 symptom.

Despite having diverse locations, the four GP practices studied showed comparable proportions of women in each age group reporting at least one 'key' symptom. In contrast, the proportion of 'key' symptoms that were 'attributable' to other diseases varied and one GP practice had no women with 'attributable' symptoms, however this assessment was vague. Due to the clinical judgement required on a case-by-case basis for this category, it should not be over-interpreted for this study. Moreover, clinical diagnosis can be delayed due to misattribution of symptom aetiology. It is unclear why the proportion of women that consulted relative to list size in each age group differed across surgeries. This was unexpected and possibly due to the short period examined. Although the women were analysed in two different age bands, very few differences were observed between the smaller more postmenopausal age group that and the larger mixed menopausal status age group.

No other published studies have investigated the incidence of typical ovarian cancer symptoms in primary care in women aged 45-74 using medical record data. Goff *et al.* collected survey symptom data on women presenting to primary care, however they included women as young as 15 who probably have a different spectrum of symptoms.[75] The only other published study that has investigated 'targeted ovarian cancer screening' was a prospective pilot study which randomised GP practices to being able to refer (or not refer) women aged ≥45 with symptoms potentially related to ovarian cancer for immediate CA125 testing and transvaginal ultrasound.[111] Study compliance was poor and only 39 out of 79 practices made referrals during the study period. Referral patterns were also extremely variable between GPs. Over the recruitment period that spanned almost 2½ years, only 317 women were referred. Clearly, this is much lower than the present study's estimated numbers. Since no cases of ovarian cancer were identified in the study population, no indication of whether or not symptoms may facilitate earlier detection could be inferred.

Current referral criteria for rapid access gynaecological-oncology clinics (relevant to ovarian cancer) only include two symptoms: suspicious pelvic mass and PMB (if not on HRT). Thus, there are many ovarian cancer symptoms that do not meet rapid referral guidelines. 'Targeted screening' may provide a useful tool to utilise when these symptoms are present but suspicion of ovarian cancer is not sufficient to prompt referral or more expensive investigations.

Potential limitations of this study include that GP practices were limited to those with EMIS LV and therefore may not be representative of all practices across the UK. Also, there may be bias from GP surgeries that were willing to participate, as they may have been more organised practices with better quality notes. Another drawback is that the study was small and only one week's worth of consultations was examined. This period is too short to take into account seasonal variations in consultations (for example, all patients over 65 are offered free flu vaccinations in the winter months). In addition, the prevalence of 'key' symptoms in the population of women aged 45-74 that did not consult in the census week remains unknown. Consultation rates have been shown to be higher in rural than in inner-city and urban areas[155] and no rural practices were included in the current study. Another issue to consider, is that recording in GP notes tends to be diagnosis driven,[24] and actual reporting of symptoms in primary care may in fact be underestimated by this study. Also, the prevalence of 'key' symptoms in the non-consulting population could be high. If women were to be encouraged to seek health-care for these symptoms, the associated GP workload could become much greater.

However, importantly this study has provided much needed data on ovarian cancer symptom prevalence in primary care, albeit limited. The results will help to design future studies and provides 'ballpark' figures.

Further work is planned to gain longitudinal data to confirm these pilot findings, and to refine the potential testing thresholds.

6.6 Conclusions

According to these pilot data, 40.5% (95%CI 44.0%, 59.7%) of women aged 45-74 present to primary care in one year with a 'new' ovarian cancer symptom. An even smaller proportion were estimated to have an Index 2 symptom over one year (35.4% [95%CI 27.5%, 44.0%]). Crucially, restricting women to those who were actually eligible for testing reduced the numbers considerably. This is clearly something that should be built in to any future studies assessing GP workload and ovarian cancer symptoms.

Although there are other issues to consider, these data are encouraging since the numbers are considerably less than those required for population screening and in reality, stricter criteria would be applied which should further reduce these figures. However, these data do not take into account any potential increases in symptom reporting at primary care from increased ovarian cancer symptom awareness of women in the general population. Further work needs to be done to find out if 'targeted screening' would have an acceptable NNI (number needed to investigate).

7 CHAPTER 7: Summary & General Discussion

The following discussion will summarise the key findings of this thesis, and it will highlight what has been added to the knowledge-base of symptoms and the events prior to ovarian cancer diagnosis. It will also cover the potential implications and identify the areas that require future work.

The key findings of this thesis include:

Any Symptom (possibly related to ovarian cancer)

- Most women with ovarian cancer have symptoms at diagnosis.
- Women with ovarian cancer start to have more problem visits to the GP approximately 6 months before diagnosis.
- No case-control differences were observed beyond 15 months prior to diagnosis. Symptoms provide a maximum (diagnostic) lead time of about 15 (i.e. 14.99) months (12.99 months (actual) lead time if 2 months are allowed for from referral to diagnosis).
- Delays in diagnosis were identified, 50% of symptomatic women (according to our definition) had at least 9.8 months (IQR 5.0-12.3) between symptom onset and diagnosis (combined data). Furthermore, total and provider delays may be longer in the UK, however this needs to be confirmed.
- 64%-81% (depending on data source) of women with ovarian cancer reported at least one symptom possibly related to ovarian cancer over 1 year (3 to 14.99 months before diagnosis).
- However, 20%-47% of women in the general population would also have these symptoms over 1 year (depending on source of data, study design and precise list of symptoms) and therefore would be offered 'targeted screening'.
- Based on estimates using pilot data, limiting women in the general population to those who would actually be eligible for 'targeted screening' over 1 year, approximately halved the numbers.

Symptom Index

- Using a crude symptom index (to improve specificity), 45%-74% of women with ovarian cancer had symptoms at least 3 months prior to diagnosis.
- 13%-35% (depending on source of data and study design) of women in the general population had an index symptom over the course of 1 year.
- After allowing for time from referral to diagnosis, 26%-56% of women with ovarian cancer had an actual symptom lead time of at least 3 months.

Methodological Issues

- Timing of symptoms is an important consideration in study design given that symptom reporting is markedly reduced >3 months prior to diagnosis.
- Complexities of symptoms research and related terminology have not been fully appreciated, and need to be addressed in future studies.

Symptoms that were found to be highly associated with ovarian cancer included:

- Increased abdominal size
- Loss of appetite
- Pelvic/abdominal pain or discomfort
- Lump in abdomen
- Constipation
- Fatigue
- Abdominal bloating
- Weight loss

Other symptoms that were associated with ovarian malignancy included:

- Change in bowel habit
- Irregular vaginal bleeding
- Postmenopausal bleeding
- Vaginal discharge
- Nausea or vomiting
- Urinary other symptoms

Back pain was found to be inversely associated with ovarian cancer, hence should probably only be used in future studies as a negative discriminatory factor for cases, if at all.

Importantly, this thesis has shown that consideration of the timing of symptoms is critical. Symptom data that are collected retrospectively need to be carefully evaluated in the light of symptoms as a process over time. Recently, the Department of Health (UK) released ovarian cancer key messages for health professionals,[157] in which much emphasis was placed upon the severity, frequency and persistence of ovarian cancer symptoms. However, these are attributes of symptoms that were described in hindsight (i.e. at diagnosis), and there are no prospective data to confirm these claims. Ovarian cancer symptoms may have very different severity and frequency when they first appear compared with at diagnosis. Also, some symptoms do not need to be

sustained temporally to be salient (e.g. postmenopausal bleeding, acute episodes of severe pain).

The Goff index[108] was developed using data on symptoms *at* diagnosis. When the same index was applied to our data, performance for symptoms at diagnosis was similar, but performance for symptoms that started between 3-14 months before diagnosis was markedly worse. Therefore, the provocative series of studies published by the Goff group may have sparked premature optimism for a symptoms-based diagnostic tool for ovarian cancer.[8, 75, 103, 108] Nevertheless, it is encouraging that the same or similar symptoms in the Goff index were found to be salient in the crude indices explored for this thesis. Furthermore, it should be pointed out that the possibility of developing a symptoms tool to facilitate earlier diagnosis remains viable. However, it is paramount that retrospectively collected symptom data are viewed from a real life perspective (i.e. as events that may change as a function of time).

Prospective data would greatly advance our knowledge of what would truly be feasible. As such, Goff's prospective validation work on the symptoms index is keenly awaited. However, there is an important distinction between Goff's ongoing work and the approach suggested in this thesis. Goff's index is being assessed as a questionnaire which is administered to women attending primary care. This could produce very different results from offering 'testing' based on symptoms spontaneously reported by women at primary care. Also, the Goff index includes symptoms that are present for up to 12 months, which could be prone to recall error and potentially could greatly increase the proportion of women that would require 'testing' by including 'older' symptoms.

At University College London (UCL), another ongoing prospective study in ovarian cancer symptoms is taking place. The study aims to prospectively evaluate symptoms in apparently healthy women participating in the UKCTOCS ovarian cancer screening trial. The study is using an ovarian cancer symptoms questionnaire that was recently validated using the EORTC guidelines (Penny Allen, Institute for Women's Health, UCL, personal communication). The results from this study should provide extremely useful prospective data on symptoms in ovarian cancer and data on the relationship between CA125 and symptoms.

Interpretation of symptom data and the development of any symptoms-based tool will need to be considered carefully. One of the main issues is how the symptom data were elicited as this would have an impact on what setting the tool could be sensibly used. In general, there are several different ways in which symptoms could be used to

bring forward diagnosis. Each would produce dramatically different sensitivity and specificity. The 3 main approaches include:

1) Offer 'testing' based on a list of symptoms spontaneously reported by women at the GP

2) Offer 'testing' based on a list of symptoms specifically asked about by the GP when women present to primary care (for any reason)

3) Encourage women to visit their GP for symptoms on a list

The balance will be between sensitivity, lead time and specificity (or GP workload). Using the first approach, one would expect lower sensitivity but better specificity. The second would have higher sensitivity but lower specificity. The third would have higher sensitivity but substantially lower specificity.

The use of continuation odds ratios in this thesis was a novel approach which has not been applied before in ovarian cancer symptoms. They provided a measure that combined the advantages of a hazard ratio and an odds ratio. The major advantage of the continuation odds was that it allowed us to quantify symptom lead time by examining the proportion of incident symptoms in cases and controls over time. This allowed the retrospectively collected data to be considered in a prospective manner.

Another aspect of timing that emerged from this thesis was that ovarian cancer symptoms develop within 15 months of diagnosis. It seems increasingly unlikely that symptoms which manifest more than two years before diagnosis are associated with ovarian cancer, except perhaps in rare cases. Exacerbation of symptoms from pre-existing co-morbidities is entirely conceivable and any symptoms-based strategy may need to accommodate for this (in order to avoid diagnostic overshadowing). However, for the most part it would seem that the critical period for symptoms data collection is within 15 months prior to diagnosis.

Prompt diagnosis after presentation for symptoms is an ideal outcome and has the additional benefit of providing reassurance to patients. Any delays in diagnosis are likely to cause anxiety and frustration, yet with ovarian cancer the advancement of diagnosis by one month is unlikely to influence morbidity or survival. Equally, an expedition of two months may not produce any clinically relevant benefit. Results from the case-control study showed that it is crucial to omit symptoms reported in the 3 months before diagnosis as there could be a clustering of symptoms experienced

during this time. Certainly, in one other study, the period 1-3 months before diagnosis showed much higher symptom prevalence than any of the earlier periods examined.[20]

It remains controversial as to whether or not ovarian cancer behaves in a progressive manner (i.e. develops stepwise from stage I to IV). To date, researchers have been unable to distinguish between early and late stage disease via the presence or absence of any particular symptoms. It is entirely plausible that some early versus late tumours may arise out of completely different biological pathways. Moreover, there is no concrete evidence to suggest that women detected with advanced disease actually experienced symptoms when disease was in its 'early' stages. This assumption is at the crux of any symptoms-based diagnostic tool. Currently, the detection of advanced ovarian malignancies at an earlier time point would not translate into gains in survival or mortality. However, a proportion of women may benefit from improved quality of life as a result of smaller disease volume and optimal surgical clearance, in addition to the psychological benefits of being detected 'earlier'. Ultimately, as the time taken to get from localised tumour to disseminated disease is unknown, the impact of any period of diagnostic advancement is also unknown.

Symptoms (and data relating to symptoms) are highly subjective, and therefore rather more complicated to study than is often appreciated. Much of the research performed to date has failed to recognise the weaknesses stemming from this subjectivity. For example, data on symptom duration and delays are hugely dependent on which symptoms are deemed to be attributable to ovarian cancer. If constipation that started 5 years before diagnosis is considered to be relevant, delay in diagnosis and symptom duration will be correspondingly long. The systematic review in Chapter 2 outlined some of the methodological issues associated with carrying out and evaluating symptom studies in ovarian cancer. Methods of symptom elicitation and collection have a significant bearing on outcome measures. Simply asking if patients had symptoms before diagnosis may yield inadequate data due to preconceptions about which symptoms are attributable to ovarian cancer. Moreover, mislabelling and misinterpretation of symptoms seem to be a significant problem.

Patient and provider delays varied markedly with each data source in our study. As demonstrated in the systematic review, delays data in ovarian cancer overall have been extremely variable. Again, this is attributable to the subjectivity in identifying symptoms that are associated with ovarian cancer, and whether or not calculations are based on all women or symptomatic women only. The methods used for calculation of delays needs to be standardised and details should be explicitly described in study methods. Given the evidence (or lack of it), education of women and GP's seems to be

premature at this stage. While the ovarian cancer key messages released by the Department of Health[157] may be useful in that GPs may be more likely to consider ovarian cancer in women who present with the symptoms listed, there is still no evidence to show that this would bring diagnosis forward or have any positive impact at all. There are many unanswered questions and more robust research which takes into account the various methodological issues is desperately needed.

Crucially, this thesis has provided data on symptoms in the general population (primary care consulting and non-consulting). Table 7-1 compares the proportion of women in the general population who would require 'testing' over the course of 1 year according to each of the studies in this thesis. The proportion of women aged 45-74 in the general population that would also require 'testing' over one year appears to be around 35%-41% in the GP workload data (see Table 7-1). Llimiting women to those who would actually be tested in a clinical setting approximately halved the numbers. Only about 23% (95%CI 16%, 32%) of women with a 'new' 'key' symptom were estimated to be actually eligible for 'testing'. In the case-control study, the proportion varied greatly with each data source, despite being based on the same women. This highlights again the need for careful consideration of data source and methodology when estimating symptom specificity. GP note and interview data were the most similar to the GP workload data. Theoretically, self-reported data represent a mix of symptoms for which women have sought health care and those for which they have not. Thus, the specificity of self-reported data is expected to be lower than that of GP note data, however this was not observed. This may be due to overestimation in the extrapolation (GP workload) or perhaps could reflect difficulties in recalling transient symptoms in the self-reported data. After the application of a symptoms index, the proportion of women to be 'tested' reduced across all women (both studies). Examination of GP workload using longitudinal data is expected to produce even smaller estimates. In addition, 'targeted screening' would most sensibly be offered based on incident (not prevalent) symptoms, with a minimum screening interval. Both of these factors should further reduce the proportion of women in the general population who would require 'testing'.

Table 7-1 Comparison of Estimates for the Proportion of Women in the General Population who would be Offered 'Targeted Screening' Over 1 Year (95%CI) - Controls (Case-Control Study) versus Women in GP Workload Study

	Controls from Case-Control Study			GP Workload Study	
	Questionnaire % (95%CI)	Interview % (95%CI)	GP Notes % (95%CI)	New % (95%CI)	Actually Eligible** % (95%CI)
Any Symptom*	20.1% (15.5%, 25.5%)	36.8% (28.4%, 45.9%)	46.7% (40.1%, 53.4%)	40.5% (32.5%, 49.0%)	23.2% (16.2%, 31.5%)
Index 2 Symptom***	13.4% (9.6%, 18.1%)	26.4% (18.9%, 35.0%)	21.6% (16.4%, 27.5%)	35.4% (27.5%, 44.0%)	14.3% (8.6%, 21.9%)

*Any symptom possibly related to ovarian cancer (as per Appendix VI) incident over 3-14.99 months before consent for case-control study. Any new (last 3 months) 'key' symptom (as per Appendix X) estimated over 1 year for GP workload study.

**Women who would actually be eligible for testing based on a new 'key' symptom. Excludes women who would not be offered testing because symptoms attributable to other diseases, pre-existing symptom, referred to ultrasound or other specialists, in palliative care, already would have been tested at previous visits etc.

***Any one of increased abdominal size, lump in abdomen, loss of appetite, pelvic/abdominal pain or discomfort, constipation, fatigue, bloating, nausea or vomiting, weight loss

The concept of identifying a symptom index using multiple symptoms is attractive given that the positive predictive value (PPV) of individual symptoms is likely to be extremely low (less than 1%). A symptom index could provide a useful differential upon which the threshold for suspicion could be lowered, such that women who are at theoretically higher risk of disease will undergo a simple 'screening' procedure. Symptoms of ovarian cancer are predominantly abdominal and gastrointestinal. The normal investigations for abdominal symptoms include barium meal or enema, colonoscopy and sigmoidoscopy. Such tests are expensive and can be painful, therefore the inclusion of a CA125 blood test as part of a standard differential diagnosis for GI symptoms may be easy to justify.

Importantly, 'targeted screening' as proposed in this thesis, is not totally dependent on the success of CA125 testing. The principle of using vague symptoms to prompt a blood test for ovarian cancer can be supported by any blood-borne marker testing that may become clinically available in the future. However, there will need to be clear protocols developed if CA125 is to be used.

Regrettably, symptoms-based strategies remain unproven and it is prudent that decision-making and dissemination of information is founded on empirical evidence. Currently, there is insufficient evidence to advocate that women should be made aware of the symptoms and encouraged to consult primary care givers for symptoms. Ultimately, the goal of cancer screening is to reduce mortality. In cancer screening, it is critical to consider what is being detected. If screening (symptoms-based or conventional) is only able to detect disease that would have had a good prognosis without screen-detection and/or disease that has a poor prognosis despite being screen-detected, the likelihood of a mortality benefit is low. In the absence of evidence, it is misleading to propagate the idea that acting on symptoms will save lives. Ovarian cancer symptoms research should be continued with scientific rigour in a logical manner which ensures that all of the lessons learnt so far are incorporated, and that hasty conclusions are not drawn.

The future of any paradigm based on symptoms is contingent on the elucidation of several key unanswered questions:

- Is symptom lead time sufficiently long to translate into benefits in survival, mortality, quality of life or psychological well-being?
- What proportion of women in the general population (consulting and non-consulting) will also require testing? (i.e. NNI [number needed to investigate])
- Will increasing symptoms awareness increase GP workload?

- What is the relationship between the potential second line tests and symptoms? (i.e. will ovarian cancer be detected)

Assuming that these questions are answered and 'targeted screening' remains feasible in light of these answers, there are still many more general (but significant) hurdles that would need to be overcome. These include:

- The heterogeneity of ovarian cancer – will all (or at least a reasonable proportion of) tumours benefit from symptoms-based efforts?
- Increasing symptoms awareness is not the same as changing behaviour. How feasible is it to get women and GPs to act on symptoms to produce these benefits?
- What would be done about negative investigations but persistent symptoms?

In the pilot study that examined the feasibility of 'targeted screening', GP compliance was poor.[111] The reasons for this will need to be carefully examined. Early detection may be hindered by women who delay presenting to healthcare even armed with knowledge of symptoms, and by patients who do not attend referral visits.[158] Also, unravelling the relationship between symptoms and CA125 levels, and between symptoms and ovarian cancer survival, is integral to the goals listed above. Finally, a crucial unanswered question for the UK is why do British women have lower survival rates than the rest of the world?

Future Work

For future research (or any symptoms tool that relies on women's ability to interpret symptoms), it is imperative that the language and terminology used is clear and consistent. Also, the tools used to elicit symptom data should be carefully considered. The recently validated ovarian cancer symptoms questionnaire should prove invaluable for addressing some of these measurement issues.

As mentioned previously, there are two ongoing prospective studies in ovarian cancer symptoms (validated symptoms questionnaire in healthy volunteers [UK] and the Goff index in primary care [US]). In tandem with this work, more research should be dedicated to confirming the pilot data in this study regarding the proportion of women in the general population that would need 'targeted screening'.

Recently, a study investigated the GP acceptability of blood testing based on symptoms in ovarian cancer (Megan Goldsmith *et al.*, Primary Care Group Oxford, unpublished data). The main findings were that the concept of 'targeted screening' was foreign to GPs and many felt that such a scheme would delay diagnosis. Clearly,

there are many more steps that need to be taken before ‘targeted screening’ could be implemented.

More work is also needed to better define delays in diagnosis, and to identify barriers in the diagnostic process. Any studies investigating delays should allow for the mandatory phases in the diagnostic pathway. Specifically, efforts should be directed at standardising the definition and quantification of ‘delays’ in diagnosis. Moreover, the identification of more appropriate terms to describe the interval between symptom onset and first GP visit, first GP visit and diagnosis, and symptom onset and diagnosis, is long overdue.

Given the current evidence, efforts to increase symptoms awareness in women and GPs are probably premature. It would be prudent to wait for stronger evidence (in the form of the ongoing prospective studies), before any strong recommendations for increasing awareness are made.

In the future, perhaps electronic patient record systems could play a valuable role by generating automatic flags to prompt GPs to offer testing based on symptoms and act as reminders for repeat testing.

Lastly, in Part II, the case-control data were extremely complex, and many more avenues remain unexplored that were beyond the scope of this thesis. Planned analyses include a combined analysis, examination of the relationship between symptoms and survival (all of the women have been flagged with the Office of National Statistics [ONS]).

Although the development of a screening test via proteomics or other genetically-based tools is more likely to produce a much needed survival shift in ovarian cancer, this could be years away from clinical practice. Also, the efficacy of population screening in ovarian cancer will not be clear until 2015. Therefore, symptoms research continues to play a major role in the potential for early detection of ovarian cancer. Even if symptoms are ultimately proven to be an ineffective tool for expediting diagnosis in ovarian cancer, this would also constitute an important finding, as scarce resources could be directed to other areas.

8 Appendices

8.1 Appendix I Health Changes/Symptoms Questionnaires

Order inserted:

Amendment 2 Questionnaire Version 2 dated January 2006 (first version in use for current thesis project, telephone interview consent added, version and date was not changed in error)

Amendment 3 Questionnaire Version 3 dated June 2006 (version used in most of the women in this thesis, designed to improve data quality and completion)

Section 7: PHYSICAL CHANGES

7.1 Have you had any of the following in the *past year*?

If so, please indicate the severity, frequency (number of days per month), and duration.

	Severity (mild/ moderate/ severe)	No. of days per month	Duration (months e.g. OCT-DEC)	On-going (tick)	Did you tell your GP about this? If so, when?
Pelvic or Abdominal pain				☐	
Back pain				☐	
Indigestion				☐	
Nausea or vomiting				☐	
Weight loss (unplanned) or appearance of weight loss				☐	
Abdominal bloating (abdomen feels full & tight)				☐	
Abdominal swelling (an actual increase in abdominal size/girth)				☐	
Able to feel a lump in the abdomen				☐	
Urine (passing urine more than usual or feeling an urgent need to go)				☐	
Constipation				☐	
Diarrhoea				☐	
Irregular periods				☐	
Bleeding after menopause				☐	
Pain during intercourse				☐	
Bleeding with intercourse				☐	
Fatigue				☐	
Leg swelling				☐	
Other changes *please specify*:				☐	

Please tick if you had **NONE OF THE ABOVE:** ☐

MAY WE CONTACT YOU AGAIN FOR FURTHER TELEPHONE INTERVIEW: Yes ☐ No ☐

If yes, preferred contact telephone number: _____

2265627639 *Risk Prediction and Early Detection of Ovarian Cancer*

Section 3: HEALTH CHANGES

3.1 Please complete the table for the past 12 months.
If you have been diagnosed with ovarian cancer in the past, please complete for the year before diagnosis.

For each "Yes" please complete the entire row	Have you had this?	When did you first have this?	How many months have you had/did you have this?	Do you still have this?	Did you tell your GP? If so, when?	When this FIRST started: About how many days per month did you have this? 1 - 4 days	5-15 days	16-31 days	How severe was it? Mild	Moderate	Severe
Example ***Indigestion***	(Y)/N	*Aug 2005*	8	Y(N)	(Y)/N *Nov 2005*	✓				✓	
Pelvic or abdominal pain or discomfort	Y/N			Y/N	Y/N						
Back pain	Y/N			Y/N	Y/N						
Indigestion	Y/N			Y/N	Y/N						
Loss of appetite	Y/N			Y/N	Y/N						
Nausea or vomiting	Y/N			Y/N	Y/N						
Weight loss (unplanned) or appearance of weight loss	Y/N			Y/N	Y/N	n/a					
Abdomen feels bloated	Y/N			Y/N	Y/N						
Increase in abdominal size	Y/N			Y/N	Y/N						
Able to feel a lump in the abdomen	Y/N			Y/N	Y/N	n/a			n/a		
Urinating more often or feeling an urgent need to go	Y/N			Y/N	Y/N						
Constipation	Y/N			Y/N	Y/N						
Diarrhoea	Y/N			Y/N	Y/N						
Irregular vaginal bleeding	Y/N			Y/N	Y/N						
Fatigue	Y/N			Y/N	Y/N						
Other (please specify): e.g. *leg swelling*	Y/N			Y/N	Y/N						

Please tick if you had **NONE OF THE ABOVE**: ☐

May we contact you again for further telephone interview: ☐ **Yes** ☐ **No**
If yes, preferred contact telephone number: _____

6215

Additional Information on Health Changes

8.2 Appendix II UKOPS Health Changes Guidelines

UKOPS Health Changes Completion Guidelines

Key points

- Please ensure the **entire row** is completed for each "yes"
- Symptoms / health changes should be included regardless of perceived relationship to cancer or relevance to the study
- Symptoms / health changes should be recorded for the **last 12 months** (or for the 12 months prior to ovarian cancer diagnosis if applicable).
- Severity and frequency should be rated for when the symptom / health change **first started**
- If the subject reports no symptoms please ensure the tick box for "none of the above" is checked
- Please ensure one of the tick boxes for telephone interview is checked

Definitions

Pelvic / abdominal pain or discomfort	Pain, pressure or discomfort around the pelvis or abdomen. Exclude pain with bowel motions, intercourse, periods – these should be listed under "Other"
Back pain	Pain or discomfort in the back area particularly in the lower and mid back areas. Flank or shoulder pain should be listed under "Other"
Indigestion	A vague feeling of abdominal discomfort after eating. Possibly including belching.
Loss of appetite	Eating difficulties including not feeling hungry or being unable to eat much
Nausea or vomiting	Nausea is the sensation of having an urge to vomit. Vomiting is forcing the contents of the stomach up through the oesophagus and out of the mouth
Weight loss	Actual loss of weight or appearance of weight loss
Abdomen feels bloated	A feeling of distension or fullness in the abdomen. Abdomen may feel full and tight.
Increase in abdominal size	An actual increase in abdominal girth. Clothing may feel tighter.
Able to feel lump in abdomen	Feeling a lump in the abdomen
Urinating more often than usual or feeling an urgent need to go	An increase in frequency of passing urine compared to previously or feeling an urgent need to pass urine which often results in some leakage of urine
Constipation	Infrequent or hard stools or difficult passing stools.
Diarrhoea	More frequent bowel motions or loose watery stools.
Irregular Vaginal Bleeding	Irregular vaginal bleeding including postmenopausal bleed
Fatigue	A feeling of weariness, tiredness, or lack of energy
Other	Any other symptom, e.g. leg swelling, vaginal discharge, breathlessness

Guide to rating Severity

Mild	Easy to tolerate, did not bother you much
Moderate	Hard to tolerate, bothered you somewhat
Severe	Intolerable, bothered you a lot

Version 2 dated Sept 2006

8.3 Appendix III Telephone Interview Script for Cases

Hello, may I speak with [patient name] please…

If patient comes to the phone:

My name is [interviewer name], I'm calling with regard to the ovarian cancer study you are taking part in, is this a convenient time to speak? If not I can call back at a time that would suit you better…

If no to timing: [interviewer to arrange a time that best suits the subject]

If no to further interview: That's fine. Thank you for your time, take care.

If yes: You agreed to do an optional telephone interview as part of the study; it usually takes between 10-30 minutes. Are you still happy to do the interview? It's completely up to you….

If no: That's fine. Thank you for your time, take care.

If yes: Thank you, we really appreciate it. I'd like to arrange a suitable time to call you back to perform the interview. When would be best?

After time arranged:

Thank you again, I will speak to you then. Take care.

Interview:

Hello it's [interviewer name] speaking, may I speak to [patient name] please.

Once patient comes to the phone:

Hello, it's [interviewer name] calling to do the interview for the ovarian cancer study. Is it still okay for you do to the interview today?

If no to further interview: That's not a problem at all. Thank you for your time, take care.

If no to timing: [interviewer to arrange another a time that best suits the subject]

If yes:

Thank you. If you feel tired or wish to stop at any time, please let me know.

The research we're doing is to try to find out if there's any way we can diagnose ovarian cancer earlier by examining the events that occur before diagnosis. If you feel tired or wish to stop at any time, please let me know. The interview is about any symptoms or health problems you may have noticed in the 12 months before you had your operation/started chemotherapy (mention specific month). You answered similar questions in the clinic, but we're hoping to get a bit more information using this interview. Do you know if you will have any further treatment? [If yes ask for start date of further treatment]

It's important that you let me know if you have experienced any of these regardless of whether or not you think they have anything to do with your diagnosis.

Go through symptom questions on telephone interview form.

Did you experience any other changes or is there anything else you would like to mention?

Patients will be asked to give details about the path of diagnosis.

Changes in severity over time

Thank you very much for your time, I really appreciate it and it's really helped us. Is there anything you'd like to ask me?

Maintaining focus of interview

I'm sorry to hear that, from what you've told me it sounds like you have had a difficult time. Do you feel ready/able to continue with the interview?

There is a lot of uncertainty, symptoms are non-specific and the vast majority of women with these symptoms will not have ovarian cancer

I can tell you've been thinking about this. Do you feel ready/able to continue with the interview?

8.4 Appendix IV Telephone Interview Script for Controls

Hello, may I speak with [volunteer name] please…

If volunteer comes to the phone:

My name is [interviewer name], I'm calling with regard to the ovarian cancer study you volunteered for, is this a convenient time to speak? If not I can call back at a time that would suit you better…

If no to timing: [interviewer to arrange a time that best suits the subject]

If no to further interview: That's fine. Thank you for your time, take care.

If yes: You agreed to do an optional telephone interview as part of the study; it should take about 10 minutes. Are you still happy to do the interview? It's completely up to you…

If no: That's fine. Thank you for your time, take care.

If yes: Thank you, we really appreciate it. I'd like to arrange a suitable time to call you back to perform the interview. When would be best?

After time arranged:

Thank you again, I will speak to you then. Take care.

Interview:

Hello it's [interviewer name] speaking, may I speak to [volunteer name] please.

Once volunteer comes to the phone:

Hello, it's [interviewer name] calling to do the interview for the ovarian cancer study. Is it still okay for you do to the interview today?

If no to further interview: That's not a problem at all. Thank you for your time, take care.

If no to timing: [interviewer to arrange another a time that best suits the subject]

If yes:

The research we're doing is to try to find out if there's any way we can diagnose ovarian cancer earlier by examining events that occur before diagnosis and comparing them to that of women without cancer. The interview is about any symptoms or health problems you may have noticed in the 12 months before you were recruited to the study. The symptoms that I'm going to ask you about are quite common in post-menopausal women. You answered similar questions at the clinic, but we're hoping to get a bit more information using this interview.

Go through symptom questions on telephone interview form.

Thank you very much for your time, I really appreciate it and it's really helped us.

8.5 Appendix V Telephone Interview Form

TELEPHONE INTERVIEW (Version 2 dated Jan 2007)

UKOPS REFERENCE: _____ **DOB:** _____ _____

	When did you first have this? When did you last have this?	Did you tell your GP (or others)? If so, when?	Did you take any medication for this?	Severity when started Severity changes	Frequency
Pelvic or Abdominal pain or discomfort				Mild☐ Mod☐ Sev☐	Constant☐ Intermittent☐ 1-4 days 5-15 days 16-31 days
**Abdominal bloating (abdomen feels full & tight)				Mild☐ Mod☐ Sev☐	Constant☐ Intermittent☐ 1-4 days 5-15 days 16-31 days
**Abdominal swelling (an actual increase in abdominal size/girth)				Mild☐ Mod☐ Sev☐	Constant☐ Intermittent☐ 1-4 days 5-15 days 16-31 days
Able to feel a lump in the abdomen				Mild☐ Mod☐ Sev☐	Constant☐ Intermittent☐ 1-4 days 5-15 days 16-31 days
Back pain				Mild☐ Mod☐ Sev☐	Constant☐ Intermittent☐ 1-4 days 5-15 days 16-31 days
Indigestion				Mild☐ Mod☐ Sev☐	Constant☐ Intermittent☐ 1-4 days 5-15 days 16-31 days
Loss of appetite or feeling full quickly				Mild☐ Mod☐ Sev☐	Constant☐ Intermittent☐ 1-4 days 5-15 days 16-31 days
Nausea or vomiting				Mild☐ Mod☐ Sev☐	Constant☐ Intermittent☐ 1-4 days 5-15 days 16-31 days
Weight loss (unplanned) or appearance of weight loss				Mild☐ Mod☐ Sev☐	

Mild=Symptom easily tolerable, did not bother you much
Moderate=Symptom hard to tolerate, bothered you somewhat
Severe=Symptom intolerable, bothered you a lot.

TELEPHONE INTERVIEW (Version 2 dated Jan 2007)

UKOPS REFERENCE: _____ **DOB:** _____ _____

	When did you first have this? When did you last have this?	Did you tell your GP (or others)? If so, when?	Did you take any medication for this?	Severity when started Severity changes	Frequency
Urine (passing urine more than usual or feeling an urgent need to go)				Mild☐ Mod☐ Sev☐	Constant☐ Intermittent☐ 1-4 days 5-15 days 16-31 days
Constipation (infrequent or difficult bowel motion, hard stool)				Mild☐ Mod☐ Sev☐	Constant☐ Intermittent☐ 1-4 days 5-15 days 16-31 days
Diarrhoea (↑frequency, more liquid or loose stool)				Mild☐ Mod☐ Sev☐	Constant☐ Intermittent☐ 1-4 days 5-15 days 16-31 days
Bleeding after menopause				Mild☐ Mod☐ Sev☐	Constant☐ Intermittent☐ 1-4 days 5-15 days 16-31 days
Fatigue				Mild☐ Mod☐ Sev☐	Constant☐ Intermittent☐ 1-4 days 5-15 days 16-31 days
Leg swelling				Mild☐ Mod☐ Sev☐	Constant☐ Intermittent☐ 1-4 days 5-15 days 16-31 days
Other changes:				Mild☐ Mod☐ Sev☐	Constant☐ Intermittent☐ 1-4 days 5-15 days 16-31 days
None	Extra symptoms: *Shoulder pain* *Dizziness* *Change in taste* *Night sweats*				

Which symptom prompted you to visit the GP? _____

Route of diagnosis: _____

Mild=Symptom easily tolerable, did not bother you much
Moderate=Symptom hard to tolerate, bothered you somewhat
Severe=Symptom intolerable, bothered you a lot.

8.6 Appendix VI UKOPS Database Rules

GP Visit Coding

GP Visits:

- Surgery visits
- Home visits
- A&E visits
- Out-of-hours visits
- Telephone consultations
- NHS triage calls

Visit Exclusions:

- Visits for blood only
- Telephone calls where test results are reported ONLY (i.e. no recording of symptoms or referrals)

Each visit was categorised as either problem or routine. If more than one consultation or visit occurred on the same day, it was recorded as one visit (e.g. nurse and GP visit on same day, or telephone consultation and surgery visit on same day).

Problem Visits included:

- Visits/consultations where the main reason seemed to be for a symptom or problem, including follow-up/review of undiagnosed problems. Once problems were diagnosed, these became routine review visits.

Routine Visits included:

- Nurse visits
- Medication reviews
- Immunizations/vaccine/injections
- Asthma/hypertension/diabetes reviews
- Well women clinics (pelvic vaginal examination [PV], smear, breast examination)

If a routine and problem visit took place on the same day, problem code overrode routine. If a problem that would normally constitute a separate visit was mentioned at a routine visit, the visit was classified as a problem.

Visit data were extracted up until date of consent for controls and 'cut-off' date for cases. 'Cut-off date' was date that pelvic or ovarian aetiology was first suspected.

Symptom Coding

Any sign or symptom recorded in interview, questionnaire or in GP notes, was considered to be a symptom regardless of perceived or documented aetiology. For GP notes, symptoms were extracted from both read codes and free text. If a diagnosis was recorded without symptoms, the diagnosis was entered into the database as a 'symptom' if the condition is typically associated with key ovarian cancer symptoms (e.g. women with cystitis usually experience urinary frequency).

Relevance to ovarian cancer was determined in line with findings from previous studies and discussion with the clinical investigators. Key symptoms were those in the questionnaire list, symptoms 'other possibly related' were any symptoms not in the key symptom list that may be related to ovarian cancer. One exception to this was change in bowel habit which was included as a key symptom since this overlaps with constipation and diarrhoea. Many symptoms recorded were obviously not related to ovarian cancer (such as itchy eyes, ear ache etc.). The decision was made not to database these individually, instead they were recorded as 'other unrelated' with no details of onset date, duration, severity etcetera. If more than one unrelated symptom was mentioned on any given source (questionnaire or interview) or visit date (GP notes) they were databased as a single symptom 'other unrelated'.

The purpose of the following list was to create guidelines for databasing symptoms from the free text in GP notes and the additional symptoms on the interview and questionnaire. It was not intended to be an exhaustive list that includes every possible description of each symptom.

Symptom Definitions

Main Symptoms (Possibly Related to Ovarian Cancer)

Abdominal swelling/Increased abdominal size/abdominal distension

- Abdominal swelling due to weight gain (weight gain coded also)
- Dragging sensation in abdomen
- Hard abdomen
- Hard tummy
- Heavy solid feeling in abdomen

Abdominal bloating (includes bloating associated with eating)

- Abdominal fullness

- ‘Blown up’
- Feeling full

Abdominal/pelvic pain or discomfort

- Abdominal pain worse when eats
- Colicky abdominal pain
- Felt something moving inside tummy (possible ascites)
- Felt something moving inside (possible ascites)
- Suprapubic pain/discomfort
- Wind pains
- Upper abdominal pain
- Discomfort in hypochondrium
- Upper left quadrant pain
- Lower abdominal pain
- Iliac fossa pain
- Salpingitis

Back pain

- Flank pain
- Sciatica (many women self-report this as back pain)
- Sub-scapular/rib pain

Change in bowel habit

- Change in stool size/consistency
- Mucous in stool
- Stringy stool
- Urgency to pass motion

Constipation

- Could not pass stool
- Difficulty passing stool/difficulty passing motion
- Not emptying bowels regularly, coded as ‘constipation’ unless indication that this was a change in bowel habit
- Painful defecation
- Sensation of not emptying bowel
- Tenesmus

Diarrhoea

- Explosive watery stool
- Explosive bowel motions
- Faecal urgency
- Loose stool
- Viral gastroenteritis

Fatigue

- Lack of energy
- Lethargy/lethargic
- Malaise
- Feeling unwell
- Generally unwell
- Feels sluggish
- Feels sick
- Feels weak
- Tiredness

Indigestion

- Dyspepsia
- Heartburn
- Acid reflux/reflux
- Epigastric discomfort
- Epigastric pain (if alone) & not part of abdominal/pelvic pain
- Borborygmi
- Belching/burping
- Hiccups
- Regurgitation/regurgitation of bile
- Reflux of bile into mouth
- Oesophagitis

Loss of appetite

- Anorexia

Irregular vaginal bleeding (see also PV bleed)

- Post-coital bleeding

- Post-menopausal bleeding (on questionnaire only)

Urinary frequency or urgency

- Frequency (including nocturia)
- Urgency

Vomiting/Nausea

- Retching

Leg swelling

- Peripheral oedema
- Thigh swelling/upper thigh swelling

Other Possibly Related to Ovarian Cancer

GI Other

- Faecal incontinence
- Dampness from back passage
- IBS-like symptoms
- Flatulence
- Flatus
- Passing wind/increased wind/wind
- Variable bowel habit

Mood

- Low mood
- Anxiety
- Agitated
- Distressed
- Stress
- Depressed/seems depressed/depression
- Irritable
- Restlessness
- Panic attack

PR bleed

- Black stool

- Blood in stool
- Bleeding on defecation
- Dark blood stain in stool

PV/PR Bleeding unspecified

- Blood in toilet
- Blood on wiping
- Occasional bleeding

Difficulty breathing

- Breathlessness
- Dyspnoea
- SOB
- SOB on activity or exertion
- NOT wheezing

DVT-related

- Calf pain
- DVT
- Leg cramps
- Leg pain/tender calves/thigh pain
- Leg spasms

Urinary other

- Dysuria
- Incontinence (all types)/leakage/stress/latch key/urge
- UTI/Coliform UTI/possible infection/Cystitis
- Retention/Could not pass urine/difficulty passing urine/poor stream/reduced urine output/retention with full bowel Outflow obstruction Nocturnal difficult passing urine
- Bladder pressure symptoms
- Dark urine/concentrated urine/darker colour & smell change
- Smelly urine
- Haematuria/blood in urine
- Urinary symptoms
- Nocturnal enuresis

N.B. Although all women with urge incontinence have urgency, if they have this diagnosis they would have had investigations to demonstrate that they had detrusor instability, so this symptom was not included as urinary urgency.

Pain Radiation

- Buttock to leg pain
- Hip to leg pain

Utero-Vaginal Prolapse

- Bulge in vagina
- Cystocoele
- Felt as if pelvis was going to fall out
- Sensation of lump between legs/vaginal lump/vulval lump
- Sinking feeling in pelvis
- Vaginal fullness

Sleep disorders

- Insomnia
- Poor sleep
- Sleep disturbance

Groin/Loin Pain

- Groin pain
- Loin pain/ache

Other miscellaneous

- Painful intercourse
- Vaginal discharge
- Ankle swelling
- Dysmenorrhoea
- Hip pain
- Lump in neck (could be lymph node related in advanced cases)
- Menorrhagia
- Painful lower ribs
- PE
- Pleurisy
- Weight gain (including from ascites as women may think weight gain)

Other Unrelated to Ovarian Cancer

Vasovagal

- Black out
- Black out
- Collapse
- Cold & clammy
- Dizziness/dizzy
- Faint
- Feels faint
- Giddy/giddiness
- Vertigo
- Wobbly
- Acne/spotty skin
- Asymmetry of breasts
- Bad taste in mouth
- Bleeding under skin on leg
- Boil of vulva
- Breast lump
- Breast pain
- Breast swelling
- Change in taste
- Chest pain/ache lower R chest/lower chest pain
- Cough
- Dysphagia/oesophageal dysmotility/solid food stuck in oesophagus
- Eczematous rash on leg
- Elbow pain
- Fever
- Genitourinary system diseases (too unspecified to include)
- Hair falling out more
- Headache
- Hot flushes
- Itchy nipples
- Knee swelling (if knee ONLY) - leg swelling if includes leg/right knee swelling
- Leg rash

- Lump over upper left quadrant
- Lump right groin
- Memory loss
- Migraine
- Neck pain
- Nerve pain in leg
- Night sweats/sweats
- Phlebitis
- Pins & needles fingers/arms/legs
- PR lump
- Rectal pain
- Rectal mass
- Recurrent nose bleeds
- Retrosternal pain
- Rib pain
- Shivers
- Shoulder pain/pain between shoulder blades
- Spotty skin
- Submandibular mass
- Surface fatty lump on abdomen
- Thick coating on tongue
- Vaginal cyst
- Vaginal irritation
- Vaginal itch
- Vaginal pain when bowels open
- Vulval irritation
- Vulval itch
- Vulval pain
- Vulval rash
- Weight loss (intentional)

Major Symptom Groups

Abdominal Symptoms

- Pelvic/Abdominal pain or discomfort
- Abdominal bloating

- Increased abdominal size/abdominal swelling

GI Symptoms

- Change in bowel habit
- Constipation
- Diarrhoea
- GI other (flatulence, variable bowel habit, IBS symptoms)
- Indigestion
- Nausea/vomiting

Gynaecological Symptoms

- Irregular vaginal bleeding
- Postmenopausal bleeding
- Pain with intercourse
- Vaginal discharge

Urinary Symptoms

- Urinary frequency or urgency
- Dysuria
- Incontinence (all types)/leakage/stress/latch key/urge
- UTI/Coliform UTI/possible infection/Cystitis
- Retention/Could not pass urine/difficulty passing urine/poor stream/reduced urine output/retention with full bowel Outflow obstruction Nocturnal difficult passing urine
- Bladder pressure symptoms
- Dark urine/concentrated urine/darker colour & smell change
- Smelly urine
- Haematuria/blood in urine
- Urinary symptoms
- Nocturnal enuresis

General Symptoms

- Fatigue
- Loss of appetite
- Weight loss

Goff Index Symptoms

Increased Abdominal size/Bloating

- Abdominal bloating
- Increased abdominal size/abdominal swelling

Pelvic or abdominal pain

- Pelvic or abdominal pain as defined in list above

Difficulty eating or feeling full quickly

- Loss of appetite (interview data only since feeling full quickly was only specifically asked about on interview)

Dates

- Dates with only month and year but no day, were entered as the 15th of the month. E.g. Oct 05 was entered as 15 Oct 05.
- Symptom onset dates in GP notes were derived based on visit date and symptom duration if duration information was available. E.g. 'started 5/7' would have an onset date of 5 days before the visit date entered.
- Date ranges where two months were given were entered as the 1st day of the 2nd month e.g. Oct 05 – Nov 05 was entered as 01 Nov 05.
- Date ranges where >2 months were given were entered as the midpoint e.g. Jul 05 to Sep 05 midpoint would be 15th Aug 05; Jul 05 to Oct 05 would be 1st Sep 05.
- If duration was >2 years coded as '25' months
- Source date of clinic letters taken as date of clinic visit not date of letter typing
- Date of diagnosis for other significant co morbidities was taken as first date of diagnosis that appeared in notes
- Cut-off-date is date when ovarian or pelvic aetiology first suspected (includes '?pelvic mass')
- If additional text implied that dates were definitely not midpoint of the month then entry of a more sensible default date was permitted. E.g. 'Early March' was coded as 01 Mar 06
- Nonsensical symptom onset dates and 'dates GP told' on self-reported data were amended if sensible. Otherwise 'date GP told' was changed to missing. E.g. one of the case questionnaires was completed in October

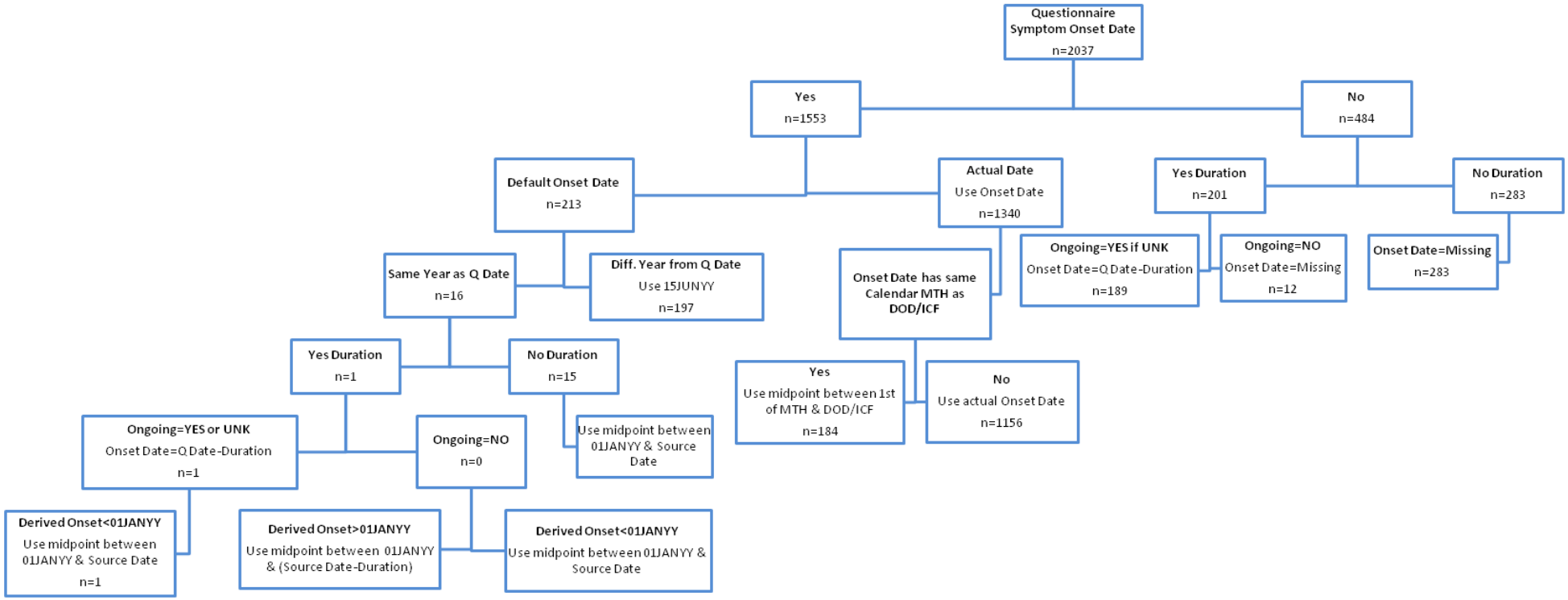
2006, with a symptom whose onset was June 2006, duration 4 months, and 'date GP told' May 2006. In this situation, the onset date was kept since it matched the duration, and the 'date GP told' was discarded.

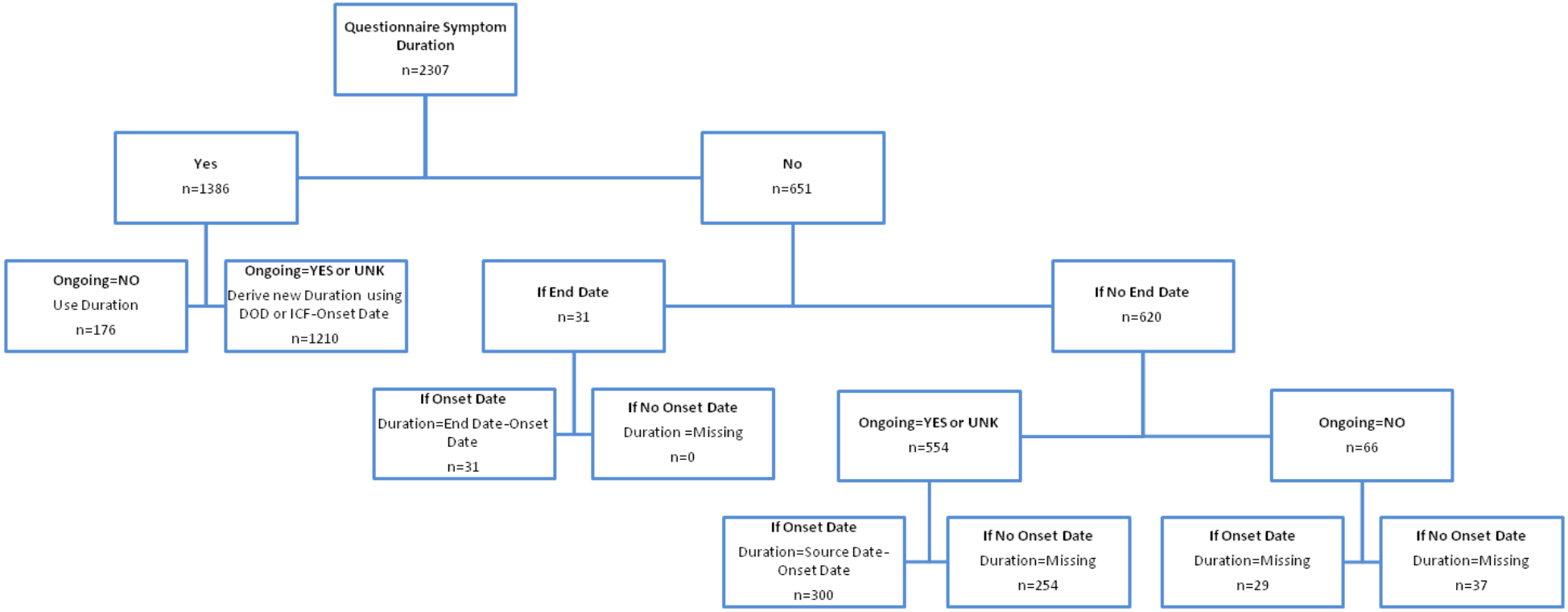
- Some questionnaires answers contained text where numbers were expected. These were dealt with in the following manner:
    - Frequency:
        - 'Most days' = 16-31 days
        - 'All' = 16-31 days
        - 'Sometimes' = missing
        - 'Variable' = missing
        - 'Couple of weeks without' = 5-15 days
    - Duration
        - '>1 year' = 13 months
- For unknown CA125 and US test dates – Date of earliest letter mentioning results was entered

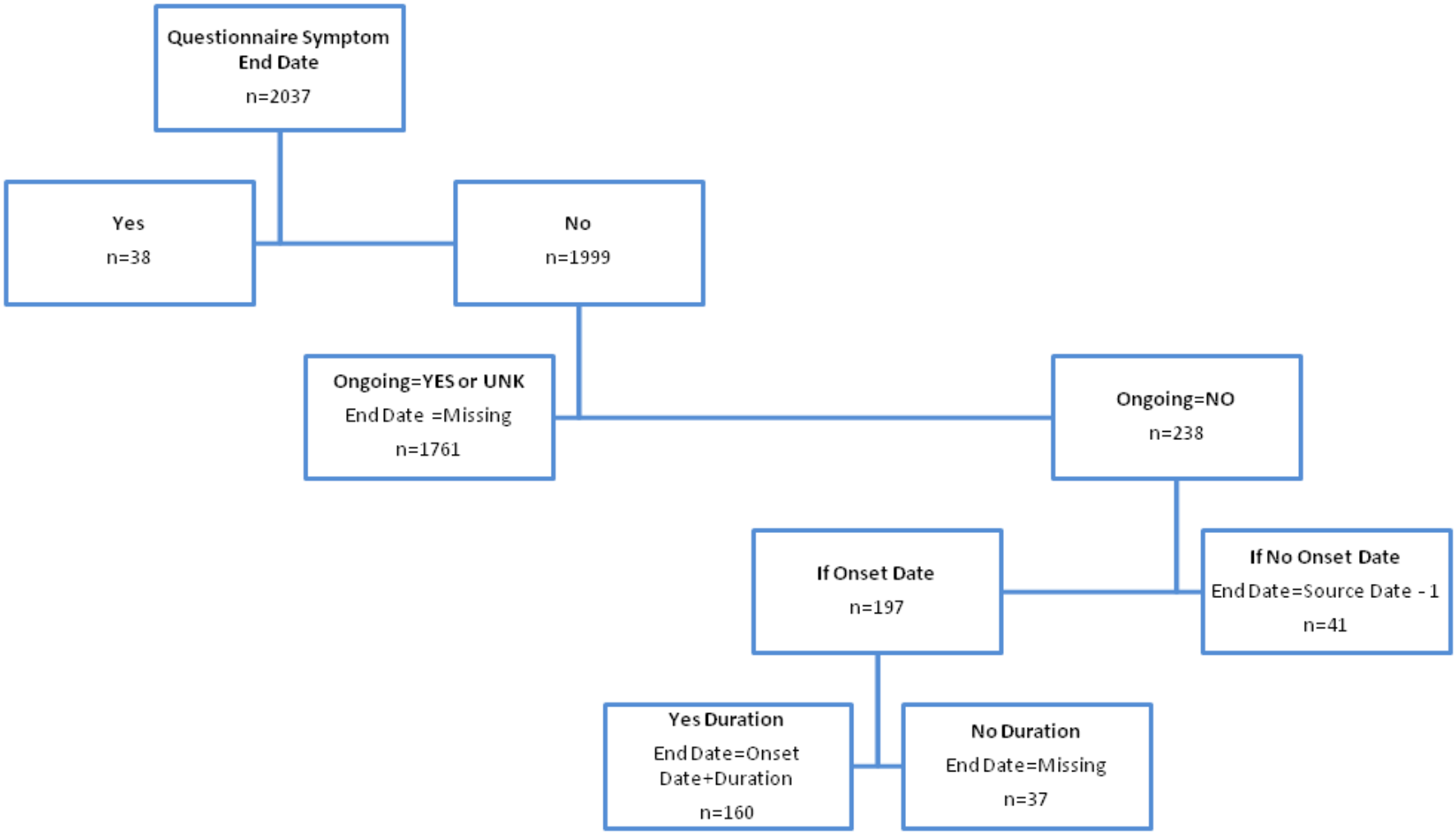
General Rules

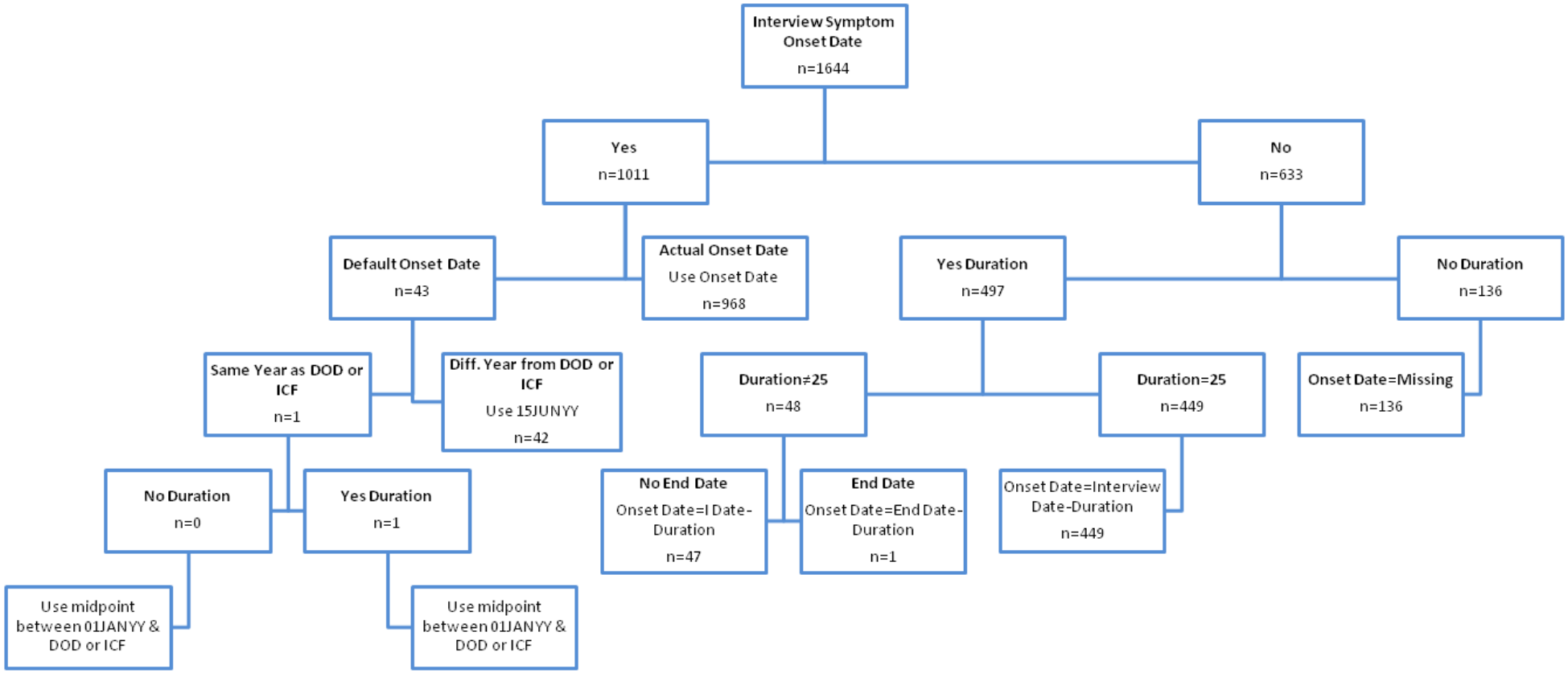
- Symptom ongoing status was assumed to be 'yes' if ongoing was missing.
- If a start date and duration were nonsensical, the start date would be trusted and a new duration was derived provided the symptom was ongoing. Derived start dates were based on date of diagnosis (for cases) and consent (for controls).
- If only a start year was provided, the midpoint between January of the start year and the date of diagnosis (or consent) was taken.
- Used midpoint for ranges where a single entry was required
- Record of blood samples taken at GP were linked to visits even if not actually taken on day of visit/request if within 5 days of visit.
- If GP was asked to order CA125 this was not recorded as GP ordered CA125
- Only databased women as 'postmenopausal' if there was documentation in the notes to reflect this (e.g. >1 year amenorrhea)
- Nurse visits for changing dressing for a problem code as routine unless a problem occurs with wound
- If no pre-treatment CA125 results were available a default value of '9999' was entered if levels were known to be abnormal or raised
- If women provided reasons for weight loss such as stress or anxiety, these were also recorded as symptoms.

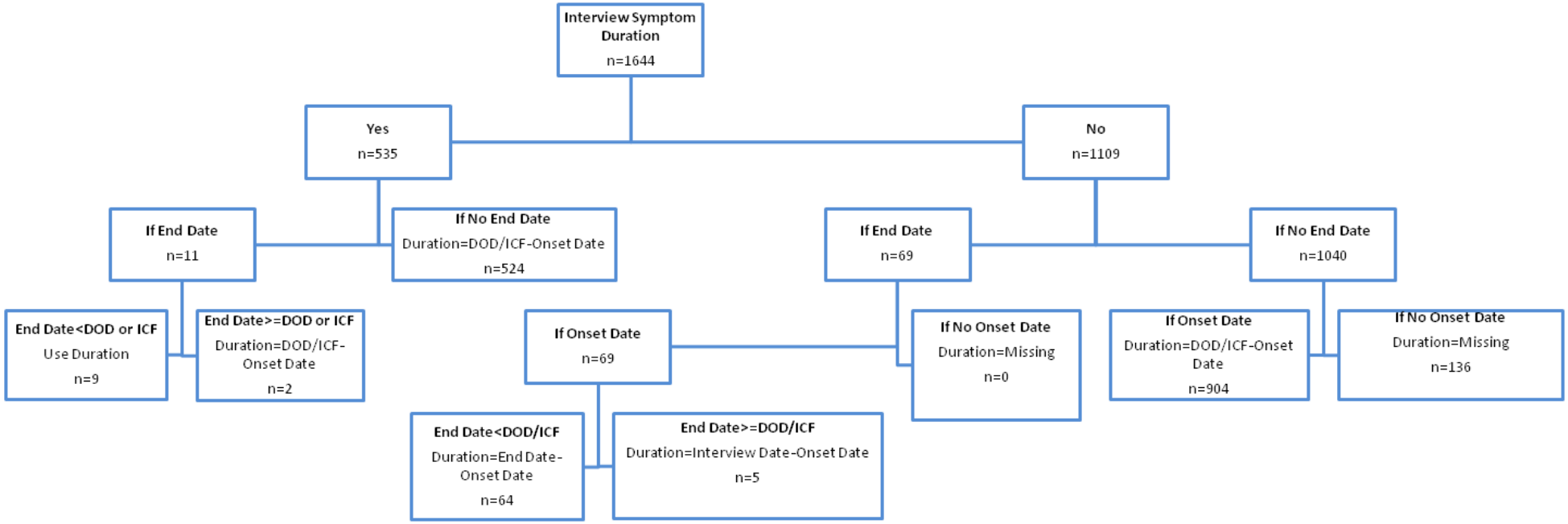
8.7 Appendix VII Data Cleaning Flowcharts

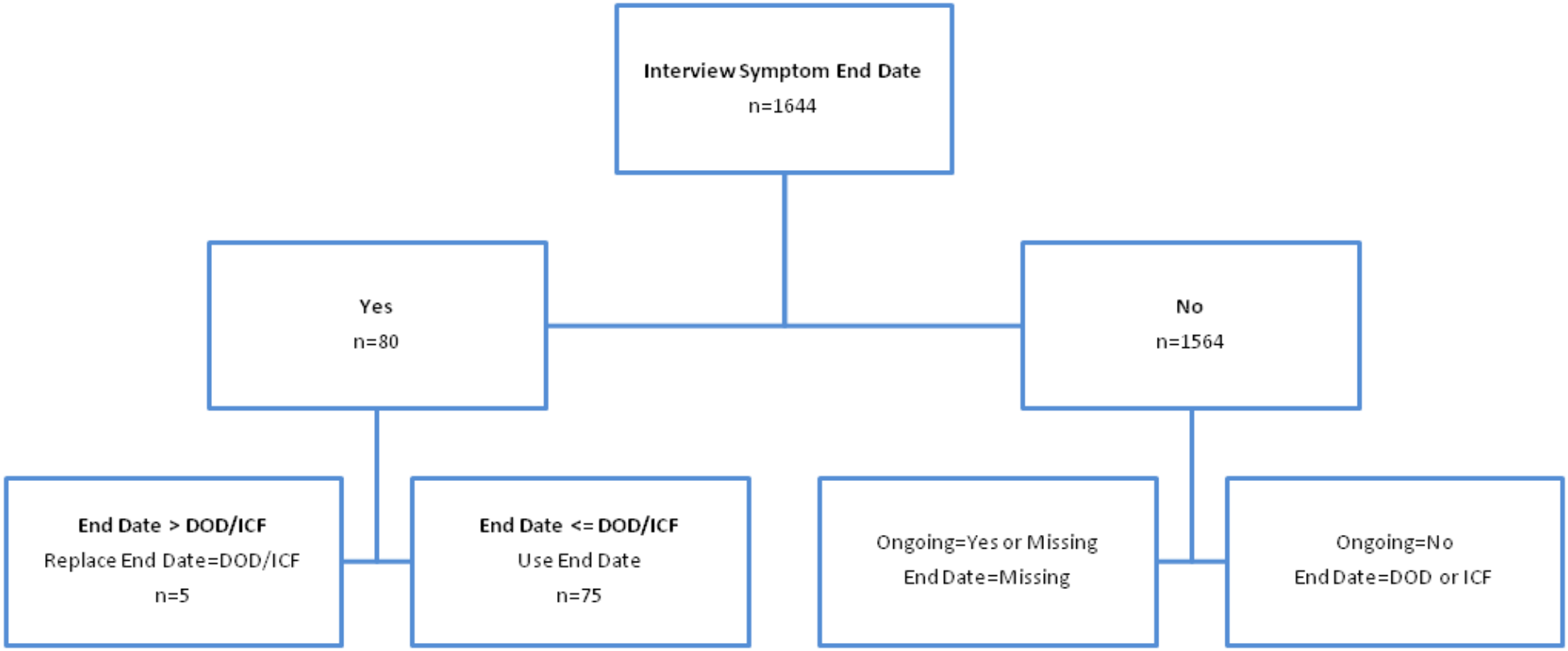



Questionnaire Symptom Duration
n=2307
Yes
n=1386
No
n=651
Ongoing=NO
Use Duration
n=176
Ongoing=YES or UNK
Derive new Duration using DOD or ICF-Onset Date
n=1210
If End Date
n=31
If No End Date
n=620
If Onset Date
Duration=End Date-Onset Date
n=31
If No Onset Date
Duration =Missing
n=0
Ongoing=YES or UNK
n=554
Ongoing=NO
n=66
If Onset Date
Duration=Source Date-Onset Date
n=300
If No Onset Date
Duration=Missing
n=254
If Onset Date
Duration=Missing
n=29
If No Onset Date
Duration=Missing
n=37

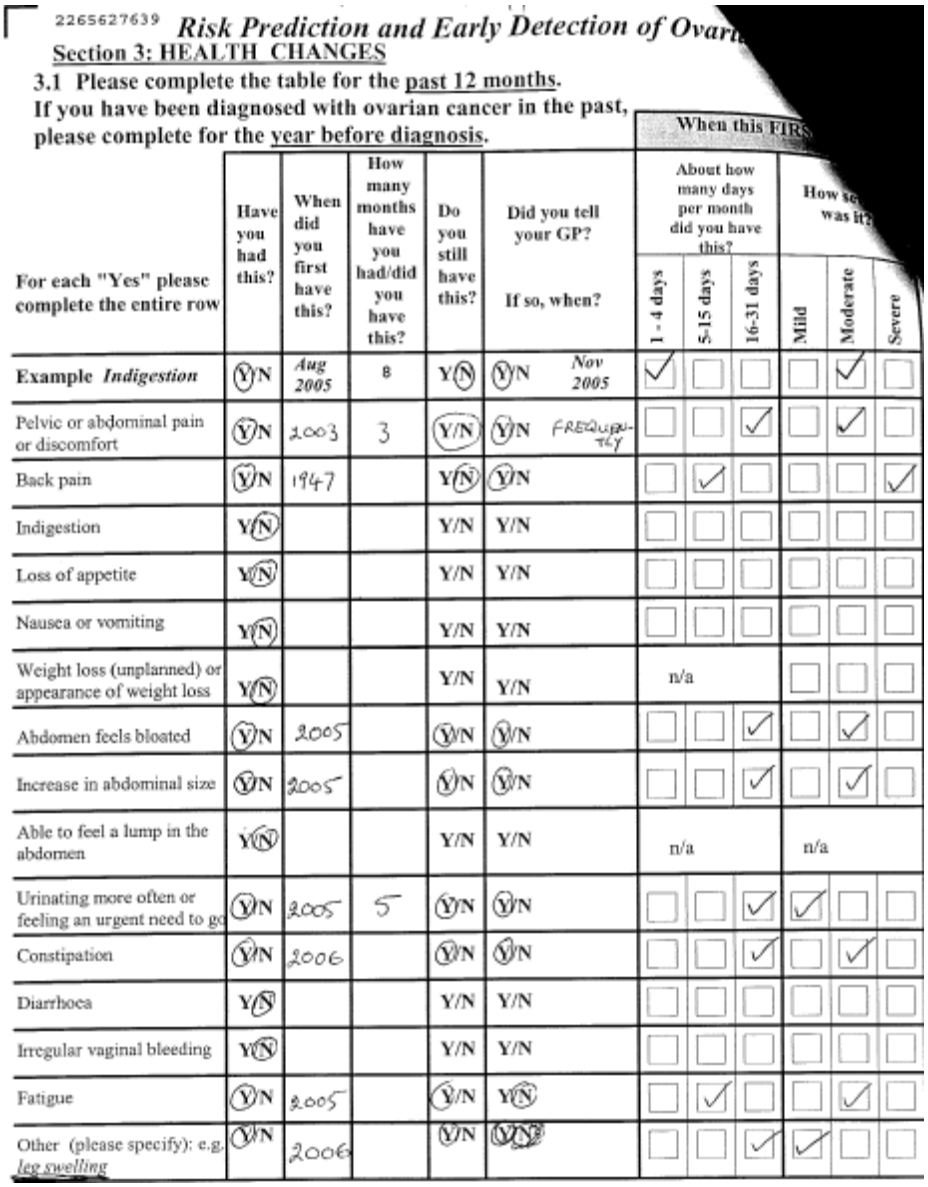

Questionnaire Symptom End Date
n=2037
Yes
n=38
No
n=1999
Ongoing=YES or UNK
End Date =Missing
n=1761
Ongoing=NO
n=238
If Onset Date
n=197
If No Onset Date
End Date=Source Date - 1
n=41
Yes Duration
End Date=Onset Date+Duration
n=160
No Duration
End Date=Missing
n=37


Interview Symptom Onset Date
n=1644
Yes
n=1011
No
n=633
Default Onset Date
n=43
Actual Onset Date
Use Onset Date
n=968
Yes Duration
n=497
No Duration
n=136
Same Year as DOD or ICF
n=1
Diff. Year from DOD or ICF
Use 15JUNYY
n=42
Duration≠25
n=48
Duration=25
n=449
Onset Date=Missing
n=136
No Duration
n=0
Yes Duration
n=1
No End Date
Onset Date=I Date-Duration
n=47
End Date
Onset Date=End Date-Duration
n=1
Onset Date=Interview Date-Duration
n=449
Use midpoint between 01JANYY & DOD or ICF
Use midpoint between 01JANYY & DOD or ICF


Interview Symptom Duration
n=1644
Yes
n=535
No
n=1109
If End Date
n=11
If No End Date
Duration=DOD/ICF-Onset Date
n=524
If End Date
n=69
If No End Date
n=1040
End Date<DOD or ICF
Use Duration
n=9
End Date>=DOD or ICF
Duration=DOD/ICF-Onset Date
n=2
If Onset Date
n=69
If No Onset Date
Duration=Missing
n=0
If Onset Date
Duration=DOD/ICF-Onset Date
n=904
If No Onset Date
Duration=Missing
n=136
End Date<DOD/ICF
Duration=End Date-Onset Date
n=64
End Date>=DOD/ICF
Duration=Interview Date-Onset Date
n=5


Interview Symptom End Date
n=1644
Yes
n=80
No
n=1564
End Date > DOD/ICF
Replace End Date=DOD/ICF
n=5
End Date <= DOD/ICF
Use End Date
n=75
Ongoing=Yes or Missing
End Date=Missing
Ongoing=No
End Date=DOD or ICF

8.8 Appendix VIII Example Questionnaire



2265627639 *Risk Prediction and Early Detection of Ovari*

Section 3: HEALTH CHANGES

3.1 Please complete the table for the past 12 months.
If you have been diagnosed with ovarian cancer in the past, please complete for the year before diagnosis.

For each "Yes" please complete the entire row	Have you had this?	When did you first have this?	How many months have you had/did you have this?	Do you still have this?	Did you tell your GP? If so, when?	When this FIRS… About how many days per month did you have this? 1 - 4 days	5-15 days	16-31 days	How se… was it? Mild	Moderate	Severe
Example ***Indigestion***	(Y)N	*Aug 2005*	8	Y(N)	(Y)N *Nov 2005*	✓				✓	
Pelvic or abdominal pain or discomfort	(Y)N	2003	3	(Y/N)	(Y)N FREQUENTLY			✓		✓	
Back pain	(Y)N	1947		Y(N)	(Y)N		✓				✓
Indigestion	Y(N)			Y/N	Y/N						
Loss of appetite	Y(N)			Y/N	Y/N						
Nausea or vomiting	Y(N)			Y/N	Y/N						
Weight loss (unplanned) or appearance of weight loss	Y(N)			Y/N	Y/N	n/a					
Abdomen feels bloated	(Y)N	2005		(Y)N	(Y)N			✓		✓	
Increase in abdominal size	(Y)N	2005		(Y)N	(Y)N			✓		✓	
Able to feel a lump in the abdomen	Y(N)			Y/N	Y/N	n/a			n/a		
Urinating more often or feeling an urgent need to go	(Y)N	2005	5	(Y)N	(Y)N			✓	✓		
Constipation	(Y)N	2006		(Y)N	(Y)N			✓		✓	
Diarrhoea	Y(N)			Y/N	Y/N						
Irregular vaginal bleeding	Y(N)			Y/N	Y/N						
Fatigue	(Y)N	2005		(Y)N	Y(N)		✓			✓	
Other (please specify): e.g. *leg swelling*	(Y)N	2006		(Y)N	(Y)N			✓	✓		

8.9 Appendix IX Worked Example of Continuation Odds Ratio

The following is an example of how the continuation odds ratio for abdominal bloating 3-5 months before diagnosis was calculated for the questionnaire data (in Table 4-7). A total of 34 cases, reported the onset of (new) abdominal bloating between 3-5 months prior to diagnosis.

Table 8-1 Number Cases and Controls with Abdominal Bloating on Questionnaire at Different Periods Before Reference Date

Months Before Reference Date	Cases	Controls
None	66	214
0-2	82	1
3-5	***34***	***2***
6-8	16	2
9-11	9	4
12-14	15	3
15-23	3	3
>2 years	11	18
Missing start date	13	21
TOTAL	**249**	**268**

Normally, an odds ratio for case-control symptom differences would be calculated as:

$$OR = \frac{odds\ of\ having\ symptom\ in\ cases}{odds\ of\ having\ symptom\ in\ controls}$$

This is equivalent to:

$$OR = \frac{cases\ with\ symptom \times controls\ without\ symptom}{cases\ without\ symptom \times controls\ without\ symptom}$$

Below are the calculations for obtaining the average number of controls 'at risk' of still getting the symptom at 3-5 months prior to diagnosis. Women who had abdominal bloating for >2 years are excluded (n=18), as are women with missing onset dates since they are treated as if their symptom started >2 years ago (n=21). Hence, the sum of these women is subtracted from total number of controls (n=268), which leaves the total number of controls who are 'at risk' of developing bloating for the first time over the 2 years prior to reference date. The total number of months that have passed from 2 years prior to reference date to 3-5 months (n=21) is multiplied by the average (control) rate of bloating per month. This is subtracted from the total number of controls who are 'at risk' of developing bloating for the first time in the 2 years prior to reference date.

$$Average\ no.of\ controls\ 'at\ risk'\ = (268 - 39) - \left(21 \times \left(\frac{15.5}{24}\right)\right)$$

$$= 215.4375$$

The number of cases ‘at risk’ is simply the number of cases who have not yet had bloating by 3-5 months prior to diagnosis.

$$No.of\ cases\ 'at\ risk'\ = (66 + 82)$$

$$= 148$$

Therefore, the continuation odds ratio (cOR) is calculated as:

$$cOR = \frac{no.of\ cases\ with\ the\ symptom\ \times average\ no.of\ controls\ 'at\ risk'(i.e.not\ yet\ reported\ the\ symptom)}{No.of\ cases\ 'at\ risk'(i.e.not\ yet\ reported\ the\ symptom)\ \times estimated\ average\ no.of\ controls\ with\ symptom\ at\ 3-5\ months}$$

$$= \frac{34 \times 215.4375}{148 \times 1.9375}$$

$$= 25.54$$

8.10 Appendix X GP Workload Symptom Lists

Table 8-2 GP Workload Symptoms

Chrisp Street	Woosehill	Northcroft	Boundary
Abdominal Bloating	Abdominal Discomfort	Abdominal Bloating	Abdominal
Abdominal Cramp	Abdominal Distension	Abdominal Distension	Bloating
Abdominal Discomfort	Abdominal Pain	Abdominal Pain	Abdominal Mass
Abdominal Distension	Acid Reflux	Acid Reflux	Abdominal Pain
Abdominal Pain	Bloated	Anorexia	Abdominal
Acute Cystitis	Bloating	Belching	Swelling
Bowel Symptoms	Borborygmi	Bowel Frequency	Acid Reflux
Constipation	Change In Stool	Bowel Irregular	Belching
Cystitis Symptoms	Constipation	Bowel Urgency	Cystitis
Diarrhoea	Diarrhoea	Change Bowel Habit	Diarrhoea
Difficulty Eating	Difficulty Emptying	Constipation	Epigastric Pain
Dyspepsia	Bowel	Diarrhoea	Fatigue
Epigastric Discomfort	Difficulty Passing	Dyspepsia	Feels Unwell
Epigastric Pain	Urine	Epigastric Discomfort	Felt Unwell
Fatigue	Fatigue	Epigastric Pain	GI Problems
Feeling Full	Feels Unwell	Fatigue	Indigestion
Feeling Unwell	Fullness In	Feels Unwell	Loss Of Appetite
Feels Strange	Epigastrium	Flatulence	Nausea
Feels Unwell	Heartburn	Food Regurgitation	Urinary Frequency
Feels Weak	Incomplete Evacuation	Heartburn	Urinary Symptoms
Gastric Problems	Bowel	Increased Bowel	Urinary Urgency
Gastric Symptoms	Indigestion	Frequency	UTI
GI Symptoms	Lethargic	Indigestion	Vomiting
Heartburn	Loss Of Appetite	Loose Bowel	Weight Loss
Indigestion	Low Energy	Loose Stool	
Loose Stool	Lower Abdominal Pain	Loss Of Appetite	
Loose Stools	Nausea	Low Energy	
Loss Of Appetite	Painful Intercourse	Lower Abdominal Pain	
Low Energy	Postmenopausal	Nausea	
Lower Abdominal Pain	bleeding	Pain On Eating	
Malaise	Post-Coital Bleeding	Postmenopausal	
Nausea	Reflux	bleeding	
Pain When Eating	Soft Stools	Recurrent UTI	
Postmenopausal	Tired	Sticky Stool	
bleeding	Urinary Frequency	Tenesmus	
Reflux	Urinary Retention	Tired	
Reflux Symptom	Urinary Symptoms	Upper Abdominal Pain	
Suprapubic Discomfort	Urinary Urgency	Urinary Frequency	
Tired	UTI Symptoms	Urinary Urgency	
Urinary Frequency	Vaginal Discharge	Vaginal Discharge	
Urinary Symptoms	Variable Bowel Habit	Variable Bowel Habit	
Urinary Urgency	Vomiting	Vomiting	
Vaginal Discharge	Watery Stools	Weight Loss	
Vomiting	Weight Loss	Wind	
Weight Loss	Wind		

Table 8-3 GP Workload Symptom Groups

Chrisp Street	Woosehill	Northcroft	Boundary
Abdominal			
Abdominal bloating Abdominal cramp Abdominal discomfort Abdominal distension Abdominal pain Lower abdominal pain Feeling full Suprapubic discomfort	Abdominal discomfort Abdominal distension Abdominal pain Abdominal bloating/bloated Fullness in epigastrium Lower abdominal pain	Abdominal bloating Abdominal pain Abdominal distension Lower abdominal pain Upper abdominal pain	Abdominal bloating Abdominal mass Abdominal swelling Abdominal pain
GI			
Bowel symptoms Constipation Diarrhoea Difficulty eating Dyspepsia Epigastric pain/discomfort GI symptoms Gastric problems Gastric symptoms Heartburn Indigestion Loose stool/loose stools Nausea Pain when eating Reflux/reflux symptoms Vomiting	Acid reflux/reflux Borborygmi Constipation Diarrhoea Difficulty emptying bowel Heartburn Incomplete evacuation of bowel Indigestion Nausea Soft stools Variable bowel habit Vomiting Watery stools Wind Change in stool	Acid reflux Anorexia Belching Bowel Frequency Bowel irregular Bowel urgency Change in bowel habit Constipation Diarrhoea Flatulence Food regurgitation Heartburn Increased bowel frequency Indigestion Loose bowel Loose stool Loss of appetite Nausea Pain on eating	Acid reflux Belching Diarrhoea GI problems Indigestion Loss of appetite Nausea Vomiting Epigastric pain

Chrisp Street	Woosehill	Northcroft	Boundary
GI Continued			
		Sticky stool Variable bowel habit Weight loss Vomiting Wind Epigastric pain/discomfort	
Urinary			
Urinary urgency Urinary frequency Acute cystitis/cystitis Urinary symptoms	Urinary retention Difficulty passing urine Urinary frequency Urinary urgency Urinary symptoms UTI symptoms	Urinary frequency Urinary urgency Recurrent UTI	Urinary frequency Urinary urgency Urinary symptoms UTI Cystitis
Gynaecological			
Postmenopausal bleeding Vaginal discharge	Painful intercourse Post-coital bleeding Postmenopausal bleeding Vaginal discharge	Vaginal discharge Postmenopausal bleeding	-
General			
Fatigue Feels strange Feeling feels/ unwell Feels weak Low energy Malaise Tired Loss of appetite Weight loss	Fatigue Feels unwell Lethargic Low energy Tired Loss of appetite Weight loss	Fatigue Feels unwell Low energy Tired Anorexia Loss of appetite Weight Loss	Feels unwell/felt unwell Loss of appetite Weight loss

Table 8-4 Consensus Symptoms

Chrisp Street	Woosehill	Northcroft	Boundary
Bloating			
Abdominal bloating Feeling full Abdominal distension	Bloating Fullness in epigastrium Abdominal distension	Bloating Abdominal distension	Abdominal bloating Abdominal swelling
Pelvic or Abdominal Pain			
Abdominal discomfort Abdominal cramp Abdominal pain Suprapubic discomfort Lower abdominal pain	Abdominal pain Lower abdominal pain	Abdominal pain Lower abdominal pain Upper abdominal pain	Abdominal pain
Difficulty Eating or Feeling Full Quickly			
Difficulty eating	-	-	-
Urinary Frequency/Urgency			
Urinary frequency Urinary urgency	Urinary frequency Urinary urgency	Urinary frequency Urinary urgency	Urinary frequency Urinary urgency

8.11 Appendix XI Innovations & Progress in Healthcare for Women Conference Abstract (7-8th April 2008)

Anita Lim

A Case-Control Study to Investigate Symptoms & Events Preceding Ovarian Cancer Diagnosis - Preliminary Results from an Interim Analysis

Anita Lim[1], Aleksandra Gentry-Maharaj[2], Nyala Balogun[2], David Mesher[1], Ian Jacobs[2], Peter Sasieni[1], Usha Menon[2]
[1]Cancer Research UK Centre for Epidemiology, Mathematics and Statistics
Wolfson Institute of Preventive Medicine, Barts and The London School of Medicine
[2]Gynaecological Oncology, UCL Elizabeth Garrett Institute of Women's Health

Background
Some women experience vague, non-specific symptoms for months before ovarian cancer diagnosis. The identification of a sensitive and specific symptom profile with sufficient lead-time to improve survival is an exciting prospect. However, the prevalence of such symptoms in postmenopausal women in the general population is likely to be high. This symptoms study is nested within UKOPS (United Kingdom Ovarian Cancer Population Study) a multi-centre study focussed on biological endpoints.

Aim
To quantify the potential for the diagnosis of ovarian cancer to be brought forward via early symptom recognition

Methods
Women aged ≥45 awaiting treatment for primary invasive ovarian cancer (cases) and women from the multimodal arm of the UKCTOCS screening study were asked to complete a symptom questionnaire. Telephone interviews to further probe symptoms were performed in a subset of cases and controls and GP notes were requested for the two year period before diagnosis or recruitment for both groups.

Results
Preliminary analysis from the first 96 cases and 56 controls showed that both cases and controls consistently report more symptoms in telephone interview than on questionnaire. This disparity was particularly marked in the control group. Bloating (odds ratio: 10.5; 95%CI, 4.8 to 23.2) and fatigue (OR 8.5; 95%CI, 3.5 to 20.7) were related to ovarian cancer in the questionnaire data, but not in the interview data (OR 1.9; 95%CI, 0.9 to 4.2, and 1.8; 95%CI, 0.8 to 3.9 respectively). Symptom prevalence recorded in GP notes for 3-6 months before diagnosis in cases was modest (abdominal / pelvic pain or discomfort (21%) bloating (5%), swelling (1%)). Anecdotal evidence from telephone interviews indicated that there are various semantic issues with symptom reporting that may contribute to the discrepancies observed between questionnaire and interview data.

Conclusions
These data suggest that reliance on symptoms reported on questionnaire alone may produce potentially misleading results. Early indications are that reducing the time between first presentation to the GP and diagnosis would make a difference in at most a quarter of cases.

9 References

1. Office for National Statistics (ONS). *Cancer Atlas of the United Kingdom and Ireland 1991-2000*. 2005 [cited 23.11.06]; Available from: http://www.statistics.gov.uk/downloads/theme_health/caUKI91_00/Ch18_Ovary.pdf.

2. Cancer Research UK. *CancerStats - Ovarian Cancer Survival Statistics*. 2006 March 2006 [cited 24.11.06]; Available from: http://info.cancerresearchuk.org:8000/cancerstats/types/ovary/survival/.

3. Cooper, N., M.J. Quinn, B. Rachet, E. Mitry, et al., *Survival from cancer of the ovary in England and Wales up to 2001.* British Journal of Cancer, 2008. **99 Suppl 1**: p. S70-2.

4. Kosary, C.L., *FIGO stage, histology, histologic grade, age and race as prognostic factors in determining survival for cancers of the female gynecological system: an analysis of 1973-87 SEER cases of cancers of the endometrium, cervix, ovary, vulva, and vagina.* Seminars in Surgical Oncology, 1994. **10**(1): p. 31-46.

5. Bell, R., M. Petticrew, and T. Sheldon, *The performance of screening tests for ovarian cancer: results of a systematic review.[see comment].* British Journal of Obstetrics & Gynaecology, 1998. **105**(11): p. 1136-47.

6. Vine, M.F., R.B. Ness, B. Calingaert, J.M. Schildkraut, et al., *Types and duration of symptoms prior to diagnosis of invasive or borderline ovarian tumor.* Gynecologic Oncology, 2001. **83**(3): p. 466-71.

7. Vine, M.F., B. Calingaert, A. Berchuck, and J.M. Schildkraut, *Characterization of prediagnostic symptoms among primary epithelial ovarian cancer cases and controls.* Gynecologic Oncology, 2003. **90**(1): p. 75-82.

8. Goff, B.A., L. Mandel, H.G. Muntz, and C.H. Melancon, *Ovarian carcinoma diagnosis.* Cancer, 2000. **89**(10): p. 2068-75.

9. Attanucci, C.A., H.G. Ball, S.L. Zweizig, and A.H. Chen, *Differences in symptoms between patients with benign and malignant ovarian neoplasms.* American Journal of Obstetrics & Gynecology, 2004. **190**(5): p. 1435-7.

10. Friedman, G.D., J.S. Skilling, N.V. Udaltsova, and L.H. Smith, *Early symptoms of ovarian cancer: a case-control study without recall bias.* Fam Pract, 2005.

11. Kirwan, J.M., D.G. Tincello, J.J. Herod, O. Frost, et al., *Effect of delays in primary care referral on survival of women with epithelial ovarian cancer: retrospective audit.[see comment].* BMJ, 2002. **324**(7330): p. 148-51.

12. Lataifeh, I., D.E. Marsden, G. Robertson, V. Gebski, et al., *Presenting symptoms of epithelial ovarian cancer.* Australian & New Zealand Journal of Obstetrics & Gynaecology, 2005. **45**(3): p. 211-4.

13. Paulsen, T., J. Kaern, K. Kjaerheim, C. Trope, et al., *Symptoms and referral of women with epithelial ovarian tumors.* International Journal of Gynaecology & Obstetrics, 2005. **88**(1): p. 31-7.

14. Thulesius, H.O., A.C. Lindgren, H.L. Olsson, and A. Hakansson, *Diagnosis and prognosis of breast and ovarian cancer--a population-based study of 234 women.* Acta Oncologica, 2004. **43**(2): p. 175-81.

15. Wikborn, C., F. Pettersson, C. Silfversward, and P.J. Moberg, *Symptoms and diagnostic difficulties in ovarian epithelial cancer.* International Journal of Gynaecology & Obstetrics, 1993. **42**(3): p. 261-4.

16. Yawn, B.P., B.A. Barrette, and P.C. Wollan, *Ovarian cancer: the neglected diagnosis.* Mayo Clinic Proceedings, 2004. **79**(10): p. 1277-82.

17. Smith, E.M. and B. Anderson, *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): p. 2727-32.

18. Wikborn, C., F. Pettersson, and P.J. Moberg, *Delay in diagnosis of epithelial ovarian cancer.* International Journal of Gynaecology & Obstetrics, 1996. **52**(3): p. 263-7.

19. Flam, F., N. Einhorn, and K. Sjovall, *Symptomatology of ovarian cancer.* Eur J Obstet Gynecol Reprod Biol, 1988. **27**(1): p. 53-7.

20. Smith, L.H., C.R. Morris, S. Yasmeen, A. Parikh-Patel, et al., *Ovarian cancer: Can we make the clinical diagnosis earlier?* Cancer, 2005.

21. Wynn, M.L., S. Chang, and L.A. Peipins, *Temporal patterns of conditions and symptoms potentially associated with ovarian cancer.* Journal of Women's Health, 2007. **16**(7): p. 971-86.

22. Kroenke, K. and L. Harris, *Symptoms research: a fertile field.* Annals of Internal Medicine, 2001. **134**(9 Pt 2): p. 801-2.

23. Kroenke, K., M.E. Arrington, and A.D. Mangelsdorff, *The prevalence of symptoms in medical outpatients and the adequacy of therapy.[see comment].* Archives of Internal Medicine, 1990. **150**(8): p. 1685-9.

24. Kroenke, K., *Studying symptoms: sampling and measurement issues.[see comment].* Annals of Internal Medicine, 2001. **134**(9 Pt 2): p. 844-53.

25. Verbrugge, L.M. and F.J. Ascione, *Exploring the iceberg. Common symptoms and how people care for them.* Medical Care, 1987. **25**(6): p. 539-69.

26. Miller, F.G. and D.L. Rosenstein, *The nature and power of the placebo effect.[see comment].* Journal of Clinical Epidemiology, 2006. **59**(4): p. 331-5.

27. Wallander, M.A., E. Dimenas, K. Svardsudd, and I. Wiklund, *Evaluation of three methods of symptom reporting in a clinical trial of felodipine.* European Journal of Clinical Pharmacology, 1991. **41**(3): p. 187-96.

28. Funch, D.P., *Predictors and consequences of symptom reporting behaviors in colorectal cancer patients.* Medical Care, 1988. **26**(10): p. 1000-8.

29. Robb, K.A., A. Miles, J. Campbell, P. Evans, et al., *Can cancer risk information raise awareness without increasing anxiety? A randomized trial.* Preventive Medicine, 2006. **43**(3): p. 187-90.

30. Office for National Statistics, *Cancer 1971-1997 [CD Rom]*. 1999: London.

31. Wald, N.J., *The Epidemiological Approach: An Introduction to Epidemiology in Medicine*. 4th ed. 2004, London: Wolfson Institute of Preventive Medicine. Royal Society of Medicine Press Ltd.

32. Office for National Statistics, *Cancer Statistics: Registrations Series MB1*. 2005: UK.

33. Averette, H.E., M.F. Janicek, and H.R. Menck, *The National Cancer Data Base report on ovarian cancer. American College of Surgeons Commission on Cancer and the American Cancer Society.* Cancer, 1995. **76**(6): p. 1096-103.

34. Koonings, P.P., K. Campbell, D.R. Mishell, Jr., and D.A. Grimes, *Relative frequency of primary ovarian neoplasms: a 10-year review.* Obstetrics & Gynecology, 1989. **74**(6): p. 921-6.

35. ISD Online. *Cancer Incidence, Mortality & Survival Data*. 2008 [cited 2008.

36. Northern Ireland Cancer Registry. *Cancer Registrations in Northern Ireland, 2003*. 2005 [cited.

37. Bankhead, C.K., S., *CancerStats - Ovarian Cancer*. 2004, Cancer Research UK: London. p. 1-6.

38. Kitchener, H.C., *Survival from cancer of the ovary in England and Wales up to 2001.* British Journal of Cancer, 2008. **99 Suppl 1**: p. S73-4.

39. Rachet, B., L.M. Woods, E. Mitry, M. Riga, et al., *Cancer survival in England and Wales at the end of the 20th century.* British Journal of Cancer, 2008. **99 Suppl 1**: p. S2-10.

40. Office for National Statistics, *Mortality Statistics: Cause, England and Wales 2006.* 2008. **Series DH2**.

41. IARC, ed. *Cancer Incidence in Five Continents*. ed. D.M. Parkin, et al. Vol. VIII. 2002, International Agency for Research on Cancer: Lyon.

42. Elmasry, K. and S.A. Gayther, *Ovarian cancer aetiology: facts and fiction.* Journal of Family Planning & Reproductive Health Care, 2006. **32**(2): p. 82-6.

43. Whittemore, A.S., R. Harris, and J. Itnyre, *Characteristics relating to ovarian cancer risk: collaborative analysis of 12 US case-control studies. IV. The pathogenesis of epithelial ovarian cancer. Collaborative Ovarian Cancer Group.* American Journal of Epidemiology, 1992. **136**(10): p. 1212-20.

44. Bosetti, C., E. Negri, D. Trichopoulos, S. Franceschi, et al., *Long-term effects of oral contraceptives on ovarian cancer risk.* International Journal of Cancer, 2002. **102**(3): p. 262-5.

45. Ness, R.B., D.W. Cramer, M.T. Goodman, S.K. Kjaer, et al., *Infertility, fertility drugs, and ovarian cancer: a pooled analysis of case-control studies.* American Journal of Epidemiology, 2002. **155**(3): p. 217-24.

46. Zhou, B., Q. Sun, R. Cong, H. Gu, et al., *Hormone replacement therapy and ovarian cancer risk: a meta-analysis.* Gynecologic Oncology, 2008. **108**(3): p. 641-51.

47. Stratton, J.F., P. Pharoah, S.K. Smith, D. Easton, et al., *A systematic review and meta-analysis of family history and risk of ovarian cancer.* British Journal of Obstetrics & Gynaecology, 1998. **105**(5): p. 493-9.

48. Aarnio, M., J.P. Mecklin, L.A. Aaltonen, M. Nystrom-Lahti, et al., *Life-time risk of different cancers in hereditary non-polyposis colorectal cancer (HNPCC) syndrome.* International Journal of Cancer, 1995. **64**(6): p. 430-3.

49. Huncharek, M., J.F. Geschwind, and B. Kupelnick, *Perineal application of cosmetic talc and risk of invasive epithelial ovarian cancer: a meta-analysis of 11,933 subjects from sixteen observational studies.* Anticancer Research, 2003. **23**(2C): p. 1955-60.

50. Gates, M.A., S.S. Tworoger, K.L. Terry, L. Titus-Ernstoff, et al., *Talc use, variants of the GSTM1, GSTT1, and NAT2 genes, and risk of epithelial ovarian cancer.* Cancer Epidemiology, Biomarkers & Prevention, 2008. **17**(9): p. 2436-44.

51. Jacobs, I., I. Stabile, J. Bridges, P. Kemsley, et al., *Multimodal approach to screening for ovarian cancer.* Lancet, 1988. **1**(8580): p. 268-71.

52. Karlan, B.Y., *The status of ultrasound and color Doppler imaging for the early detection of ovarian carcinoma.* Cancer Investigation, 1997. **15**(3): p. 265-9.

53. van Nagell, J.R., Jr., P.D. DePriest, M.B. Reedy, H.H. Gallion, et al., *The efficacy of transvaginal sonographic screening in asymptomatic women at risk for ovarian cancer.[see comment].* Gynecologic Oncology, 2000. **77**(3): p. 350-6.

54. Bast, R.C., Jr., N. Urban, V. Shridhar, D. Smith, et al., *Early detection of ovarian cancer: promise and reality.* Cancer Treatment & Research, 2002. **107**: p. 61-97.

55. Fritsche, H.A. and R.C. Bast, *CA 125 in ovarian cancer: advances and controversy.[comment].* Clinical Chemistry, 1998. **44**(7): p. 1379-80.

56. Sjovall, K., B. Nilsson, and N. Einhorn, *The significance of serum CA 125 elevation in malignant and nonmalignant diseases.* Gynecologic Oncology, 2002. **85**(1): p. 175-8.

57. Cane, P., C. Azen, E. Lopez, L.D. Platt, et al., *Tumor marker trends in asymptomatic women at risk for ovarian cancer: relevance for ovarian cancer screening.* Gynecologic Oncology, 1995. **57**(2): p. 240-5.

58. Skates, S.J., U. Menon, N. MacDonald, A.N. Rosenthal, et al., *Calculation of the Risk of Ovarian Cancer From Serial CA-125 Values for Preclinical Detection in Postmenopausal Women.* J Clin Oncol, 2003. **21**(10 Suppl): p. 206-10.

59. Jacobs, I., A.P. Davies, J. Bridges, I. Stabile, et al., *Prevalence screening for ovarian cancer in postmenopausal women by CA 125 measurement and ultrasonography.[see comment].* BMJ, 1993. **306**(6884): p. 1030-4.

60. Menon, U., A. Talaat, A.R. Jeyarajah, A.N. Rosenthal, et al., *Ultrasound assessment of ovarian cancer risk in postmenopausal women with CA125 elevation.* British Journal of Cancer, 1999. **80**(10): p. 1644-7.

61. Jacobs, I.J., S.J. Skates, N. MacDonald, U. Menon, et al., *Screening for ovarian cancer: a pilot randomised controlled trial.* Lancet, 1999. **353**(9160): p. 1207-10.

62. Buys, S.S., E. Partridge, M.H. Greene, P.C. Prorok, et al., *Ovarian cancer screening in the Prostate, Lung, Colorectal and Ovarian (PLCO) cancer screening trial: findings from the initial screen of a randomized trial.[erratum appears in Am J Obstet Gynecol. 2005 Dec;193(6):2183-4].* American Journal of Obstetrics & Gynecology, 2005. **193**(5): p. 1630-9.

63. Menon, U., A. Gentry-Maharaj, R. Hallett, A. Ryan, 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.

64. Petricoin, E.F., A.M. Ardekani, B.A. Hitt, P.J. Levine, et al., *Use of proteomic patterns in serum to identify ovarian cancer.[see comment].* Lancet, 2002. **359**(9306): p. 572-7.

65. Baggerly, K.A., J.S. Morris, S.R. Edmonson, and K.R. Coombes, *Signal in noise: evaluating reported reproducibility of serum proteomic tests for ovarian cancer.[see comment].* Journal of the National Cancer Institute, 2005. **97**(4): p. 307-9.

66. Hogg, R. and M. Friedlander, *Biology of epithelial ovarian cancer: implications for screening women at high genetic risk.* Journal of Clinical Oncology, 2004. **22**(7): p. 1315-27.

67. Eltabbakh, G.H., P.R. Yadav, and A. Morgan, *Clinical picture of women with early stage ovarian cancer.* Gynecologic Oncology, 1999. **75**(3): p. 476-9.

68. Webb, P.M., D.M. Purdie, S. Grover, S. Jordan, et al., *Symptoms and diagnosis of borderline, early and advanced epithelial ovarian cancer.* Gynecologic Oncology, 2004. **92**(1): p. 232-9.

69. Olson, S.H., L. Mignone, C. Nakraseive, T.A. Caputo, et al., *Symptoms of ovarian cancer.* Obstetrics & Gynecology, 2001. **98**(2): p. 212-7.

70. Bell, D.A., *Origins and molecular pathology of ovarian cancer.* Modern Pathology, 2005. **18 Suppl 2**: p. S19-32.

71. Friedlander, M.L., *Prognostic factors in ovarian cancer.* Semin Oncol, 1998. **25**(3): p. 305-14.

72. Bell, D.A. and R.E. Scully, *Early de novo ovarian carcinoma. A study of fourteen cases.* Cancer, 1994. **73**(7): p. 1859-64.

73. Ogawa, S., T. Kaku, S. Amada, H. Kobayashi, et al., *Ovarian endometriosis associated with ovarian carcinoma: a clinicopathological and immunohistochemical study.* Gynecologic Oncology, 2000. **77**(2): p. 298-304.

74. Scully, R.E., *Influence of origin of ovarian cancer on efficacy of screening.* Lancet, 2000. **355**(9209): p. 1028-9.

75. Goff, B.A., L.S. Mandel, C.H. Melancon, and H.G. Muntz, *Frequency of symptoms of ovarian cancer in women presenting to primary care clinics.[see comment].* JAMA, 2004. **291**(22): p. 2705-12.

76. Goldberg, R.P., P.K. Sand, and H. Beck, *Early-stage ovarian carcinoma presenting with irritative voiding symptoms and urge incontinence.* International Urogynecology Journal, 2001. **12**(5): p. 342-4.

77. Evans, J., S. Ziebland, and A. McPherson, *Minimizing delays in ovarian cancer diagnosis: an expansion of Andersen's model of 'total patient delay'.* Fam Pract, 2006.

78. Koldjeski, D., M.K. Kirkpatrick, M. Swanson, L. Everett, et al., *Ovarian cancer: early symptom patterns.* Oncology Nursing Forum Online, 2003. **30**(6): p. 927-33.

79. Bogardus, S.T., Jr., *What do we know about diverticular disease? A brief overview.* J Clin Gastroenterol, 2006. **40**(7 Suppl 3): p. S108-11.

80. Stollman, N.H. and J.B. Raskin, *Diverticular disease of the colon.* Journal of Clinical Gastroenterology, 1999. **29**(3): p. 241-52.

81. Janes, S.E.J., A. Meagher, and F.A. Frizelle, *Management of diverticulitis.* BMJ, 2006. **332**(7536): p. 271-5.

82. Department of Health, *Referral guidelines for suspected cancer*, D.o. Health, Editor. 2000. p. 62.

83. Allgar, V.L., R.D. Neal, N. Ali, B. Leese, et al., *Urgent GP referrals for suspected lung, colorectal, prostate and ovarian cancer.* British Journal of General Practice, 2006. **56**(526): p. 355-62.

84. NICE, *CG27 Referral for suspected cancer: NICE guideline*. 2005. p. 98.

85. Farquhar, M.C., S.I.G. Barclay, H. Earl, G.E. Grande, et al., *Barriers to effective communication across the primary/secondary interface: examples from the ovarian cancer patient journey (a qualitative study).* European Journal of Cancer Care, 2005. **14**(4): p. 359-66.

86. Benedet, J.L., H. Bender, H. Jones, 3rd, H.Y. Ngan, et al., *FIGO staging classifications and clinical practice guidelines in the management of gynecologic cancers. FIGO Committee on Gynecologic Oncology.* International Journal of Gynaecology & Obstetrics, 2000. **70**(2): p. 209-62.

87. Pectasides, D., E. Pectasides, A. Psyrri, and T. Economopoulos, *Treatment issues in clear cell carcinoma of the ovary: a different entity?* Oncologist, 2006. **11**(10): p. 1089-94.

88. Bhoola, S. and W.J. Hoskins, *Diagnosis and management of epithelial ovarian cancer.* Obstetrics & Gynecology, 2006. **107**(6): p. 1399-410.

89. Chan, Y.M., T.Y. Ng, P.W. Lee, H.Y. Ngan, et al., *Symptoms, coping strategies, and timing of presentations in patients with newly diagnosed ovarian cancer.* Gynecol Oncol, 2003. **90**(3): p. 651-6.

90. Igoe, B.A., *Symptoms attributed to ovarian cancer by women with the disease.* Nurse Practitioner, 1997. **22**(7): p. 122.

91. Nelson, L., A. Ekbom, and E. Gerdin, *Ovarian cancer in young women in Sweden, 1989-1991.* Gynecologic Oncology, 1999. **74**(3): p. 472-6.

92. Smith, L.H., *Early clinical detection of ovarian cancer: a review of the evidence.* Expert Review of Anticancer Therapy, 2006. **6**(7): p. 1045-52.

93. Bankhead, C.R., S.T. Kehoe, and J. Austoker, *Symptoms associated with diagnosis of ovarian cancer: a systematic review.* BJOG: an International Journal of Obstetrics & Gynaecology, 2005. **112**(7): p. 857-65.

94. Bankhead, C., *Identifying Potentially Significant Diagnostic Factors For Ovarian Cancer In Primary Care: A Qualitative And Quantitative Study*, in *Department of Primary Health Care, Faculty of Clinical Medicine*. 2005, University of Oxford: Oxford. p. 288.

95. Mantzavinos, T., D. Botsis, S. Konidaris, and P.A. Zourlas, *Review of ovarian cancer of the University of Athens.* European Journal of Gynaecological Oncology, 1988. **9**(1): p. 27-9.

96. Piura, B., M. Meirovitz, M. Bartfeld, I. Yanai-Inbar, et al., *Peritoneal papillary serous carcinoma: study of 15 cases and comparison with stage III-IV ovarian papillary serous carcinoma.* Journal of Surgical Oncology, 1998. **68**(3): p. 173-8.

97. Kennedy, C.R. and H. Gordon, *Ovarian cancer: the ten year experience of a district general hospital.* British Journal of Obstetrics & Gynaecology, 1981. **88**(12): p. 1186-91.

98. Beck, H.H., R.P. Goldberg, and P.K. Sand, *Patient Symptoms in Early Stage Ovarian Cancer.* Journal of Pelvic Surgery, 2001. **7**(2): p. 79-82.

99. Bankhead, C.R., C. Collins, H. Stokes-Lampard, P. Rose, et al., *Identifying symptoms of ovarian cancer: a qualitative and quantitative study.* BJOG, 2008. **115**(8): p. 1008-14.

100. Deligdisch, L., F. Penault-Llorca, P. Schlosshauer, A. Altchek, et al., *Stage I ovarian carcinoma: different clinical pathologic patterns.* Fertility & Sterility, 2007. **88**(4): p. 906-10.

101. Ryerson, A.B., C. Eheman, J. Burton, N. McCall, et al., *Symptoms, Diagnoses, and Time to Key Diagnostic Procedures Among Older U.S. Women With Ovarian Cancer.* Obstetrics & Gynecology, 2007. **109**(5): p. 1053-1061.

102. Behtash, N., E. Ghayouri Azar, and F. Fakhrejahani, *Symptoms of ovarian cancer in young patients 2 years before diagnosis, a case-control study.* European Journal of Cancer Care, 2008. **17**(5): p. 483-7.

103. Andersen, M.R., B.A. Goff, K.A. Lowe, N. Scholler, et al., *Combining a symptoms index with CA 125 to improve detection of ovarian cancer.* Cancer, 2008. **113**(3): p. 484-9.

104. Olsen, C.M., J. Cnossen, A.C. Green, and P.M. Webb, *Comparison of symptoms and presentation of women with benign, low malignant potential and*

invasive ovarian tumors. European Journal of Gynaecological Oncology, 2007. **28**(5): p. 376-80.

105. Black, S.S., S.L. Butler, P.A. Goldman, and M.J. Scroggins, *Ovarian cancer symptom index: possibilities for earlier detection.* Cancer, 2007. **109**(2): p. 167-9.

106. Ferrell, B., S. Smith, C. Cullinane, and C. Melancon, *Symptom concerns of women with ovarian cancer.* Journal of Pain & Symptom Management, 2003. **25**(6): p. 528-38.

107. Fitch, M., K. Deane, D. Howell, and R.E. Gray, *Women's experiences with ovarian cancer: reflections on being diagnosed.* Canadian Oncology Nursing Journal, 2002. **12**(3): p. 152-68.

108. Goff, B.A., L.S. Mandel, C.W. Drescher, N. Urban, et al., *Development of an ovarian cancer symptom index: possibilities for earlier detection.[see comment].* Cancer, 2007. **109**(2): p. 221-7.

109. Odukogbe, A.A., C.A. Adebamowo, B. Ola, O. Olayemi, et al., *Ovarian cancer in Ibadan: characteristics and management.* Journal of Obstetrics & Gynaecology, 2004. **24**(3): p. 294-7.

110. Portenoy, R.K., A.B. Kornblith, G. Wong, V. Vlamis, et al., *Pain in ovarian cancer patients. Prevalence, characteristics, and associated symptoms.* Cancer, 1994. **74**(3): p. 907-15.

111. Rufford, B.D., I. Jacobs, and U. Menon, *Feasibility of screening for ovarian cancer using symptoms as selection criteria.* BJOG: an International Journal of Obstetrics & Gynaecology, 2007. **114**(1): p. 59-64.

112. Menczer, J., A. Chetrit, and S. Sadetzki, *The effect of symptom duration in epithelial ovarian cancer on prognostic factors.* Arch Gynecol Obstet, 2008.

113. Elwood, M., *Study of effect of delays on ovarian cancer was weak.[comment].* BMJ, 2002. **324**(7345): p. 1100.

114. Smith, E.M., *Telephone interviewing in healthcare research: a summary of the evidence.* Nurse Researcher, 2005. **12**(3): p. 32-41.

115. Richards, M.A., A.M. Westcombe, S.B. Love, P. Littlejohns, et al., *Influence of delay on survival in patients with breast cancer: a systematic review.* Lancet, 1999. **353**(9159): p. 1119-26.

116. Barsky, A.J., *Forgetting, fabricating, and telescoping: the instability of the medical history.* Archives of Internal Medicine, 2002. **162**(9): p. 981-4.

117. Jenkins, P., G. Earle-Richardson, D.T. Slingerland, and J. May, *Time dependent memory decay.* American Journal of Industrial Medicine, 2002. **41**(2): p. 98-101.

118. Broder, M.S., D.E. Kanouse, and S.J. Bernstein, *Assessing symptoms before hysterectomy: is the medical record accurate?* American Journal of Obstetrics & Gynecology, 2001. **185**(1): p. 97-102.

119. Pakhomov, S.V., S.J. Jacobsen, C.G. Chute, and V.L. Roger, *Agreement between patient-reported symptoms and their documentation in the medical record.* American Journal of Managed Care, 2008. **14**(8): p. 530-9.

120. Jerant, A.F. and D.B. Hill, *Does the use of electronic medical records improve surrogate patient outcomes in outpatient settings?* Journal of Family Practice, 2000. **49**(4): p. 349-57.

121. Pringle, M., P. Ward, and C. Chilvers, *Assessment of the completeness and accuracy of computer medical records in four practices committed to recording data on computer.* British Journal of General Practice, 1995. **45**(399): p. 537-41.

122. Hamilton, W.T., A.P. Round, D. Sharp, and T.J. Peters, *The quality of record keeping in primary care: a comparison of computerised, paper and hybrid systems.* British Journal of General Practice, 2003. **53**(497): p. 929-33; discussion 933.

123. de Lusignan, S., T. Valentin, T. Chan, N. Hague, et al., *Problems with primary care data quality: osteoporosis as an exemplar.* Informatics in Primary Care, 2004. **12**(3): p. 147-56.

124. Peabody, J.W., F. Tozija, J.A. Munoz, R.J. Nordyke, et al., *Using vignettes to compare the quality of clinical care variation in economically divergent countries.* Health Services Research, 2004. **39**(6 Pt 2): p. 1951-70.

125. Shavers, V.L., C.F. Lynch, and L.F. Burmeister, *Racial differences in factors that influence the willingness to participate in medical research studies.* Annals of Epidemiology, 2002. **12**(4): p. 248-56.

126. Kroenke, K. and R.K. Price, *Symptoms in the community. Prevalence, classification, and psychiatric comorbidity.* Arch Intern Med, 1993. **153**(21): p. 2474-80.

127. Pamuk, O.N., G.E. Pamuk, and A.F. Celik, *Revalidation of description of constipation in terms of recall bias and visual scale analog questionnaire.* Journal of Gastroenterology & Hepatology, 2003. **18**(12): p. 1417-22.

128. Modugno, F., R.B. Ness, and J.E. Wheeler, *Reproductive risk factors for epithelial ovarian cancer according to histologic type and invasiveness.* Annals of Epidemiology, 2001. **11**(8): p. 568-74.

129. Belli, R.F., *The structure of autobiographical memory and the event history calendar: potential improvements in the quality of retrospective reports in surveys.* Memory, 1998. **6**(4): p. 383-406.

130. Fritz, A., C. Percy, A. Jack, K. Shanmugaratnam, et al., eds. *International Classification of Diseases for Oncology (ICD-O)*. 3rd ed. 2000, World Health Organisation: Geneva.

131. Agresti, A., *Categorical Data Analysis*. 1996, Wiley-Interscience: New York.

132. Department of Health, *Cancer Waiting Targets: A Guide (Version 5)*. 2006. p. 38.

133. Pinsky, P.F., A. Miller, B.S. Kramer, T. Church, et al., *Evidence of a healthy volunteer effect in the prostate, lung, colorectal, and ovarian cancer screening trial.* American Journal of Epidemiology, 2007. **165**(8): p. 874-81.

134. McNally, O.M., V. Wareham, D.J. Flemming, M.E. Cruickshank, et al., *The impact of the introduction of a fast track clinic on ovarian cancer referral patterns.* European Journal of Cancer Care, 2003. **12**(4): p. 327-30.

135. Kumar, H. *NAEDI Progress and Plans*. in *The National Awareness and Early Detection Initiative (NAEDI) Launch Conference*. 2008. London, UK.

136. Bjerager, M., T. Palshof, R. Dahl, P. Vedsted, et al., *Delay in diagnosis of lung cancer in general practice.* British Journal of General Practice, 2006. **56**(532): p. 863-8.

137. Moss, E.L., J. Hollingworth, and T.M. Reynolds, *The role of CA125 in clinical practice.* Journal of Clinical Pathology, 2005. **58**(3): p. 308-12.

138. Richards, M.A., P. Smith, A.J. Ramirez, I.S. Fentiman, et al., *The influence on survival of delay in the presentation and treatment of symptomatic breast cancer.* British Journal of Cancer, 1999. **79**(5-6): p. 858-64.

139. Neal, R.D. and V.L. Allgar, *Sociodemographic factors and delays in the diagnosis of six cancers: analysis of data from the "National Survey of NHS Patients: Cancer".* British Journal of Cancer, 2005. **92**(11): p. 1971-5.

140. Stratton, J.F., P. Pharoah, J.A. Tidy, and M.E. Paterson, *An analysis of ovarian tumor diameter and survival.* International Journal of Gynecological Cancer, 2000. **10**(6): p. 449-451.

141. Sightler, S.E., G.M. Boike, R.E. Estape, and H.E. Averette, *Ovarian cancer in women with prior hysterectomy: a 14-year experience at the University of Miami.[see comment].* Obstetrics & Gynecology, 1991. **78**(4): p. 681-4.

142. Bristow, R.E., R.S. Tomacruz, D.K. Armstrong, E.L. Trimble, et al., *Survival effect of maximal cytoreductive surgery for advanced ovarian carcinoma during the platinum era: a meta-analysis.* Journal of Clinical Oncology, 2002. **20**(5): p. 1248-59.

143. Griffiths, C.T., *Surgical resection of tumor bulk in the primary treatment of ovarian carcinoma.* National Cancer Institute Monographs, 1975. **42**: p. 101-4.

144. Hoskins, W.J., W.P. McGuire, M.F. Brady, H.D. Homesley, et al., *The effect of diameter of largest residual disease on survival after primary cytoreductive surgery in patients with suboptimal residual epithelial ovarian carcinoma.* American Journal of Obstetrics & Gynecology, 1994. **170**(4): p. 974-9; discussion 979-80.

145. Nustad, K., R.C. Bast, Jr., T.J. Brien, O. Nilsson, et al., *Specificity and affinity of 26 monoclonal antibodies against the CA 125 antigen: first report from the ISOBM TD-1 workshop. International Society for Oncodevelopmental Biology and Medicine.* Tumour Biology, 1996. **17**(4): p. 196-219.

146. Skates SJ, Jacobs IJ, Pauler DK, Symecko HL, et al. *Estimated duration of pre-clinical ovarian cancer from longitudinal CA125 levels*. in *Proceedings of American Association Cancer Research meeting*. 1999.

147. Markman, M., K. Webster, K. Zanotti, G. Peterson, et al., *Examples of the marked variability in the relationship between the serum CA-125 antigen level and cancer-related symptoms in ovarian cancer.* Gynecologic Oncology, 2004. **93**(3): p. 715-7.

148. Neter, J. and J. Waksberg, *A study of response errors in expenditures data from household interview.* Journal of the American Statistical Association, 1964. **59**: p. 18-55.

149. Scott, A. and S. Wordsworth, *The effects of shifts in the balance of care on general practice workload.* Family Practice, 1999. **16**(1): p. 12-7.

150. Morabia, A. and M.C. Costanza, *International variability in ages at menarche, first livebirth, and menopause. World Health Organization Collaborative Study of Neoplasia and Steroid Contraceptives.[erratum appears in Am J Epidemiol 1999 Sep 1;150(5):546].* American Journal of Epidemiology, 1998. **148**(12): p. 1195-205.

151. Royal College of General Practitioners, *General Practice in the UK: A Basic Overview*, in *RCGP Information Sheet No. 4*. 2005.

152. Egton Medical Information Systems Limited. [cited 10.01.2007]; Available from: *http://www.emis-online.com/about-us/company-profile/*.

153. Gynecologic Cancer Foundation, S.o.G.O., and American Cancer Society, *Ovarian Cancer Symptoms Consensus Statement.*, W.s.C. Network, Editor. 2007.

154. Office of Health Economics, *Compendium of Health Statistics*. 15th ed. 2004, London: Office of Health Economics.

155. Rowlands, S. and K. Moser, *Consultation rates from the general practice research database.* British Journal of General Practice, 2002. **52**(481): p. 658-60.

156. Milsom, I. and U. Molander, *Urogenital Ageing.* Journal of the British Menopause Society, 1998. **4**: p. 151-156.

157. Department of Health. *Ovarian cancer: key messages for health professionals*. 2009 [cited 2009 10.04.2009]; Available from: *http://www.dh.gov.uk/en/Healthcare/NationalServiceFrameworks/Cancer/DH_095624*.

158. Smith, L.K., C. Pope, and J.L. Botha, *Patients' help-seeking experiences and delay in cancer presentation: a qualitative synthesis.* Lancet, 2005. **366**(9488): p. 825-31.