Research

Elizabeth A Shephard and William T Hamilton

Selection of men for investigation of possible testicular cancer in primary care:

a large case–control study using electronic patient records

Abstract

Background
Testicular cancer incidence has risen over the last two decades and is expected to continue to rise. There are no primary care studies on the clinical features of testicular cancer, with recent National Institute for Health and Care Excellence (NICE) guidance based solely upon clinical consensus.

Aim
To identify clinical features of testicular cancer and to quantify their risk in primary care patients, with the aim of improving the selection of patients for investigation.

Design and setting
A matched case–control study in males aged ≥17 years, using Clinical Practice Research Datalink records.

Method
Putative clinical features of testicular cancer were identified and analysed using conditional logistic regression. Positive predictive values (PPVs) were calculated for those aged <50 years.

Results
In all, 1398 cases were available, diagnosed between 2000 and 2012, with 4956 age-, sex-, and practice-matched controls. Nine features were independently associated with testicular cancer, the top three being testicular swelling (odds ratio [OR] 280, 95% confidence interval [CI] = 110 to 690), testicular lump (OR 270, 95% CI = 100 to 740), and scrotal swelling (OR 170, 95% CI = 35 to 800). The highest PPV for 17–49-year-olds was testicular lump, at 2.5% (95% CI = 1.1 to 5.6). Combining testicular lump with testicular swelling or testicular pain produced PPVs of 17% and 10%, respectively.

Conclusion
Testicular enlargement carries a risk of cancer of 2.5% — close to the current 3% threshold in UK referral guidance. Contrary to traditional teaching, painful testicular enlargement may signify cancer. Some initial hydrocele diagnoses appear to be wrong, with missed cancers, suggesting an ultrasound may be useful when a hydrocele diagnosis is uncertain. These results support the existing NICE guidelines, and help to characterise when an ultrasound should be considered in symptomatic men.

Keywords
diagnosis; lump; pain; primary health care; testicular cancer.

INTRODUCTION

Testicular cancer is the 16th most common cancer in UK males. Nearly 2300 new cases are diagnosed annually, nearly half in those aged <35 years.[1] Incidence rates in the UK have risen by >27% since the early 1990s, and are expected to rise by 12% in the UK between 2014 and 2035.[2] Survival rates are highest in younger men: 98% of <50-year-olds survive for ≥5 years.[3] The high survival rate probably reflects the excellent response to treatment at all stages but it does not eliminate the benefit to be gained from timely diagnosis, as complications such as thromboses may arise.[4]

For testicular cancer, diagnosis requires symptomatic presentation, usually to primary care. The National Institute for Health and Care Excellence (NICE) provides evidence-based guidelines to aid GP decision making, with the aim of improving patient outcomes. The current NICE guidance for testicular cancer recommends urgent specialist referral for men with non-painful enlargement or a change in shape or texture of the testis. A direct-access ultrasound should be considered for men with persistent or unexplained other testicular symptoms.[5] These recommendations were made by reaching a consensus, as the guideline development group could find no primary care evidence on the features of testicular cancer.

GPs typically investigate testicular symptoms promptly. During 2010, only 16% of patients with testicular cancer visited their GP three or more times before referral.[6] The median time to diagnosis for testicular cancer (44 days) is one of the shortest for all cancers.[7] For one study, however, the time to diagnosis increased by >36 days when non-NICE-recommended symptoms were present.[8]

The sole primary care study of male cancer symptoms merged all types of cancer together, reporting an increased risk of testicular cancer in patients with testicular lump (185-fold), testicular pain (16-fold), and venous thromboembolism (nine-fold).[9] Symptoms of testicular cancer reported from secondary care studies include back pain, leg pain, lethargy, fatigue, weight loss, testicular swelling, testicular pain, gynaecomastia, abdominal pain, painless nodule, testicular lump, painless scrotal mass, and scrotal pain.[10–13] However, these studies were small, with patient numbers ranging between 6 and 140. Serum tumour biomarkers such as alpha-fetoprotein (AFP), human chorionic gonadotropin (hCG), and lactate dehydrogenase (LDH) have been found to be elevated in patients with testicular cancer.[14,15]

This study aimed to identify clinical features of testicular cancer in primary care and to quantify their risk using a risk assessment tool, to improve the selection of patients for investigation.

METHOD

This was a matched case–control study using primary care electronic patient records from the UK's Clinical Practice Research

EA Shephard, PhD, CPsychol, research fellow; **WT Hamilton**, MD, FRCP, FRCGP, professor of primary care diagnostics, University of Exeter Medical School, Exeter, UK.

Address for correspondence

EA Shephard, University of Exeter Medical School, College House, St Luke's Campus, Magdalen Road, Exeter EX1 2LU, UK.

How this fits in

There are no primary care studies on the testicular cancer prodrome, therefore the National Institute for Health and Care Excellence (NICE) had difficulty providing recommendations for investigation of testicular symptoms. It suggested referral for enlargement and a primary care ultrasound for persistent or unexplained other testicular symptoms, though these symptoms were not further defined. Most of the symptoms of testicular enlargement or swelling have risks of cancer around 3%, supporting NICE's recommendation for urgent referral. The authors' risk assessment tool shows that certain combinations of symptoms (some of which NICE has not considered) have risks >3%, also warranting urgent referral. The tool also shows several symptom combinations that have risks of testicular cancer in the 1–3% range for which an ultrasound is appropriate, expanding on NICE guidance.

Datalink (CPRD), mirroring the authors' previous cancer studies.[16–23] The CPRD contains anonymised medical records of patients from participating practices across the UK, representing roughly 8.8% of the population. Information is stored for clinical events such as symptoms, investigations, prescriptions, and diagnoses, together with patient demographics.

Cases and controls

A list of 22 testicular cancer diagnostic codes was collated from the CPRD master code library and used to identify patients (additional information is available from the authors upon request). Cases were selected if aged ≥16 years (though the youngest transpired to be 17), diagnosed with testicular cancer between 2000 and 2012, and had consulted their GP in the year before the diagnosis date. Up to five age-, sex-, and general practice-matched controls were assigned to each case. The first testicular cancer code was taken as the date of diagnosis, or the 'index date'. Controls were matched to their case's index date. Exclusion criteria were: controls with a previous testicular cancer diagnosis, controls not consulting their GP in the year before the case index date, cases with no controls, and cases or controls with missing relevant data. Three female cases, two with tunica vaginalis cancers, which can occur in either sex, and one with a tumour in an undescended testis, were excluded.

Selection of possible features of testicular cancer

All previously reported diagnostic features of testicular cancer identified through literature searches were studied following PubMed, EBSCO, and Google Scholar searches using the terms 'testicular cancer primary care', 'testicular cancer symptoms', and 'early signs/symptoms/features of testicular cancer'. Self-reported symptoms described on online cancer support group forums were also included, allowing for new features to be included. Libraries of codes representing each candidate feature were assembled from the CPRD master list of >100 000 medical codes. Occurrences of these codes were identified in the year before the index date. Features in <2% of cases were deemed too rare to draw useful conclusions from, and were excluded.

Records of fractures were compiled to test for any recording bias between cases and controls, with the assumption that fractures would be approximately equal in cases and controls. For laboratory tests, results outside each laboratory's normal range were considered abnormal. Patients without a test result were grouped with those having a normal result. Some tests were grouped together — inflammatory markers comprised of any of erythrocyte sedimentation rate, C-reactive protein, or plasma viscosity; liver function tests (LFTs) comprised of any of the hepatic enzymes. LDH was examined separately, as it is the only one of the three biomarkers routinely available in the full blood count test.

Analysis

The main analysis used conditional logistic regression. Features independently associated with testicular cancer in univariable analyses with a P-value of ≤0.1 were entered into multivariable analysis. Features were assembled into broad clinical groups (for example, pain: abdominal pain, groin pain, testicular pain) for the first stage of multivariable analysis, with a P-value threshold of ≤0.05 required for entry into the final stage of modelling. The final multivariable model contained all features surviving the earlier analyses, and used a P-value of ≤0.01 for final retention. Excluded variables were checked against the final model. Clinically plausible interactions were tested against the final model, also using a P-value threshold of ≤0.01.

Calculation of positive predictive values (PPVs)

Risk estimates for features associated with testicular cancer in patients consulting

in primary care can be calculated using Bayes' theorem (prior odds for a given feature multiplied by the likelihood ratio is equal to the posterior odds of having the disease). The prior odds were calculated using the age-specific national incidence rate for testicular cancer in 2008.[24] PPVs within the risk assessment tool (RAT) were calculated for men aged 17–49 years. This age range allowed for >81% of the testicular cancer patients to be represented in a single figure, maximising clinical utility. The PPVs were calculated for consulting patients aged 17–49 years only. Thus, the posterior odds were divided by 0.70, based on only 3915 out of 5601 (70%) controls in that age group having seen their GP in the previous year, while all cases had done so. PPVs for selected single and combined features in the ≥50 age group were also estimated, though with wider confidence intervals. Analyses were performed using Stata (version 14).

Power calculations were used rather than sample size calculations. The CPRD provided estimates of 1500 cases. Using a case–control ratio of 1:4, and 5% two-sided α, these numbers provided >99% power to detect a change in prevalence for a rare variable of 3% in cases and of 1% in controls. For commoner variables, this size has 95% power to detect a change in prevalence of 20% in cases and 16% in controls.

RESULTS

The CPRD provided 8300 patients: 1428 cases and 6872 controls. The application of exclusions is shown in Figure 1. A final number of 1398 cases and 4956 controls were studied. The median ages of cases and controls was 39 and 40 years, respectively. Cases consulted twice as often as controls in the year before diagnosis (8 to 4, respectively, $P<0.001$, rank-sum test).

Clinical features

A total of 31 symptoms and 22 investigations were studied. Nine remained significant in the multivariable final model. Of the 1398 cases, 938 (67%) consulted their GP with at least one of these final-model features in the year before their diagnosis. Table 1 shows the frequencies, univariable likelihood ratios, and multivariable odds ratios of the nine final-model features. There was a strong antagonistic interaction between testicular pain and testicular swelling (interaction OR 0.008, $P<0.001$), meaning the predictive power when both features are present is less than would be expected from simply multiplying the two odds ratios. Seventy-one cases (5%) and seven controls (0.1%) had a recorded LDH code. Of those, 27 were abnormally raised, all belonging to cases (2%). LDH was dropped from further analyses. In all, 178 (13%) cases had a diagnosis of either hydrocele (25), epididymis or epididymis-orchitis (150), or both (three — these were added to the hydrocele group for analysis). For epididymis/epididymis-orchitis, 106 (71%) were recorded within 3 months preceding the cancer diagnosis. For the hydrocele group ($n = 28$), 17 (61%) of these codes were recorded in the 3 months before diagnosis.

Cases were grouped according to tumour type. Of the 1398 cases, 684 (49%) had a generic code for testicular cancer, not specifying the histological subtype, 568 (41%) were seminomas, 122 (9%) were teratomas, 22 (2%) were in situ, and two (0.1%) were classed as secondary malignant neoplasms. A sensitivity analysis excluded in situ cases. The model did not significantly differ with the exclusion of these cases.
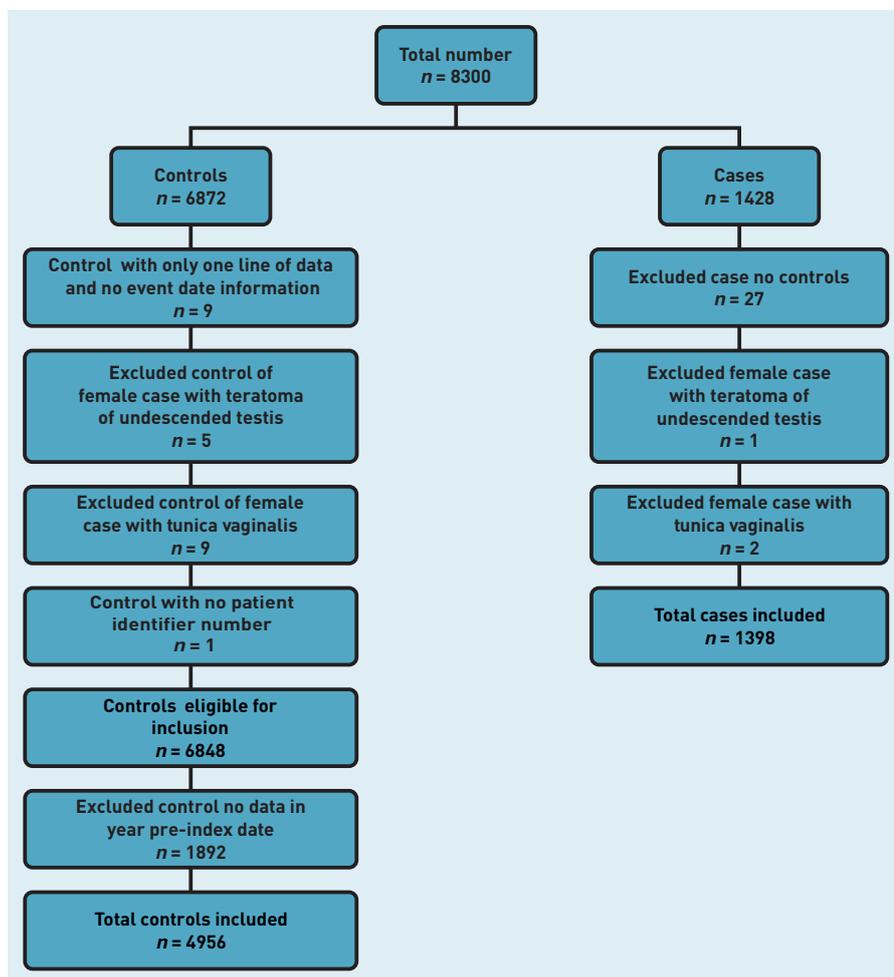
Figure 1. Application of exclusion criteria.

Table 1. Features of testicular cancer (age ≥17 years)

Feature	Cases, *n* (%) *n* = 1398	Controls, *n* (%) *n* = 4956	Likelihood ratio[a] (95% CI)	Odds ratio in multivariable analysis[b] (95% CI)
Symptoms				
Testicular swelling	354 (25)	10 (0.2)	126 (67 to 235)	280 (110 to 690)
Testicular lump	282 (20)	7 (0.1)	143 (68 to 302)	270 (100 to 740)
Testicular pain	186 (13)	32 (0.6)	21 (14 to 30)	38 (22 to 68)
Abdominal pain	67 (5)	124 (3.0)	1.9 (1.4 to 2.6)	2.5 (1.5 to 4.0)
Scrotal swelling	49 (4)	3 (0.1)	58 (18 to 185)	170 (35 to 800)
Groin pain	37 (3)	34 (0.7)	3.9 (2.4 to 6.1)	6.8 (3.3 to 14.0)
Diagnoses[c]				
Orchitis/epididymis	153 (11)	37 (0.7)	15 (10 to 21)	13 (7.8 to 23.0)
Hydrocele	28 (2)	3 (0.1)	33 (10 to 109)	28 (7.7 to 100)
Investigations				
Raised inflammatory markers	51 (4.0)	78 (2.0)	2.3 (1.6 to 3.3)	4.3 (2.5 to 7.5)
Interactions				
Testicular pain with testicular swelling				0.008 (0.001 to 0.08)[d]

[a] *The univariate likelihood ratio, showing the likelihood of having a specific symptom in a patient with testicular cancer, compared with the likelihood of having it in a patient without cancer.* [b] *In multivariate conditional logistic regression, containing all nine variables.* [c] *These are initial diagnoses, potentially misclassified. All features:* P<0.001. [d] P<0.001.

Twenty one cases (1.5%) and 76 controls (1.53%) had a record of a fracture (*P*<0.94).

Positive predictive values

PPVs for individual and combined features of testicular cancer in men aged 17–49 years are shown in Figure 2. The risk for a single symptom is shown in the top line, with paired features shown below. Testicular lump, testicular swelling, and hydrocele each produced risk estimates >1% (2.5%, 2.3%, and 1.1%, respectively). For the combined features, the highest PPVs were produced when testicular lump was recorded with other symptoms — testicular swelling (17%), testicular pain (10%), and abdominal pain (6.1%). Testicular pain and abdominal pain together also produced a PPV of 6.1%. PPVs for repeat consultations of testicular pain (1.1%) and the epididymis group (1%) are also shown. Other non-lump combinations also had risk estimates above the NICE 3% threshold for urgent referral, including testicular pain with raised inflammatory markers, or with groin pain (both 3.7%). Testicular pain had a multivariable odds ratio of 38 (95% CI = 22 to 68), and a PPV of 0.4%.

PPVs for single features in the ≥50 age group were mostly 1%, except for scrotal swelling at 2.1%. In contrast to the younger group, combining testicular lump with testicular pain or testicular swelling reduced the PPV to 0.7% and 0.3%, respectively. Testicular swelling with orchitis/epididymitis produced the highest combined PPV, of 1.3% in this ≥50 group.

DISCUSSION

Summary

This is the first study to report the clinical features of testicular cancer in primary care. Certain secondary care-reported symptoms were also associated with testicular cancer in primary care: PPVs for testicular lump and testicular swelling were both >2%. This risk increased greatly when testicular lump was combined with other symptoms, such as testicular pain (10%). Other non-lump combinations also had risk estimates above the NICE 3% threshold for urgent referral, including testicular pain with raised inflammatory markers, or with groin pain (both 3.7%). New significant findings of hydrocele and epididymis-orchitis, especially prevalent in the 3 months before testicular cancer diagnosis, suggest a possible misdiagnosis for the former, or a possible complication of the cancer for the latter.

These results largely support the UK recommendations for investigation of possible testicular cancer, though they provide additional information as to which symptoms or symptom combinations warrant ultrasound. They also identify possible sources of delay in cancer diagnosis, such as mistakenly diagnosing a hydrocele, or considering a painful testicular mass to be low risk.

Strengths and limitations

This was a very large study, with nearly 1400 patients and with ample power. The quality and validity of CPRD data has been well reported.[25–27] The CPRD contains patients from across the UK, increasing the generalisability of the results. Information regarding the stage of the cancer at diagnosis was unavailable, though arguably this is less important than for other cancers, as advanced-stage testicular cancer is often curable. Furthermore, the authors had no ethnicity data.

The study is dependent upon the quality of primary care recording of symptoms, with the possibility that there are missing data.[28,29] It is possible to record clinical details in a field that is irretrievable to researchers (the 'free-text'). This loss of data matters if recording styles are unequal between cases and controls. One study in the CPRD itself has shown relatively minor

differential recording, with recognised high-risk features of cancer being disproportionately recorded in cases.[30] This disproportion tends to inflate PPVs of high-risk features and (to a lesser extent) depress PPVs of low-risk symptoms. The methods used here for estimating PPVs are well established. However, despite the large population, several symptom combinations were too rare for reliable PPV estimates, particularly in the ≥50 years age group.

Similarly, studying existing records meant accepting the records verbatim. This was most relevant in the three variables relating to swellings — testicular lump, testicular swelling, and scrotal swelling. On paper, these mean different things, but it is possible that GPs use them interchangeably, especially the first two. The strengths of the associations with cancer for 'swelling' and 'lump' were very similar, so the distinction between the two may be relatively unimportant. However, when the enlargement was accompanied by pain, GPs seemed to use the term swelling differently from the word lump, with the former combination being much lower risk. The timings of the reported combined symptoms for each individual were not investigated, as previous analyses have shown it adds relatively little.

By identifying patient-reported symptoms using online forums in addition to literature reviews, pertinent features were unlikely to have been omitted. Finally, cases consulted more often. Thus, their doctors had more opportunity to record symptoms than for controls, potentially introducing recording bias, though the proxy for this — fractures — was reassuring.

Comparison with existing literature

Cases had approximately twice as many primary care consultations in the year before diagnosis as controls. This is somewhat at odds with a previous report that patients with testicular cancer have one of the lowest percentage of cases (16%), requiring three or more GP visits before diagnosis.[6] The findings from this paper suggest there remain opportunities for earlier diagnosis.

Testicular pain has previously been reported in primary care as having a 16-fold increased risk of testicular cancer.[9] In this study, the relatively low PPV of 0.4%, despite a high odds ratio of 38 (95% CI = 22 to 68), reflects the rarity of testicular cancer. An association with testicular cancer has not previously been reported for hydrocele or epididymis-orchitis, though the absolute risk was <1% in this study (Figure 2). Abdominal pain, previously reported in secondary care, was also associated, but with a very low PPV of 0.1%.[11] Its significance as an indicator of testicular cancer rises above the NICE threshold when accompanied by testicular pain, lump, or swelling.

A raised LDH was not seen in controls, so PPVs cannot be estimated for it; however, only 27 of the 71 cases in which the LDH was measured had a raised value (<2% of all cases). It is not possible to know if the LDH was taken because testicular cancer was suspected, or whether it was simply part of a battery of LFT tests. It is clear that a normal LDH does not rule out testicular cancer. The small number of positives do not, at this stage, support it being used as a diagnostic tool.

Implications for practice

Testicular lump and testicular swelling were each strongly associated with cancer in the <50 years age group, with PPVs of 2.5% and 2.3%, values just below the 3% threshold for urgent referral. The NICE guidance recommended urgent referral, using wording reflecting the lack of evidence at the time.[5] These results show that this recommendation was probably reasonable,

Figure 2. Positive predictive values for testicular cancer features in patients <50 years, for single and paired features.

Abdominal pain	Raised inflammatory markers	Groin pain	Orchitis/epididymis/ epididymis-orchitis	Testicular pain	Scrotal swelling	Hydrocele	Testicular swelling	Testicular lump	
0.04 0.03 to 0.05	0.05 0.03 to 0.07	0.07 0.04 to 0.12	0.3 0.2 to 0.4	0.4 0.3 to 0.7	0.7	1.1	2.3	2.5	**Risk as a single symptom**
	0.2	–	0.6	6.1	–	–	4.9	6.1	**Abdominal pain**
		–	–	3.7	–	–	–	0.4	**Raised inflammatory markers**
			0.2	3.7	–	–	–	–	**Groin pain**
			1.0	0.9	–	–	2.0	0.6	**Orchitis/epididymis/ epididymis-orchitis**
				1.1	–	–	1.0	10	**Testicular pain**
						–	3.7	3.7	**Scrotal swelling**
							–	–	**Hydrocele**
								17	**Testicular swelling**

The positive predictive value (PPV) is shown on the first line of each cell with the 95% confidence intervals (CIs) shown underneath. PPVs were not calculated if <five cases had the feature. Where <10 cases or controls had the combined features, CIs were omitted. The yellow shaded cells indicate a PPV of 1.0–1.9%; orange cells 2.0–4.9%; and red cells of ≥5.0%. The cells showing the same feature vertically and horizontally represent a second attendance with the same investigation.

especially as the guidance explicitly asked clinicians to use their judgement in deciding upon the merits of referral. Contrary to traditional teaching, painful testicular enlargement remains a strong predictor of cancer. In the >50 years age group, testicular enlargement collectively still produces a PPV of >3% (testicular lump 0.6%, testicular swelling 0.4%, and scrotal swelling 2.1%). However, pain appeared to reduce the risk from enlargement in contrast to younger men. This may represent an increased frequency of infection.

Recorded hydrocele, epididymis/orchitis, or testicular pain (especially when these lead to a second consultation) are important features of testicular cancer in primary care. It is plausible that either a hydrocele or infection could arise as a complication of testicular cancer. Even so, the fact that these presumptive or possibly erroneous diagnoses have an association with cancer makes them candidates for investigation. NICE guidance suggested ultrasound should be considered for 'unexplained or persistent' testicular symptoms. This study suggests that recurrent testicular pain, unresolving epididymis-orchitis, or hydrocele should come within this recommendation.

Other isolated features, shown in the upper row of Figure 2, are very low risk, and probably do not warrant initial investigation or referral, though good clinical practice would suggest some agreed form of review netting could be offered. However, non-testicular features, such as abdominal pain, groin pain, and raised inflammatory markers, all warrant referral when combined with testicular pain.

Ethical approval

Ethical approval was given by the Independent Scientific Advisory Committee — protocol 09-110.

Provenance

Freely submitted; externally peer reviewed.

Competing interests

William Hamilton was clinical lead on the 2015 revision of the NICE guidance on investigation of suspected cancer. His contribution to this article is in a personal capacity, and does not represent the view of the Guideline Development Group, or of NICE itself. Elizabeth Shephard has declared no competing interests.

Discuss this article

Contribute and read comments about this article: **bjgp.org/letters**

REFERENCES

1. Cancer Research UK. Testicular cancer statistics. 2018. http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/testicular-cancer (accessed 22 Jun 2018).

2. Cancer Research UK. Testicular cancer incidence statistics. 2016. http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/testicular-cancer/incidence#heading-Six (accessed 15 Jun 2018).

3. Cancer Research UK. Testicular cancer statistics. Testicular cancer survival. 2018. http://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/testicular-cancer#heading-Two (accessed 22 Jun 2018).

4. Blom JW, Vanderschoot JP, Oostindier MJ, *et al.* Incidence of venous thrombosis in a large cohort of 66 329 cancer patients: results of a record linkage study. *J Thromb Haemost* 2006; **4(3):** 529–535.

5. National Institute for Health and Care Excellence. *Suspected cancer: recognition and referral. NG12.* 2017. http://www.nice.org.uk/guidance/NG12 (accessed 15 Jun 2018).

6. Lyratzopoulos G, Neal RD, Barbiere JM, *et al.* Variation in number of general practitioner consultations before hospital referral for cancer: findings from the 2010 National Cancer Patient Experience Survey in England. *Lancet Oncol* 2012; **13:** 353–365.

7. Neal RD, Din NU, Hamilton W, *et al.* Comparison of cancer diagnostic intervals before and after implementation of NICE guidelines: analysis of data from the UK General Practice Research Database. *Br J Cancer* 2014; **110(3):** 584–592.

8. Din N, Ukoumunne O, Rubin G, *et al.* Age and gender variations in cancer diagnostic intervals in 15 cancers: analysis of data from the UK Clinical Practice Research Datalink. *PLoS One* 2015; **10(5):** e0127717.

9. Hippisley-Cox J, Coupland C. Symptoms and risk factors to identify men with suspected cancer in primary care: derivation and validation of an algorithm. *Br J Gen Pract* 2013; DOI: https://doi.org/10.3399/bjgp13X660724.

10. Gascoigne P, Mason MD, Roberts E. Factors affecting presentation and delay in patients with testicular cancer: results of a qualitative study. *Psycho-Oncology* 1999; **8(2):** 144–154.

11. Oliver RT. Factors contributing to delay in diagnosis of testicular tumours. *Br Med J (Clin Res Ed)* 1985; **290(6465):** 356.

12. Ozturk C, Fleer J, Hoekstra HJ, Hoekstra-Weebers JE. Delay in diagnosis of testicular cancer; a need for awareness programs. *PLoS One* 2015; **10(11):** e0141244.

13. Toklu C, Ozen H, Sahin A, *et al.* Factors involved in diagnostic delay of testicular cancer. *Int Urol Nephrol* 1999; **31(3):** 383–388.

14. Germà-Lluch JR, Garcia del Muro X, Maroto P, *et al.* Clinical pattern and therapeutic results achieved in 1490 patients with germ cell tumours of the testis: the experience of the Spanish Germ Cell Cancer Group (GG). *Eur Urol* 2002; **42(6):** 553–562.

15. Milose JC, Filson CP, Weizer AZ, *et al.* Role of biochemical markers in testicular cancer: diagnosis, staging, and surveillance. *Open Access J Urol* 2011; **4:** 1–8.

16. Shephard E, Stapley S, Neal RD, *et al.* Clinical features of bladder cancer in primary care. *Br J Gen Pract* 2012; DOI: https://doi.org/10.3399/bjgp12X654560.

17. Stapley S, Peters TJ, Neal RD, *et al.* The risk of oesophagogastric cancer in symptomatic patients in primary care: a large case–control study using electronic records. *Br J Cancer* 2013; **108(1):** 25–31.

18. Shephard EA, Neal RD, Rose P, *et al.* Clinical features of kidney cancer in primary care: a case–control study using primary care records. *Br J Gen Pract* 2013; DOI: https://doi.org/10.3399/bjgp13X665215.

19. Walker S, Hyde C, Hamilton W. Risk of breast cancer in symptomatic women in primary care: a case–control study using electronic records. *Br J Gen Pract* 2014; DOI: https://doi.org/10.3399/bjgp14X682873.

20. Shephard E, Neal R, Rose P, *et al.* Quantifying the risk of myeloma from symptoms reported in primary care patients: a large case–control study using electronic records. *Br J Gen Pract* 2015; DOI: https://doi.org/10.3399/bjgp15X683545.

21. Shephard EA, Neal RD, Rose PW, *et al.* Quantifying the risk of non-Hodgkin lymphoma in symptomatic primary care patients aged ≥40 years: a large case–control study using electronic records. *Br J Gen Pract* 2015; DOI: https://doi.org/10.3399/bjgp15X684793.

22. Shephard EA, Neal RD, Rose PW, *et al.* Quantifying the risk of Hodgkin lymphoma in symptomatic primary care patients aged ≥40 years: a case–control study using electronic records. *Br J Gen Pract* 2015; DOI: https://doi.org/10.3399/bjgp15X684805.

23. Shephard EA, Neal RD, Rose PW, *et al.* Symptoms of adult chronic and acute leukaemia before diagnosis: large primary care case–control studies using electronic records. *Br J Gen Pract* 2016; DOI: https://doi.org/10.3399/bjgp16X683989.

24. Office for National Statistics. Cancer registration statistics 2008. http://webarchive.nationalarchives.gov.uk/20160106061901/http://www.ons.gov.uk/ons/publications/re-reference-tables.html?edition=tcm%3A77-394268 (accessed 22 Jun 2018).

25. Khan NF, Harrison SE, Rose PW. Validity of diagnostic coding within the General Practice Research Database: a systematic review. *Br J Gen Pract* 2010; DOI: https://doi.org/10.3399/bjgp10X483562.

26. Thiru K, Hassey A, Sullivan F. Systematic review of scope and quality of electronic patient record data in primary care. *BMJ* 2003; **326:** 1070.

27. Shephard E, Stapley S, Hamilton W. The use of electronic databases in primary care research. *Fam Pract* 2011; **28(4):** 352–354.

28. Sollie A, Sijmons RH, Helsper C, Numans ME. Reusability of coded data in the primary care electronic medical record: a dynamic cohort study concerning cancer diagnoses. *Int J Med Inform* 2017; **99:** 45–52.

29. Tate AR, Martin AGR, Ali A, Cassell JA. Using free text information to explore how and when GPs code a diagnosis of ovarian cancer: an observational study using primary care records of patients with ovarian cancer. *BMJ Open* 2011; **1(1):** e000025.

30. Price SJ, Stapley SA, Shephard E, *et al.* Is omission of free text records a possible source of data loss and bias in Clinical Practice Research Datalink studies? A case–control study. *BMJ Open* 2016; **6(5):** e011664.