

## ARTICLE OPEN



## Epidemiology

# The impact of changing risk thresholds on the number of people in England eligible for urgent investigation for possible cancer: an observational cross-sectional study

Sarah F. Moore<sup>1</sup>✉, Sarah J. Price<sup>1</sup>, Sarah Chowienczyk<sup>2</sup>, Jennifer Bostock<sup>3</sup> and Willie Hamilton<sup>1</sup>

© The Author(s) 2021

**BACKGROUND:** Expediting cancer diagnosis may be achieved by targeted decreases in referral thresholds to increase numbers of patients referred for urgent investigation.

**METHODS:** Clinical Practice Research Datalink data from England for 150,921 adults aged  $\geq 40$  were used to identify participants with features of possible cancer equating to risk thresholds  $\geq 1\%$ ,  $\geq 2\%$  or  $\geq 3\%$  for breast, lung, colorectal, oesophago-gastric, pancreatic, renal, bladder, prostatic, ovarian, endometrial and laryngeal cancers.

**RESULTS:** The mean age of participants was 60 (SD 13) years, with 73,643 males (49%). In 2016, 8576 consultation records contained coded features having a positive predictive value (PPV) of  $\geq 3\%$  for any of the 11 cancers. This equates to a rate of 5682/100,000 patients compared with 4601/100,000 Suspected Cancer NHS referrals for these cancers from April 2016–March 2017. Nine thousands two hundred ninety-one patient-consultation records had coded features equating to a  $\geq 2\%$  PPV, 8% more than met  $PPV \geq 3\%$ . Similarly, 19,517 had features with a  $PPV \geq 1\%$ , 136% higher than for  $PPV \geq 3\%$ .

**CONCLUSIONS:** This study estimated the number of primary-care patients presenting at lower thresholds of cancer risk. The resource implications of liberalising this threshold to 2% are modest and manageable. The details across individual cancer sites should assist planning of English cancer services.

*British Journal of Cancer* (2021) 125:1593–1597; <https://doi.org/10.1038/s41416-021-01541-4>

## INTRODUCTION

Prompt diagnosis of symptomatic cancer is a priority for patients, the public and governments worldwide. This is particularly so in the UK, which has worse survival from cancer than most comparator countries [1, 2]. In the UK, most patients with symptoms of possible cancer report initially to primary care. In England, selection of patients for specialist investigations for possible cancer is supported by recommendations in the 2015 'Suspected cancer: recognition and referral,' NG12 guideline published by National Institute for Health Care Excellence [3]. This guidance describes presenting features (symptoms, signs and abnormal blood test results) warranting investigation. It established the principle of selecting patients for cancer investigation based on the estimated cancer risk, and chose a threshold of 3% risk warranting urgent investigation. This value was chosen, as a balance between the benefit of liberalised investigation (compared with the previous value of  $\sim 5\%$ ) and the potential harms of overwhelming clinical services, or over-investigation [3]. Furthermore, for colorectal, lung and ovarian cancers, NG12 supports testing overseen within primary care at levels below 3%, using faecal immunochemical testing (FIT), chest X-ray and Cancer Antigen 125 (Ca125), respectively. Times to diagnosis of patients

with the lower-risk presenting features of colorectal cancer have reduced since NG12 was introduced [4].

Efforts continue to improve England's cancer outcomes, with an explicit target of 75% of new cancers to be diagnosed at stage I or II (i.e. probably curable) by 2028, from the current level of around 54% [2]. One possible intervention to facilitate this would be to lower the threshold risk for urgent investigation to 2%, or even 1%, for all cancers, which could expedite the diagnosis for some patients. Such a change would place further demand for investigations, on top of a considerable recent increase [5]. Demand for lower gastrointestinal endoscopy for all indications doubled from 2012 to 2017 [6]. Imaging figures are similar [7]. Furthermore, there has been unprecedented strain on services during the COVID-19 pandemic, which severely curtailed diagnostic services [8]. Cancer diagnosis is also costly: over £1bn is spent annually in the UK on it [9]. Therefore, any change to the threshold risk for cancer investigations requires estimates of the potential increased demand.

In this study, we estimated how many patients in primary care in England would be eligible for urgent cancer investigations either overseen in primary care or performed in secondary care at

<sup>1</sup>University of Exeter Medical School, College of Medicine & Health, St Luke's Campus, Magdalen Road, Exeter, UK. <sup>2</sup>Royal Devon and Exeter Hospital, Barrack Road, Exeter, UK. <sup>3</sup>Policy Research Unit in Cancer Awareness, Screening and Early Diagnosis, Queen Mary University of London, Mile End Rd, Bethnal Green, London, UK. ✉email: s.moore3@exeter.ac.uk

Received: 16 April 2021 Revised: 27 July 2021 Accepted: 24 August 2021

Published online: 16 September 2021

a current risk threshold of  $\geq 3\%$ . We extended this to provide estimates for putative  $\geq 1\%$  and  $\geq 2\%$  risk thresholds.

## METHODS

### Study design and setting

This was an observational cross-sectional study set in primary care using data from the UK's Clinical Practice Research Datalink (CPRD GOLD). The study period was 1 January 2015 to 31 December 2016, to allow full capture of participants meeting criteria in 2016. We studied participants' records for the symptoms, signs and abnormal test results (collectively termed "features" hereafter) for the eleven most common internal cancers: bladder, breast, colon or rectum, endometrium, kidney, larynx, lung, oesophago-gastric, ovary, pancreas and prostate. External cancers that are associated with easily visible signs on clinical examination, such as skin cancers, and haematological cancers, commonly diagnosed through blood tests, were not studied.

### Sample selection criteria

A random sample of 200,000 CPRD participants was supplied by the CPRD, using the following inclusion criteria: aged 40 or over; continuous registration at a CPRD general practice for at least 1 year before 1 January 2016; and complete CPRD records of up-to-date standard throughout the study period. Only participants registered at practices in England were included in our analyses, as there are different diagnostic pathways and guidance across the devolved nations. This left a final study population of 150,921.

### Variables

**Participant characteristics.** We calculated age from the year of birth. Month and day of birth are not available with CPRD data (to protect anonymity) so we assigned a birthday of 1 July. Smoking status adopted the NG12 categories of "never-smoker" and "ever-smoker".

**Identification of features.** We derived a list of features of cancer from the systematic reviews published in NG12 for each of the cancer sites (Appendix A). For laryngeal cancer, no primary-care evidence was available when NG12 was being drawn up so features representing  $\geq 1\%$  risk of this cancer were identified from a subsequently published study [10].

We assembled CPRD code lists representing each chosen feature, mostly from a previous study [4], supplemented by a small number of new ones, all created using robust methods [11]. These code lists are available in Appendix B.

We searched each participant's medical records using the code lists, and participant-level variables were created for the presence/absence of each feature. In line with usual practice, we interpreted the absence of a code for a feature as an absence of the feature itself [12].

For abnormal blood test results, we compared a participant's result with the local laboratory's reference range. The exceptions to this were tests for anaemia, prostate-specific antigen (PSA) and CA125. For PSA, age-specific values were used:  $\geq 3.0$  ng/ml for participants aged 50–69 years and 4.0 ng/ml for  $\geq 70$  years [13]. To define anaemia, we used haemoglobin level  $< 130$  g/L in men and  $< 120$  g/L in women [14]. The threshold for CA125 was 35 U/L, as per NG12.

**Identification of participants who meet criteria in NG12.** We matched each participant's features with the criteria for urgent investigation in NG12 by cancer site (Appendix A). Appendix A also describes the relevant study or meta-analysis underpinning the selection of features, and identifies the positive predictive value (PPV) of each feature (or combination thereof) for each cancer and the population to which it should apply.

For criteria using paired or recurrent features, we required the second feature to occur within one year of the first, following the methodology in the literature underpinning many of the risk estimates [15].

Within cancer sites, participants were assigned an index date; namely, the date of the earliest time they met any criterion for that site between 1 January 2016 and 31 December 2016. For paired or recurrent features, the index date was the later of the two dates. Participants could have an index date for more than one cancer site.

**Assignment to risk bands.** Within cancer sites, participants were assigned to a risk band (1.0–1.9%, 2.0–2.9%,  $\geq 3\%$ ) according to the PPV on the index date. Participants with "suspected-cancer" codes, but without a coded

specific feature, were assumed to meet criteria for urgent investigation under the current NG12 guideline and were added to the  $\geq 3\%$  group. Participants were not assigned to a risk band for any cancers with which they had been diagnosed prior to 1 January 2016. For colorectal, lung and ovarian cancers, triage testing exists within primary care. Test positivity rates in the primary-care setting are estimated as: 15.9% (FIT) [16], 6% (chest X-ray), [Bradley, BJGP, in press] and CA125 (6.8%), respectively [17]. For all three tests, a positive result carries a risk of  $\geq 3\%$  plus a recommendation of specialist referral, so the appropriate percentage was assigned to the  $\geq 3\%$  group.

Within cancer sites, numbers qualifying at each risk band represent numbers of unique participants. The totals across cancer sites comprise multiple contributions from individual participants.

### Analysis

Simple descriptive statistics were used to report the cumulative numbers of participants meeting criteria at the  $\geq 1\%$ ,  $\geq 2\%$  and  $\geq 3\%$  risk thresholds within cancer sites. The numbers of participants meeting criteria at each threshold were then converted to a rate per 100,000 and reported with 95% confidence intervals. We estimated the rate (per 100,000 of the population) in England eligible for urgent cancer investigation using totals across all cancer sites for each risk band. These estimates are based on Office for National Statistics data on the English population aged over 40 (27.5 million in 2016; 13.2 million males, and 14.3 million females) [18]. The numbers of actual numbers of 2-week-wait appointments in England for 2016 are reported for comparison [19].

### Study size

There were 1.88 million urgent suspected-cancer appointments in April 2016–March 2017, representing 6.8% of the English population of 27.5 million aged over 40 [18, 20]. A sample of 150,000 gives a margin of error of 0.2 percentage points around an estimate of 6.8%.

### Patient and public involvement

A member of the Patient and Public Involvement (PPI) group affiliated with the Policy Research Unit (PRU) gave feedback on the lay summary of ISAC ethics protocol application to the CPRD for this project and is a co-author on this manuscript. The PRU's PPI group comprises lay members with an interest as public, patients and/or carers of those with cancer or suspected cancer, and brings the public perspective to our research.

## RESULTS

### Participants

Our sample consisted of 150,921 participants registered at a general practice in England. Participant characteristics are described in Table 1. Comparison to Office for National Statistics UK population data from mid-2016 showed that the sample received was representative in terms of age and sex of the UK population [18].

### Participants meeting specific cancer risk thresholds

The features of cancer are included in Appendix A. These are listed by cancer site, identifying criteria for urgent investigation, the positive predictive value (PPV) of each feature (or combination thereof) for that cancer and the population to which they should apply. Appendix A also cites the relevant study or meta-analysis underpinning the selection.

**Table 1.** Participant demographics.

	Number of participants Number (%)	Age in mid-2016 (years) Mean (SD)
Male	73,643 (48.8)	59.8 (13.6)
Female	77,278 (51.2)	61.5 (14.4)
Total	150,921 (100.0)	60.7 (13.9)

**Table 2.** Number of participants with features of possible cancer in selected risk bands (and comparison to PPV  $\geq$  3%), in 150,921 English patients aged  $\geq$ 40 years.

	PPV $\geq$ 3%						PPV $\geq$ 2%						PPV $\geq$ 1%			
	Number qualifying	Additional number qualifying for PC testing	Estimated number positive tests	Adjusted number qualifying for referral after PC testing	Adjusted number/100,000 population after PC testing [95% CI]	Number qualifying	Additional number qualifying for PC testing	Estimated number positive tests	Adjusted number qualifying for referral after PC testing	Adjusted number/100,000 population after PC testing [95% CI]	% change from PPV $\geq$ 3%	Number qualifying	Adjusted number/100,000 population after PC testing [95% CI]	% change from PPV $\geq$ 3%	Number/100,000 population [95% CI]	% change from PPV $\geq$ 3%
Colorectal	5125	1776	282	5407	3583 [3489, 3678]	5440	1461	232	5672	3758 [3663, 3855]	5	6901	3758 [3663, 3855]	28	4573 [4468, 4679]	28
Prostate	894			894	592 [554, 632]	1046			1046	693 [652, 736]	17	1126	693 [652, 736]	26	746 [703, 791]	26
Lung	803	2503	30	833	552 [515, 591]	963	2343	28	991	657 [617, 699]	19	3306	657 [617, 699]	297	2191 [2117, 2266]	297
Breast	293			293	194 [173, 218]	311			311	206 [183, 231]	6	474	206 [183, 231]	62	314 [286, 344]	62
Kidney	263			263	174 [154, 197]	263			263	174 [154, 197]	0	961	174 [154, 197]	265	637 [597, 678]	265
Bladder	246			246	163 [143, 184]	246			246	163 [143, 185]	0	841	163 [143, 185]	242	557 [520, 596]	242
Upper GI	211			211	140 [122, 160]	249			249	165 [145, 187]	18	5044	165 [145, 187]	2291	3342 [3252, 3434]	2291
Ovary	128	31	2	130	86 [72, 102]	149	10	1	150	99 [84, 117]	15	159	99 [84, 117]	22	105 [90, 123]	22
Endometrium	109			109	72 [59, 87]	124			124	82 [68, 98]	14	244	82 [68, 98]	124	162 [142, 183]	124
Larynx	111			111	74 [61, 89]	142			142	94 [79, 111]	28	142	94 [79, 111]	28	94 [79, 111]	28
Pancreas	78			78	52 [41, 65]	97			97	64 [52, 78]	24	319	64 [52, 78]	309	211 [189, 236]	309
Totals	8261 <sup>a</sup>			8576	5682 [5566, 5800]	9030 <sup>b</sup>			9291	6156 [6035, 6279]	8	19517 <sup>c</sup>	6156 [6035, 6279]	136	12932 [12763, 13102]	136

The risk bands are cumulative. Thus a participant in the  $\geq$ 3% risk band also appears in the  $\geq$ 2% risk band and the  $\geq$ 1% risk band.

<sup>a</sup>Upper GI upper gastrointestinal (oesophagus or stomach), PC primary care.

<sup>b</sup>Total of 8261 in 6898 unique participants.

<sup>c</sup>Total of 9030 in 7453 unique participants.

<sup>d</sup>Total of 19,517 in 10,649 unique participants.

## Outcomes

*Numbers qualifying for urgent investigation.* The cumulative numbers of participants with features of cancer at each threshold of increasing risk are reported as absolute numbers and per 100,000, by cancer site (Table 2). For the three cancer sites (colorectal, lung and ovary) where urgent investigation in primary care is available at the lower threshold of  $\geq 1\%$ , an additional column (under the heading 'Additional number qualifying for PC testing') shows the number of participants who would meet these lower threshold criteria. The estimated number of those qualifying for Primary Care testing who would have had a positive test result is shown under 'Estimated number positive results'. This is added to the numbers meeting the higher-risk criteria of  $\geq 2\%$  or  $\geq 3\%$  to give an adjusted number in our sample and an adjusted number/100,000 population meeting criteria for referral. The percentage change compared to the reference category of  $PPV \geq 3\%$  (the current English threshold for urgent investigation) is reported for thresholds  $\geq 2\%$  and  $\geq 1\%$ .

Comparison of estimates of participants eligible for urgent investigation to the actual numbers of patients referred urgently to secondary care in 2016–2017 in England showed a ratio of 1.2 between study estimates and actual numbers of appointments across the cancer sites for which NHS data were reported.

## DISCUSSION

### Key results

This study estimated that people consulted English primary care with features representing a  $\geq 3\%$  risk for one of the 11 commonest cancers at a rate of 5682 (95% CI 5566–5800) per 100,000 adults aged  $\geq 40$  in 2016. Liberalisation of criteria for referral from a  $PPV \geq 3\%$  to  $\geq 2\%$  across the 11 cancer sites would lead to 6156 (95% CI 6035–6279) patients per 100,000 meeting criteria, an 8% increase compared to  $PPV \geq 3\%$ . Further reduction to a threshold  $PPV$  of  $\geq 1\%$  would have increased numbers eligible for urgent investigation to 12,932 (95% CI 12,763–13,102) patients per 100,000, a 136% increase compared to  $PPV \geq 3\%$ .

### Strengths and limitations

This study used CPRD data, which is considered representative of primary care in England, the target population for application of NG12 [12, 21]. Furthermore, our code lists to identify the features of possible cancer were comprehensive and developed by robust published methods: most had been used before [4, 11]. However, there are limitations to CPRD data. Some symptoms of possible cancer will not have been recorded, or were recorded in the 'free text' area of the notes, which is unavailable to researchers for fear of compromising anonymity. This missing data will have reduced power, although we had sufficient power to give narrow confidence intervals for our primary analyses. Missing data may also have introduced misclassification bias for symptoms. Analysis of coded data alone underestimates the true number of participants with a symptom, by up to 30% for non-site-specific symptoms such as abdominal pain, and by 20% for symptoms that are specific and highly predictive of a particular cancer [22]. Thus our estimates of criteria based solely on symptoms, particularly for the 2% and 1% thresholds, are likely to be moderate underestimates. In contrast, test results are automatically coded into the CPRD record, effectively eliminating missing data for that aspect.

We did not attempt to quantify any possible benefits from a liberalisation of the threshold for specialist investigation. The relationship between timeliness of symptomatic cancer diagnosis and improvements in survival has not been studied empirically, although a higher cancer referral rate has been associated with reduced cancer mortality [23]. This relationship is complicated by there being at least three mechanisms whereby survival could be

improved: a stage shift, a within-stage shift or an avoidance of complications precipitating emergency admission. One modelling study of the COVID-19-related diagnostic delays estimated a 2-month delay on breast, colorectal and bladder cancers to lead to a reduction in 1-year net survival of 1.4–5.0%, 6.6–10.7% and 9.1–11.6%, respectively, where the ranges represent different age bands from 40 years upwards [8]. If harms from delayed diagnoses were of a similar magnitude to benefits from expedited diagnosis, then an improvement of even 1 week in timeliness of diagnosis could yield a survival benefit, perhaps as much as 1%. Such improvements in timeliness of diagnosis of symptomatic cancer are achievable, as seen following the introduction of NG12 in 2015 [4]. The benefit would not just be experienced by the modest percentage of cancer patients whose sole presenting features were in the 2.0–2.9% range, but also by those who presented with low-risk features before developing higher-risk features. However, we would caution against assuming that an 8% increase in resource use would yield an 8% increase in—for example—improved stage at diagnosis. It may do so—but it may not.

The disparity between the rate of patients eligible for specialist cancer investigation and the rate of actual appointments offered in 2016 was 20%. This is no surprise: it has long been recognised that GPs do not slavishly adhere to guidance, including NG12. Indeed, NG12 explicitly condones this, supporting GPs in making a holistic assessment of the patient's risk—ideally coming to a shared decision about the need for investigation.

### Use of the results—implications for policy and practice

There are powerful policy implications from this study, particularly relating to any liberalisation of criteria for urgent investigation. Based on our 11 cancers, it appears that a liberalisation to a 2% threshold would have a modest impact on investigative resource use, of around 8%. This increase in resources would not occur evenly across all specialities: in percentage terms, patients with possible pancreatic or laryngeal cancer would see the largest increase.

For colorectal, lung and ovarian cancers, any increase in demand for specialist diagnostic services should be mitigated by primary-care testing. This principle is already present in NG12—explicitly supporting investigation of possible cancer in primary care for risks below 3% where a good primary-care test is available [3]. We incorporated such testing in our methods. However, primary-care testing is not available for all cancers. For upper gastrointestinal cancers there is no current primary-care test that can triage patients with low-risk features, so the additional investigative burden would fall on endoscopy services, already under considerable strain [24]. The resource implications of lowering the investigation threshold to 1% are much larger. In particular, eligibility for possible gastrointestinal cancer testing would increase over 30 fold.

Lowering the threshold also has other implications beyond the scope of this study, in terms of opportunity costs, and potentially increasing patient anxiety and unnecessary investigation.

## CONCLUSIONS

These results provide estimates of the impact of changing referral thresholds for suspected cancer on numbers of patients referred across 11 different cancer sites, which can be used to inform future NHS policy in England. The detail of this study across individual cancer sites and criteria will help future planning of cancer services.

### DATA AVAILABILITY

The anonymised participant data from this study are not available, in line with the CPRD's data security policy. CPRD code libraries are available from the authors on request.

## REFERENCES

- Allemani C, Matsuda T, Carlo VD, Harewood R, Matz M, Nikšić M, et al. Global surveillance of trends in cancer survival 2000–14 (CONCORD-3): analysis of individual records for 37 513 025 patients diagnosed with one of 18 cancers from 322 population-based registries in 71 countries. *The Lancet*. 2018;391:1023–75.
- NHS England. The NHS Long Term Plan. 2019. <https://www.longtermplan.nhs.uk/wp-content/uploads/2019/08/nhs-long-term-plan-version-1.2.pdf>.
- NICE. Suspected cancer: recognition and referral [NG12]. 2015. <http://www.nice.org.uk/guidance/NG12>.
- Price S, Spencer A, Zhang X, Ball S, Lyratzopoulos G, Mujica-Mota R, et al. Trends in time to cancer diagnosis around the period of changing national guidance on referral of symptomatic patients: A serial cross-sectional study using UK electronic healthcare records from 2006–17. *Cancer Epidemiol*. 2020;69:101805.
- Cancer Services - Public Health England. <https://fingertips.phe.org.uk/profile/cancerservices/data>.
- Richards M, Thorlby R, Fisher R, Turton C. Unfinished business. 2018. <https://www.health.org.uk/publications/unfinished-business>.
- 2020 Delivery. Horizon Scanning. 2015. [https://www.cancerresearchuk.org/sites/default/files/horizon\\_scanning\\_-\\_final.pdf](https://www.cancerresearchuk.org/sites/default/files/horizon_scanning_-_final.pdf).
- Sud A, Torr B, Jones ME, Broggio J, Scott S, Loveday C, et al. Effect of delays in the 2-week-wait cancer referral pathway during the COVID-19 pandemic on cancer survival in the UK: a modelling study. *Lancet Oncol*. 2020; 21:1035–44.
- Hamilton W. Diagnosing symptomatic cancer in the NHS. *BMJ*. 2015;351:h5311.
- Shephard EA, Parkinson MAL, Hamilton WT. Recognising laryngeal cancer in primary care: a large case-control study using electronic records. *Br J Gen Pract*. 2019;69:e127.
- Watson J, Nicholson BD, Hamilton W, Price S. Identifying clinical features in primary care electronic health record studies: methods for codelist development. *BMJ Open*. 2017;7:e019637.
- Herrett E, Gallagher AM, Bhaskaran K, Forbes H, Mathur R, van Staa T, et al. Data resource profile: Clinical Practice Research Datalink (CPRD). *Int J Epidemiol* 2015;44:827–36.
- NICE. Information about PSA testing. 2017. <https://cks.nice.org.uk/topics/prostate-cancer/diagnosis/psa-testing/>.
- NICE. What investigations should I arrange to confirm iron deficiency anaemia? 2018. <https://cks.nice.org.uk/topics/anaemia-iron-deficiency/diagnosis/investigations/>.
- Hamilton W. The CAPER studies: five case-control studies aimed at identifying and quantifying the risk of cancer in symptomatic primary care patients. *Br J Cancer*. 2009;101(Suppl 2):S80–6.
- Bailey SER, Abel GA, Atkins A, Byford R, Davies S-J, Mays J, et al. Diagnostic performance of a faecal immunochemical test for patients with low-risk symptoms of colorectal cancer in primary care: an evaluation in the South West of England. *Br J Cancer*. 2021;124:1231–6.
- Funston G, Hamilton W, Abel G, Crosbie EJ, Rous B, Walter FM. The diagnostic performance of CA125 for the detection of ovarian and non-ovarian cancer in primary care: a population-based cohort study. *PLoS Med*. 2020;17: e1003295.
- Office of National Statistics. Estimates of the population for the UK, England and Wales, Scotland and Northern Ireland: mid 2016 edition. 2020. Accessed 1/10/2020. <https://www.ons.gov.uk/peoplepopulationandcommunity/populationandmigration/populationestimates/timeseries/ukpop/pop>.
- Samuels M, Dai D, McDonnell P. Waiting Times for Suspected and Diagnosed Cancer Patients 2016-17 Annual Report. NHS England website, in electronic PDF format only: NHS England; 2017.
- Statistics » Cancer Waiting Times Annual Report, 2016–17 [Internet]. [cited 2021 Sep 14]. Available from: <https://www.england.nhs.uk/statistics/statistical-work-areas/cancer-waiting-times/cwt-annual-reports/cwt-annual-report-2016-17/>.
- Rubin G, Berendsen A, Crawford SM, Domett R, Earle C, Emery J, et al. The expanding role of primary care in cancer control. *Lancet Oncol*. 2015;16:1231–72.
- Price SJ, Stapley SA, Shephard E, Barraclough K, Hamilton WT. 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:e011664.
- Møller H, Gildea C, Meechan D, Rubin G, Round T, Vedsted P. Use of the English urgent referral pathway for suspected cancer and mortality in patients with cancer: cohort study. *BMJ*. 2015. <https://www.bmj.com/content/351/bmj.h5102>
- Health Services Management Centre. SCOPING THE FUTURE. 2015. [https://www.cancerresearchuk.org/sites/default/files/scoping\\_the\\_future\\_-\\_final.pdf](https://www.cancerresearchuk.org/sites/default/files/scoping_the_future_-_final.pdf).

## ACKNOWLEDGEMENTS

Not applicable.

## AUTHOR CONTRIBUTIONS

All authors were involved in drafting and revising the manuscript, approved the final version and agree to be accountable for all aspects of the work in ensuring that questions related to the accuracy and integrity of the work are appropriately investigated and resolved. In addition: SFM designed the study, compiled the tables of evidence (Appendix A), created statistical models to analyse the data and interpreted the results. SJP conceived the study, created statistical models to analyse the data and interpreted the results. SC conceived the study and interpreted the results. JB conceived the study. WH conceived the study and interpreted the results.

## FUNDING

This research was funded by the National Institute for Health Research (NIHR) Policy Research Programme, conducted through the Policy Research Unit in Cancer Awareness, Screening and Early Diagnosis, PR-PRU-1217-21601. SFM is funded by an NIHR Academic Clinical Fellowship, ACF-2015-23-501. The views expressed are those of the authors and not necessarily those of the NIHR or the Department of Health and Social Care.

## ETHICS APPROVAL

Ethics approval for the use of anonymised electronic patient records in this study was granted from the Clinical Practice Research Datalink's (CPRD) Independent Scientific Advisory (ISAC) Board (protocol 19\_124R2A) on 11 July 2019.

## COMPETING INTERESTS

All authors have completed the Unified Competing Interest form (available on request from the corresponding author) and declare: SFM reports grants from National Institute for Health Research (NIHR) Academic Clinical Fellowship ACF-2015-23-501, during the conduct of the study; SJP reports grants from NIHR Policy Research Programme, conducted through the Policy Research Unit in Cancer Awareness, Screening and Early Diagnosis. PR-PRU-1217-21601, during the conduct of the study; WH reports grants from University of Exeter, during conduct of the study; remaining authors received no support from any organisation for the submitted work; no financial relationships with any organisations that might have an interest in the submitted work in the previous three years, no other relationships or activities that could appear to have influenced the submitted work.

## ADDITIONAL INFORMATION

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s41416-021-01541-4>.

**Correspondence** and requests for materials should be addressed to Sarah F. Moore.

**Reprints and permission information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021