BMJ Open

Relationship between ethnicity and stage at diagnosis in England: a national analysis of six cancer sites

Anna Fry,[1,2] Becky White ,[2,3] Diana Nagarwalla ,[2] Jon Shelton,[2] Ruth H Jack [4]

To cite: Fry A, White B, Nagarwalla D, *et al*. Relationship between ethnicity and stage at diagnosis in England: a national analysis of six cancer sites. *BMJ Open* 2023;**13**:e062079. doi:10.1136/bmjopen-2022-062079

► Prepublication history and additional supplemental material for this paper are available online. To view these files, please visit the journal online (http://dx.doi.org/10.1136/bmjopen-2022-062079).

Received 16 February 2022
Accepted 25 November 2022

[1]National Cancer Registration and Analysis Service, NHS Digital, London, UK
[2]Cancer Intelligence, Cancer Research UK, London, UK
[3]ECHO (Epidemiology of Cancer Healthcare & Outcomes), Department of Behavioural Science & Health, Institute of Epidemiology & Health Care, University College London, London, UK
[4]Centre for Academic Primary Care, Lifespan and Population Health, School of Medicine, University of Nottingham, Nottingham, UK

Correspondence to
Diana Nagarwalla;
diana.nagarwalla@cancer.org.uk

ABSTRACT

Objectives Cancer stage at diagnosis is a determinant of treatment options and survival. Previous research has shown differences in barriers to presentation with cancer between ethnic groups. The completeness and quality of cancer stage and ethnicity data has improved markedly over recent years in England, allowing for comparison of stage distributions at diagnosis between ethnic groups. This study aimed to assess relationships between ethnic group and two outcomes: unknown stage cancer and late stage (stages 3 and 4) cancer, after adjustment for confounders.

Design and setting A retrospective secondary data analysis using data from NHS Digital's National Cancer Registration and Analysis Service and Hospital Episode Statistics records from 2012 to 2016.

Participants This study analysed newly diagnosed breast, colon, non-small cell lung cancer (NSCLC), ovary, prostate and uterine cancers in white British, Caribbean, African, Chinese and Asian patients aged 15–99 in England.

Results Caribbean, African and Asian women with breast or ovarian cancer, Caribbean and African women with uterine or colon cancer, Caribbean women with NSCLC and Caribbean men with colon cancer had increased odds of late-stage disease at diagnosis compared with the white British cohort. In contrast, Caribbean and African men with prostate cancer had decreased odds of late-stage cancer. Where stage was known, there were variations in late-stage cancer by ethnic group.

Conclusions Low symptom awareness and barriers to presentation can cause delays, resulting in later stage diagnosis. Targeted intervention campaigns to help raise awareness of cancer signs and symptoms and the benefits of early diagnosis, along with removing barriers to appropriate referrals, could help to improve these inequalities.

STRENGTHS AND LIMITATIONS OF THIS STUDY

⇒ This study used national-level data and granular ethnicity definition where possible.
⇒ The analysis was adjusted for confounders, including comorbidities, which may impact the timing of an individual's cancer diagnosis; previous studies do not consistently account for confounding, or cover a range of sites.
⇒ Information on an individual's education, religion, country of birth and language spoken at home was not available; these characteristics have been shown to have an impact on symptom knowledge and understanding of the healthcare system.
⇒ The requirement of hospital admission in defining Charlson Comorbidity Index score means that comorbidities managed in primary care are not accounted for; individuals with no Hospital Episode Statistics record are more likely to have unknown ethnicity and missing comorbidity.

INTRODUCTION

Cancer stage at diagnosis is a strong determinant of viable treatment options and survival.[1] Any delays between a first symptom and diagnosis can lead to the disease progressing and a later cancer stage at diagnosis. Previous research, particularly on breast cancer, suggests there are differences in proportions of patients diagnosed at different stages by ethnic group.[2 3] Delays could be due to patients (delaying seeking help), doctors (delaying ordering investigative tests) or system delays, either in primary or secondary care.[4] Several studies have reported differences in barriers to presentation with cancer between ethnic and socioeconomic groups, including difficulties organising and attending appointments, other practical and service barriers and emotional barriers.[5–9]

Historically, the completeness of national cancer stage information in English population-based cancer registry data has not been high enough to allow for robust assessment of links between stage at diagnosis and ethnicity. In recent years, the completeness of both stage and ethnicity information has improved substantially, and now allows for investigation of the association between ethnicity and stage at diagnosis for common cancers. Although cancer stage data is improving in completeness, some tumours remain unstaged. This may indicate that the stage was never ascertained, as certain tumour morphologies cannot be staged, it might

have been inappropriate to stage the patient (eg, if the patient was too frail for surgery), the patient died prior to staging, or alternatively, data was not transferred to the national registry, either due to incomplete recording or the diagnosis was made outside of the National Health Service (NHS).[1] Old age has been found to be associated with missing stage independent of comorbidities and short-term mortality.[10]

Previous research on associations between ethnicity and stage at diagnosis in England has been limited geographically and to certain cancer sites. Lyratzopoulos *et al* assessed multiple cancer sites, however, missing ethnicity data meant ethnic variation in stage at diagnosis was not explored.[11]

This paper contributes to the cancer literature by assessing, for the first time, the associations between ethnicity and stage at diagnosis across England for six cancer sites, adjusting for patient case-mix and using a granular ethnicity definition where possible.

METHODS

Study population

Patient information was obtained from NHS Digital's National Cancer Registration and Analysis Service[12] for incident cases of six cancer sites: female breast (International Classification of Diseases 10th Revision (ICD-10) C50), colon (C18–C19), non-small cell lung cancer (NSCLC) (C33–C34 cases excluding confirmed small cell lung cancer, defined in online supplemental table 1), ovary (C56–C57 and female C48 cases excluding ICD-O-2 8800–8806, 8963, 8990, 8991, 9040–9044, 8811–8921, 9120–9373, 9530–9582), prostate (C61) and uterine (C54–C55). Cancer sites were chosen on the basis of high staging completeness and sufficient counts in the smaller ethnic groups.

The five ethnic categories used were: white British, Caribbean, African, Chinese and Asian. Indian, Bangladeshi and Pakistani patients were grouped into Asian, in order to boost statistical validity due to low numbers for certain combinations of site and ethnic group. This group did not include individuals with any other Asian background. The use of granular ethnicity groups where possible is particularly important because ethnicity subgroups have been found to have different risk factor prevalence and awareness.[13] Self-reported ethnic group from cancer registration data was primarily used, as this provided a single capture of ethnicity at the time of diagnosis. Where missing, it was supplemented with information from inpatient and outpatient Hospital Episode Statistics (HES) records. Ethnicity in HES is recorded separately for each hospital episode, and multiple ethnic groups may be recorded. In order to assign a single ethnicity from HES records for each registration without a recorded ethnicity, the most commonly recorded ethnicity within HES data was assigned (or the most recent of these in the event of a tied number of records). Patients with mixed ethnicities and other ethnic groups in the HES 16-group classification were excluded from this analysis given previous research that showed low validity of recording of these groups in HES when compared with self-reported survey data, perhaps as a result of differences between observed and self-reported inputs, or changes in how a person self-identifies over time.[14]

All newly diagnosed cancer registrations in people aged 15–99 diagnosed in 2012–2016 (years of high-quality stage and ethnicity data) in England were included in this analysis. Patients were assigned to quintiles of the income domain of the Indices of Deprivation 2015 (an area-based measure of deprivation), using their postcode of residence at diagnosis.[15] Cancer stage at diagnosis was mapped to tumour, node, metastases groups and categorised into early (stages 1–2), late (stages 3–4) or unknown stage. Tumours have an assigned Route to Diagnosis, summarising the pathway to diagnosis into one of eight routes: screen detected, 2-week wait, general practitioner (GP) referral, other outpatient/inpatient elective, emergency presentation, death certificate only and unknown.[16] Cases recorded from death certificate only were excluded. The comorbidity index (0, 1, 2, 3+) was derived based on the Charlson Comorbidity Index lookup table[17] using inpatient HES data, with the same methodology as described by Maringe *et al*, but using only inpatient data and a different time window: from 27 months to 3 months prior to the cancer diagnosis.[18] Individuals with missing comorbidity data, including those who had no record in HES, were included in the 0 category. England's Regions as defined by the Office for National Statistics were assigned to each individual based on their postcode of residence. In the 2012–2016 cohort, 3% of the patients analysed had multiple tumours; a patient may therefore be included multiple times in this analysis if they had multiple tumours.

Patient and public involvement

This work uses data that has been provided by patients and collected by the NHS as part of their care and support. No patients or public were involved in the design and conduct of this research.

Statistical analysis

Logistic regression models were used to assess the relationships between ethnicity and two outcomes: if the tumour had an unknown stage at diagnosis (the 'unstaged' model) and, if the stage was known, whether the tumour was diagnosed at early or late stage. All models were stratified by sex and adjusted for 10-year age band, comorbidity index, deprivation quintile, year of diagnosis and region. The model predicting unknown stage at diagnosis is further controlled for Route to Diagnosis since route has been found to vary by ethnicity and deprivation.[19 20] The fully adjusted late-stage model did not adjust for Route to Diagnosis, due to the relationship between certain routes and ethnic groups resulting in potential overadjustment and therefore increasing potential for bias. A sensitivity analysis of the late-stage model further adjusted for Route to Diagnosis is included in the online supplemental appendix.

ORs and 95% CIs were assessed in comparison to the baseline white British group. Likelihood ratio (LR) tests were used to assess the significance of ethnicity in the models.

Analyses were carried out in Stata V.15 (StataCorp, College Station, Texas, USA).

RESULTS

After exclusion criteria were applied, there were 786 596 diagnoses of the six cancer sites of interest between 2012 and 2016, 743 659 (94.5%) of which occurred in patients with known ethnicity. The distribution of detailed ethnic groups is presented in online supplemental table 2. Patients with known ethnicity had characteristics which were generally similar to those whose ethnicity was not recorded (online supplemental table 3). Those with a comorbidity index of 0, and those in the least deprived quintiles were more likely to have no ethnicity recorded compared with the other comorbidity groups and more deprived groups. Of the 786 596 cohort,634 712 (81%) had comorbidity index of 0 and 97.2% (616 963) of these tumours had a link to HES data.

A total of 696 875 tumours occurred in patients who were members of the five ethnic groups included in this analysis—the final study cohort (93.7% of those with known ethnicity). A breakdown of demographic features of this cohort is shown in online supplemental table 4.

Unknown cancer stage at diagnosis

LR tests were performed to determine which factors were significant in whether cancers were recorded with an unknown stage. Overall, ethnicity was found to be a significant factor in breast (p=0.0101), prostate (p=0.0023) and male colon cancer (p=0.0172), driven by one or two ethnic groups in each site (table 1). The unadjusted proportions of unstaged disease in each ethnic group were very similar for prostate (13%) and breast cancer (9–11%). Differences in the adjusted odds of unstaged cancer between the ethnic groups were therefore largely consistent with random noise, or small and clinically minor differences. For colon cancer, the unadjusted proportion of unstaged cancer ranged from 7% (African men) to 14% (Chinese men). However, there were small numbers of men diagnosed with colon cancer in these ethnic groups (284 and 161, respectively) and the CIs around the adjusted estimates were wide (95% CI 0.37 to 0.94 and 95% CI 0.91 to 2.30, respectively). All models were stratified by sex and adjusted for age, year, deprivation, comorbidity, region and Route to Diagnosis (online supplemental table 5).

Late cancer stage at diagnosis

Ethnic group was found to be a statistically significant predictor of late-stage diagnosis overall for all six cancer sites, namely among women for breast, ovary, uterine, NSCLC and colon cancer and among men for prostate cancer (table 2). The only models that did not demonstrate statistically significant variation between ethnic groups were NSCLC and colon cancer in men. These models were stratified by sex and adjusted for age, year, comorbidity, deprivation and region (online supplemental table 6).

After adjustment, Caribbean and African women were significantly more likely to be diagnosed with late-stage breast (OR 1.27 (95% CI 1.12 to 1.43); OR 1.71 (95% CI 1.51 to 1.95), respectively), ovarian (OR 1.48 (95% CI 1.02 to 2.14); OR 1.85 (95% CI 1.26 to 2.71)) and uterine cancer (OR 2.17 (95% CI 1.74 to 2.71); OR 2.19 (95% CI 1.61 to 2.97)), compared with white British women. Asian women had increased odds of late-stage breast (OR 1.12 (95% CI 1.03 to 1.22)) and ovarian cancer (OR 1.21 (95% CI 1.02 to 1.44)). Caribbean women had increased odds of late-stage NSCLC (OR 1.62 (95% CI 1.20 to 2.19)). Late-stage colon cancer was more likely among Caribbean men (OR 1.22 (95% CI 1.01 to 1.47)) and Caribbean and African women (OR 1.37 (95% CI 1.11 to 1.68); OR 1.42 (95% CI 1.07 to 1.88)). Conversely, Caribbean (OR 0.74 (95% CI 0.68 to 0.80)) and African (OR 0.79 (95% CI 0.71 to 0.88)) men with prostate cancer were less likely to be diagnosed at a late stage than white British men.

Adjustment for confounding variables results in no changes in statistical significance of ethnic group overall for any site apart from colon cancer among men (where the p value increased from 0.0016 to 0.0649 after adjustment). However, adjustment did alter the significance of results for specific ethnic groups, for example, Caribbean and African women with ovarian cancer (which became significant after adjustment).

The sensitivity analysis (online supplemental table 7) tested the late-stage model after further adjustment for Route to Diagnosis. Ethnicity was a significant predictor of late-stage disease for the same sites as in the fully adjusted model. All patterns observed in the fully adjusted model were also significant after adjustment for route, aside from the increased odds of late-stage ovarian cancer for Caribbean and Asian women, which were no longer significant. After adjustment for Route to Diagnosis, Chinese women had increased odds of late-stage colon cancer.

DISCUSSION

This retrospective analysis examined the relationship between ethnicity and stage at diagnosis in 696 875 patients in England. Findings indicate that Caribbean, African and Asian women with breast or ovarian cancer, Caribbean and African women with uterine or colon cancer, Caribbean women with NSCLC and Caribbean men with colon cancer were more likely to be diagnosed at a late stage compared with the white British cohort. Comparatively, Caribbean and African men with prostate cancer had decreased odds of late-stage cancer.

Ethnic group was only a predictor of unstaged disease in breast, prostate and male colon cancer. This provides reassurance that differential rates of unknown stage are unlikely to be introducing bias into the late-stage comparisons and supports the use of individuals with known stage data. The results do however suggest that some ethnic groups with small numbers of cancer cases may be less likely to have

Table 1 Percent of patients with unstaged disease and odds ratio before and after adjustment

			Total	White British	Caribbean	African	Chinese	Asian	p-value
Breast (female)		%	11	11	9	9	11	9	
		OR†		Ref	0.81 (0.69 to 0.95)	0.81 (0.67 to 0.97)	1.05 (0.83 to 1.34)	0.86 (0.79 to 0.95)	0.0004
		Adjusted OR		Ref	0.79 (0.67 to 0.94)	0.91 (0.76 to 1.10)	1.29 (1.00 to 1.65)	0.95 (0.86 to 1.05)	0.0101
Ovary		%	15	15	16	15	14	15	
		OR		Ref	1.05 (0.70 to 1.57)	0.96 (0.62 to 1.50)	0.88 (0.49 to 1.57)	0.96 (0.78 to 1.17)	0.9792
		Adjusted OR		Ref	0.94 (0.61 to 1.46)	1.07 (0.67 to 1.70)	1.20 (0.66 to 2.21)	1.04 (0.84 to 1.29)	0.9590
Uterine		%	10	10	11	11	10	11	
		OR		Ref	1.15 (0.85 to 1.56)	1.20 (0.80 to 1.81)	1.02 (0.53 to 1.96)	1.15 (0.95 to 1.40)	0.4935
		Adjusted OR		Ref	0.82 (0.58 to 1.15)	1.23 (0.79 to 1.92)	1.24 (0.61 to 2.52)	1.14 (0.92 to 1.41)	0.3910
Prostate		%	13	13	13	13	13	13	
		OR		Ref	0.93 (0.84 to 1.04)	1.00 (0.87 to 1.15)	0.99 (0.69 to 1.42)	1.00 (0.89 to 1.12)	0.7900
		Adjusted OR		Ref	0.82 (0.73 to 0.92)	1.10 (0.95 to 1.27)	0.93 (0.64 to 1.37)	0.89 (0.79 to 1.01)	0.0023
NSCLC	Men	%	8	8	9	7	9	8	
		OR		Ref	1.08 (0.83 to 1.41)	0.87 (0.53 to 1.42)	1.14 (0.70 to 1.85)	0.90 (0.73 to 1.11)	0.7477
		Adjusted OR		Ref	0.96 (0.72 to 1.26)	0.94 (0.57 to 1.56)	1.16 (0.70 to 1.91)	0.78 (0.62 to 0.97)	0.2081
	Women	%	9	9	8	7	7	10	
		OR		Ref	0.88 (0.58 to 1.35)	0.73 (0.37 to 1.43)	0.77 (0.41 to 1.41)	1.18 (0.88 to 1.57)	0.5291
		Adjusted OR		Ref	0.68 (0.44 to 1.05)	0.82 (0.41 to 1.63)	0.74 (0.39 to 1.37)	1.00 (0.75 to 1.35)	0.3197
Colon	Men	%	12	12	11	7	14	11	
		OR		Ref	0.93 (0.71 to 1.22)	0.61 (0.39 to 0.95)	1.27 (0.82 to 1.98)	0.95 (0.76 to 1.18)	0.1710
		Adjusted OR		Ref	0.76 (0.57 to 1.01)	0.59 (0.37 to 0.94)	1.45 (0.91 to 2.30)	0.92 (0.73 to 1.16)	0.0172
	Women	%	13	14	12	10	13	12	
		OR		Ref	0.90 (0.68 to 1.19)	0.69 (0.45 to 1.05)	0.94 (0.57 to 1.53)	0.87 (0.69 to 1.09)	0.2820
		Adjusted OR		Ref	0.96 (0.72 to 1.29)	0.98 (0.63 to 1.52)	1.17 (0.70 to 1.97)	0.98 (0.77 to 1.25)	0.9766

Adjustments made for 10-year age band, comorbidity index, deprivation quintile, year of diagnosis and region.
*OR (95% CI).
CI, confidence interval; NSCLC, non-small cell lung cancer; OR, odds ratio.

Table 2 Percent of patients with late-stage disease (stages 3 and 4) and odds ratio before and after adjustment

			Total	White British	Caribbean	African	Chinese	Asian	p-value
Breast (female)		%	15	15	21	27	13	17	
		OR†		Ref	1.49 (1.33 to 1.67)	2.05 (1.81 to 2.31)	0.87 (0.68 to 1.10)	1.17 (1.08 to 1.26)	<0.0001
		Adjusted OR		Ref	1.27 (1.12 to 1.43)	1.71 (1.51 to 1.95)	0.84 (0.66 to 1.07)	1.12 (1.03 to 1.22)	<0.0001
Ovary		%	61	61	65	62	46	55	
		OR		Ref	1.18 (0.84 to 1.65)	1.01 (0.71 to 1.44)	0.53 (0.34 to 0.82)	0.77 (0.66 to 0.90)	0.0007
		Adjusted OR		Ref	1.48 (1.02 to 2.14)	1.85 (1.26 to 2.71)	0.99 (0.62 to 1.58)	1.21 (1.02 to 1.44)	0.0013
Uterine		%	18	18	36	33	14	18	
		OR		Ref	2.52 (2.05 to 3.11)	2.20 (1.64 to 2.94)	0.73 (0.41 to 1.32)	1.00 (0.85 to 1.18)	<0.0001
		Adjusted OR		Ref	2.17 (1.74 to 2.71)	2.19 (1.61 to 2.97)	0.83 (0.46 to 1.51)	1.04 (0.87 to 1.23)	<0.0001
Prostate		%	43	43	36	34	38	42	
		OR		Ref	0.73 (0.67 to 0.79)	0.68 (0.61 to 0.75)	0.82 (0.63 to 1.07)	0.93 (0.86 to 1.02)	<0.0001
		Adjusted OR		Ref	0.74 (0.68 to 0.80)	0.79 (0.71 to 0.88)	0.85 (0.65 to 1.12)	0.97 (0.89 to 1.06)	<0.0001
NSCLC	Men	%	74	74	76	78	75	73	
		OR		Ref	1.14 (0.94 to 1.37)	1.26 (0.91 to 1.74)	1.04 (0.74 to 1.46)	0.98 (0.86 to 1.12)	0.4288
		Adjusted OR		Ref	1.04 (0.86 to 1.26)	1.10 (0.79 to 1.53)	0.90 (0.64 to 1.27)	0.98 (0.85 to 1.12)	0.9113
	Women	%	70	70	79	71	71	72	
		OR		Ref	1.69 (1.26 to 2.27)	1.05 (0.72 to 1.53)	1.09 (0.76 to 1.55)	1.12 (0.92 to 1.37)	0.0085
		Adjusted OR		Ref	1.62 (1.20 to 2.19)	0.90 (0.61 to 1.31)	0.96 (0.67 to 1.37)	1.09 (0.89 to 1.34)	0.0174
Colon	Men	%	55	55	62	64	52	56	
		OR		Ref	1.32 (1.09 to 1.58)	1.45 (1.13 to 1.87)	0.89 (0.64 to 1.25)	1.04 (0.90 to 1.21)	0.0016
		Adjusted OR		Ref	1.22 (1.01 to 1.47)	1.26 (0.97 to 1.62)	0.84 (0.60 to 1.18)	0.95 (0.81 to 1.10)	0.0649
	Women	%	56	55	64	66	65	55	
		OR		Ref	1.46 (1.19 to 1.79)	1.59 (1.20 to 2.10)	1.50 (1.03 to 2.17)	0.98 (0.84 to 1.16)	<0.0001
		Adjusted OR		Ref	1.37 (1.11 to 1.68)	1.42 (1.07 to 1.88)	1.40 (0.96 to 2.03)	0.92 (0.78 to 1.09)	0.0009

Adjustments made for 10-year age band, comorbidity index, deprivation quintile, year of diagnosis and region.
*OR (95% CI).
CI, confidence interval; NSCLC, non-small cell lung cancer; OR, odds ratio.

a stage recorded. Chinese women with breast cancer and men with colon cancer were the smallest groups with these cancers and had the highest ORs of not having a stage at diagnosis recorded. However, these were either of borderline significance (p=0.048 breast cancer) or not statistically significant (p=0.11 colon cancer). If stage is missing at random, and the odds of being unstaged is unrelated to the true stage (eg, being late stage), estimates of late-stage disease will be unbiased.[10] However, if tumours are unstaged due to being a particular stage (eg, if the negative impact of biopsy is thought to outweigh the benefit of having precise stage information where the treatment options are clear), stage data will be missing not at random (MNAR). The comparisons of the odds of late-stage diagnosis in the different ethnic groups could be biased if stage data are MNAR and vary by ethnic group.

Of those with a known stage, ethnicity was a significant predictor of late stage for women with breast, ovary, uterine, NSCLC and colon cancer, and for men with prostate cancer. Caribbean and African women had over double the odds of late-stage uterine cancer than white British women, which is in line with studies from the USA.[21 22] Caribbean, African and Asian women also had higher odds of being diagnosed with late-stage breast and ovarian cancer. In contrast to this, Caribbean and African men had over 20% decreased odds of late-stage prostate cancer compared with white men. Prostate cancer incidence is higher among black men than white men in England[23–25] and research has shown that black men have a higher likelihood of receiving prostate-specific antigen testing, independent of family history and education status,[26 27] which increases the proportion of early-stage prostate cancers.[28] Greater awareness of the risk of prostate cancer among both GPs and black men supported by cancer charity campaigns is likely to be driving this increase. However, earlier research using diagnoses from 1998 to 2003 did not find differences in the proportions of late-stage prostate cancer between ethnic groups, suggesting that more recent developments may have had an impact.[29] African men have increased odds of unstaged prostate cancer; however, it is not clear why there is a difference between Caribbean and African men in this regard and therefore caution is required when interpreting late-stage prostate cancer results for African men.

Adjusting for confounding variables strengthened and attenuated relationships between ethnicity and late-stage disease depending on the cancer site, and in some cases affected the statistical significance of these relationships. Differing age distributions in the ethnic groups had the largest effect on the adjusted estimates relative to other confounding variables. Overall, the sensitivity analysis suggests that the ethnic group differences in late-stage disease are not fully explained by Route to Diagnosis. For the sites with a screening programme (breast and colon), adjustment for differences in route between ethnic groups shows that non-white patients diagnosed via the same route still have increased odds of late-stage disease.

The observed differences by ethnic group may be explained by a range and interaction of many factors including both patient-level and system-level factors.[30 31] The finding of later stage diagnosis among ethnic minorities may be linked to poorer symptom awareness. Individuals in ethnic minority groups have also been found to recognise fewer cancer symptoms than white individuals, with symptom recognition by ethnicity reducing in the following order: white British, Chinese, black Caribbean and Indian, Pakistani, black African and Bangladeshi.[7 8] A longer time interval between symptoms and presentation to a GP ('patient interval') could result in cancer progression and therefore later stage disease. Longer delays have been observed among people with lower symptom knowledge and more barriers to presentation.[5] Minority ethnic groups have previously been shown to have more barriers to presentation than white individuals, including embarrassment, worry and practical barriers such as difficulty in making appointments, cultural barriers and different attitudes to healthcare system usage, including negative attitudes towards the GP.[7 8 32] Cancer fatalism, a factor causing possible delay to diagnosis, has also been found to be higher in ethnic minority women compared with white British women, with African and Indian women more fearful of cancer.[33]

A study of patients with cancer in England found that mixed, black and Asian patients with cancer were more likely to have reported visiting a family doctor on three or more occasions before being referred to hospital.[30] These results were unaffected by adjustment for type of cancer, age, sex and socioeconomic deprivation. Delays within the primary care system may therefore be contributing to a later diagnosis for some ethnic groups. It is not known whether there are differences in delays between initiating symptom investigations, or in tests being performed and results being made available. A report on cancer and black and minority ethnic communities in England found a lack of cancer education and knowledge of available support services among these groups, calling for cultural competence training of healthcare providers and embedding of diversity into the design of cancer services.[34]

In terms of ethnicity and stage, breast cancer has been most studied. Caribbean, African and Asian women with breast cancer were more likely to be diagnosed at late stage compared with white British women, with a similar association observed in recent US publications.[35 36] A study from England found that women from ethnic minorities had greater odds of less favourable tumour characteristics of breast cancer, including stage, histological grade, oestrogen receptor status and HER2 status.[37] Black women have been found to do less breast checking in comparison to white women,[5] and barriers such as stigma are related to the woman's country of birth, age and how long she has lived in the UK.[6 36] In one study, black women reported limited awareness of breast cancer, services and treatment options.[38] Additionally, biological variation in cancers between ethnic groups could potentially contribute to differing stage distributions.[39] There is a possibility that certain ethnic groups or exposures are linked with particular subtypes or morphologies with more aggressive disease, resulting in

later stage cancer. An example of this is hormone receptor negative status where triple negative breast cancer, which grows quickly and is associated with poor survival, is more common in black women.[39 40]

Strengths and limitations

The key strength of this analysis is the use of national-level data, using granular ethnicity where possible. This has been made possible through improvements in data completeness and quality and allows for investigation of distinct ethnic groups which are commonly grouped in other studies. Adjustment was also made for other confounders including comorbidities. Comorbidity could feasibly result in a patient already being known to and engaged with the healthcare system, and therefore a potentially incidental cancer finding with earlier presentation, or alternatively in less diagnostic testing, or a longer process of cancer diagnosis due to ill health and other disease symptoms.

There are some limitations to the analysis that need to be considered. There was no control for tumour subtype or morphology, which may be associated with ethnic group, and potentially related to staging. Deprivation level, which has been found to be linked to cancer symptom awareness and barriers to presentation, was adjusted for in this analysis.[41 42] However, an area-based measure of income deprivation was used as opposed to the individual's socioeconomic status. Certain information was unavailable for the analysis, for example, data on education, religion, country of birth and language spoken at home. Symptom knowledge has been previously shown to be lower among those who do not use English at home,[8] while an understanding of the healthcare system can be better in those born in the UK.[6] Another limitation of this analysis is the generic split of early and late stage for all cancer sites, when it could be more appropriate to define late stage for specific sites, depending on differences in patient outcomes by stage. Since ethnicity and comorbidity data are obtained from HES, individuals with no HES record are more likely to have unknown ethnicity and missing comorbidity. This explains the observation of individuals with Charlson index of 0 being less likely to have ethnic group recorded. The requirement of hospital admission in defining Charlson score means that diseases managed in primary care are not accounted for. Additionally, this analysis is based on diagnoses in residents in England where an NHS is in place with free healthcare at the point of use, and therefore outcomes may not be representative of international patterns. The sensitivity analysis adjusted for Route to Diagnosis. One limitation in this sensitivity analysis is that screening eligibility was not adjusted for; in screening-eligible age groups, there may be different patterns of staging due to the screening programmes.

CONCLUSION

This study found variations in late-stage cancer by ethnic group. Certain cancer sites such as breast have a clear symptom signature,[43 44] meaning ethnicity-related disparities in symptom awareness could greatly impact stage distributions. Other sites in this analysis, including colon, lung and ovarian cancers, have been found to have a broader symptom signature.[43] Given the differences in stage at diagnosis between ethnic groups, potentially resulting from delays to diagnosis, this analysis supports the consideration of targeted interventions to raise the awareness of early diagnosis and its benefits among non-white ethnic groups. Ensuring healthcare systems are not creating barriers that prevent people from particular ethnic groups being seen and referred appropriately without delays is an important area to focus on in the future. Others have suggested the practical delivery of these interventions through community networks, local media and faith settings, as well as by culturally competent patient navigators or advocates who can increase understanding of cancer.[34 38] Removal of barriers to appropriate referrals which will benefit early diagnosis could also help to reduce the observed inequalities.

The importance of the observed relationships between ethnicity and late-stage diagnosis is also highlighted by differential cancer survival between different ethnic groups. For example, black women in England have been found to have poorer breast cancer survival than white women.[2 45] A recent analysis of survival found low symptom awareness was linked to lower cancer survival, and more barriers were linked with lower breast cancer survival.[7] Understanding these mechanisms of delayed diagnosis, and therefore late-stage disease, is important to reduce ethnic group survival disparities.

Acknowledgements This work uses data that has been provided by patients and collected by the National Health Service (NHS) as part of their care and support. The data is collated, maintained and quality assured by the National Cancer Registration and Analysis Service, which is part of The Health and Social Care Information Centre, trading as NHS Digital. RHJ acknowledges the financial support of the National Institute for Health Research Nottingham Biomedical Research Centre. The authors acknowledge and thank colleagues who helped with the inception and development of this work, namely Clare Pearson, Bella Carneiro and Nick Ormiston-Smith.

Contributors All authors participated in the study conception and design, and/or the analysis and interpretation of the results. Conception and design: AF and RHJ. Analysis and interpretation: AF, BW, JS and RHJ. Writing manuscript: AF, JS, BW, RHJ and DN. Supervision: RHJ and JS. All authors critically reviewed drafts of the manuscript, read and approved the final manuscript. JS is the guarantor and accepts full responsibility for the work, had access to the data and controlled the decision to publish.

Funding The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests None declared.

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Patient consent for publication Not applicable.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data may be obtained from a third party and are not publicly available. All data analysed during this study are available from NHS Digital's National Cancer Registration and Analysis Service, however, restrictions apply to the availability of these data.

ORCID iDs
Becky White http://orcid.org/0000-0002-0643-7890
Diana Nagarwalla http://orcid.org/0000-0001-8716-5446
Ruth H Jack http://orcid.org/0000-0003-4009-020X

REFERENCES

1 McPhail S, Johnson S, Greenberg D, *et al*. Stage at diagnosis and early mortality from cancer in England. *Br J Cancer* 2015;112 Suppl 1:S108–15.
2 Jack RH, Davies EA, Møller H. Breast cancer incidence, stage, treatment and survival in ethnic groups in South East England. *Br J Cancer* 2009;100:545–50.
3 NCIN. *Ethnicity and stage at diagnosis*, 2016.
4 Olesen F, Hansen RP, Vedsted P. Delay in diagnosis: the experience in Denmark. *Br J Cancer* 2009;101 Suppl 2:S5–8.
5 Forbes LJL, Atkins L, Thurnham A, *et al*. Breast cancer awareness and barriers to symptomatic presentation among women from different ethnic groups in East London. *Br J Cancer* 2011;105:1474–9.
6 Jones CEL, Maben J, Lucas G, *et al*. Barriers to early diagnosis of symptomatic breast cancer: a qualitative study of black African, black Caribbean and white British women living in the UK. *BMJ Open* 2015;5:e006944.
7 Niksic M, Rachet B, Warburton FG, *et al*. Ethnic differences in cancer symptom awareness and barriers to seeking medical help in England. *Br J Cancer* 2016;115:136–44.
8 Waller J, Robb K, Stubbings S, *et al*. Awareness of cancer symptoms and anticipated help seeking among ethnic minority groups in England. *Br J Cancer* 2009;101 Suppl 2:S24–30.
9 Williams ED, Whitaker KL, Piano M, *et al*. Ethnic differences in barriers to symptomatic presentation in primary care: a survey of women in England. *Psychooncology* 2019;28:2336–43.
10 Di Girolamo C, Walters S, Benitez Majano S, *et al*. Characteristics of patients with missing information on stage: a population-based study of patients diagnosed with colon, lung or breast cancer in England in 2013. *BMC Cancer* 2018;18:492.
11 Lyratzopoulos G, Abel GA, Brown CH, *et al*. Socio-Demographic inequalities in stage of cancer diagnosis: evidence from patients with female breast, lung, colon, rectal, prostate, renal, bladder, melanoma, ovarian and endometrial cancer. *Ann Oncol* 2013;24:843–50.
12 Henson KE, Elliss-Brookes L, Coupland VH, *et al*. Data resource profile: National cancer registration dataset in England. *Int J Epidemiol* 2020;49:16–16h.
13 Marlow LAV, Robb KA, Simon AE, *et al*. Awareness of cancer risk factors among ethnic minority groups in England. *Public Health* 2012;126: . :702–9.
14 Saunders CL, Abel GA, El Turabi A, *et al*. Accuracy of routinely recorded ethnic group information compared with self-reported ethnicity: evidence from the English cancer patient experience survey. *BMJ Open* 2013;3:e002882.
15 Ministry of Housing Communities & Local Government,. *English indices of deprivation*, 2015.
16 Elliss-Brookes L, McPhail S, Ives A, *et al*. Routes to diagnosis for cancer - determining the patient journey using multiple routine data sets. *Br J Cancer* 2012;107:1220–6.
17 Quan H, Li B, Couris CM, *et al*. Updating and validating the Charlson comorbidity index and score for risk adjustment in hospital discharge Abstracts using data from 6 countries. *Am J Epidemiol* 2011;173:676–82.
18 Maringe C, Fowler H, Rachet B, *et al*. Reproducibility, reliability and validity of population-based administrative health data for the assessment of cancer non-related comorbidities. *PLoS One* 2017;12:e0172814.
19 Jack RH, Møller H, Robson T, *et al*. Breast cancer screening uptake among women from different ethnic groups in London: a population-based cohort study. *BMJ Open* 2014;4:e005586.
20 NCIN. *Routes to diagnosis by stage, 2012-2013*, 2016.
21 Halpern MT, Ward EM, Pavluck AL, *et al*. Association of insurance status and ethnicity with cancer stage at diagnosis for 12 cancer sites: a retrospective analysis. *Lancet Oncol* 2008;9:222–31.
22 Madison T, Schottenfeld D, James SA, *et al*. Endometrial cancer: socioeconomic status and racial/ethnic differences in stage at diagnosis, treatment, and survival. *Am J Public Health* 2004;94:2104–11.
23 Ben-Shlomo Y, Evans S, Ibrahim F, *et al*. The risk of prostate cancer amongst black men in the United Kingdom: the process cohort study. *Eur Urol* 2008;53:99–105.
24 Lloyd T, Hounsome L, Mehay A, *et al*. Lifetime risk of being diagnosed with, or dying from, prostate cancer by major ethnic group in England 2008-2010. *BMC Med* 2015;13:171.
25 Perez-Cornago A, Key TJ, Allen NE, *et al*. Prospective investigation of risk factors for prostate cancer in the UK Biobank cohort study. *Br J Cancer* 2017;117:1562–71.
26 Littlejohns TJ, Travis RC, Key TJ, *et al*. Lifestyle factors and prostate-specific antigen (PSA) testing in UK Biobank: implications for epidemiological research. *Cancer Epidemiol* 2016;45:40–6.
27 Nderitu P, Van Hemelrijck M, Ashworth M, *et al*. Prostate-Specific antigen testing in inner London general practices: are those at higher risk most likely to get tested? *BMJ Open* 2016;6:e011356.
28 Moore AL, Dimitropoulou P, Lane A, *et al*. Population-Based prostate-specific antigen testing in the UK leads to a stage migration of prostate cancer. *BJU Int* 2009;104:1592–8.
29 Jack RH, Davies EA, Møller H. Prostate cancer incidence, stage at diagnosis, treatment and survival in ethnic groups in south-east England. *BJU Int* 2010;105:1226–30.
30 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;13:353–65.
31 Lyratzopoulos G, Vedsted P, Singh H. Understanding missed opportunities for more timely diagnosis of cancer in symptomatic patients after presentation. *Br J Cancer* 2015;112 Suppl 1:S84–91.
32 Marlow LAV, McGregor LM, Nazroo JY, *et al*. Facilitators and barriers to help-seeking for breast and cervical cancer symptoms: a qualitative study with an ethnically diverse sample in London. *Psychooncology* 2014;23:749–57.
33 Vrinten C, Wardle J, Marlow LA. Cancer fear and fatalism among ethnic minority women in the United Kingdom. *Br J Cancer* 2016;114:597–604.
34 Fazil Q. *Cancer and black and minority ethnic communities, in better health Briefing 47*. London: Race Equality Foundation, 2018.
35 Iqbal J, Ginsburg O, Rochon PA, *et al*. Differences in breast cancer stage at diagnosis and cancer-specific survival by race and ethnicity in the United States. *JAMA* 2015;313:165.
36 Jones CE, Maben J, Jack RH, *et al*. A systematic review of barriers to early presentation and diagnosis with breast cancer among black women. *BMJ Open* 2014;4:e004076.
37 Gathani T, Reeves G, Broggio J, *et al*. Ethnicity and the tumour characteristics of invasive breast cancer in over 116,500 women in England. *Br J Cancer* 2021;125:611–7.
38 Bamidele O, Ali N, Papadopoulos C, *et al*. Exploring factors contributing to low uptake of the NHS breast cancer screening programme among black African women in the UK. *Divers Equal Health Care* 2017;14.
39 Copson E, Maishman T, Gerty S, *et al*. Ethnicity and outcome of young breast cancer patients in the United Kingdom: the POSH study. *Br J Cancer* 2014;110:230–41.
40 Jack RH, Davies EA, Renshaw C, *et al*. Differences in breast cancer hormone receptor status in ethnic groups: a London population. *Eur J Cancer* 2013;49:696–702.
41 Moffat J, Hinchliffe R, Ironmonger L, *et al*. Identifying anticipated barriers to help-seeking to promote earlier diagnosis of cancer in Great Britain. *Public Health* 2016;141:120–5.
42 Niksic M, Rachet B, Warburton FG, *et al*. Cancer symptom awareness and barriers to symptomatic presentation in England--are we clear on cancer? *Br J Cancer* 2015;113:533–42.
43 Koo MM, Hamilton W, Walter FM, *et al*. Symptom signatures and diagnostic timeliness in cancer patients: a review of current evidence. *Neoplasia* 2018;20:165–74.
44 Lyratzopoulos G, Abel GA, Barbiere JM, *et al*. Variation in advanced stage at diagnosis of lung and female breast cancer in an English region 2006-2009. *Br J Cancer* 2012;106:1068–75.
45 Møller H, Henson K, Lüchtenborg M, *et al*. Short-Term breast cancer survival in relation to ethnicity, stage, grade and receptor status: national cohort study in England. *Br J Cancer* 2016;115:1408–15.