The Royal Society of Medicine

Journal of the Royal Society of Medicine; 2016, Vol. 109(8) 303–309
DOI: 10.1177/0141076816651037

Risk of individual malignant neoplasms in patients with sickle cell disease: English national record linkage study

Olena O Seminog, Oyindamola I Ogunlaja, David Yeates and Michael J Goldacre
Unit of Health-Care Epidemiology, Nuffield Department of Population Health, University of Oxford, Oxford OX3 7LF, UK
Corresponding author: Michael J Goldacre. Email: michael.goldacre@dph.ox.ac.uk

Abstract

Objective: Case reports suggest that there may be an increased risk of some cancers associated with sickle cell disease. However, population-based studies are scarce and there is no comprehensive enumeration of the risks across the whole range of site-specific cancers. Our aim was to provide this.

Design: We used an English national dataset of linked statistical records of hospital admissions and deaths from 1999 to 2011 to undertake a retrospective cohort study.

Setting: England.

Participants: Records of all hospital admissions in England with SCD or with conditions included in the control cohort.

Main outcome measures: Rate ratios were calculated comparing rates of cancer in a sickle cell disease cohort and a control cohort, confining the analyses to people whose ethnicity was recorded as Black.

Results: Comparing the sickle cell disease cohort with the cohort without sickle cell disease, the rate ratio for all cancers combined was 2.1 (95% confidence interval 1.7–2.5). There were significantly high rate ratios for haematological malignancies, including Hodgkin's lymphoma (rate ratio 3.7, 1.5–8.4), non-Hodgkin's lymphoma (2.6, 1.3–4.8), multiple myeloma (5.5, 2.8–10.1), lymphoid leukaemia (3.3, 1.3–8.0) and myeloid leukaemia (10.0, 4.6–21.5). Four solid tumours showed elevated rate ratios: colon cancer (2.8, 1.2–5.5), non-melanoma skin cancer (4.4, 1.3–12.2), kidney cancer (5.4, 2.3–11.5) and thyroid cancer (5.1, 1.3–15.4).

Conclusions: The risk of some malignancies may be raised in patients with sickle cell disease. However, this study was based on administrative data without the scope to validate these against patients' full clinical records. Our findings need confirmation or refutation. If confirmed, work to elucidate, at the genetic and molecular level, why people with sickle cell disease have elevated risks of individual cancers might make contributions to the fundamental understanding of carcinogenesis.

Keywords

Sickle cell disease, haematological malignancies, cancer risk, record linkage studies, epidemiology

Introduction

Patients with sickle cell disease (SCD) have a gene that codes for an abnormal form of haemoglobin – Haemoglobin S.[1] The disease is usually seen in people from Sub-Saharan Africa, India, parts of the Middle East and Mediterranean, and their descendants. It is now the most common genetic disorder in England.[2] Case reports and case series have reported an increased risk of haematological malignancies in patients with SCD.[3–9] However, they are limited in the range of cancers studied, and they do not provide comparisons of cancer risk with the general population. We aimed to quantify the scale of risk of cancers overall, and individually, in SCD patients, by comparing them with a population without SCD, using a national dataset of linked hospital and mortality records in England. This study is part of a broader programme of work on associations between different diseases,[10] including studies of the possibility of altered risk of cancers in people with some genetic disorders.[11–13]

Methods

Hospital Episode Statistics (HES) is the routinely collected all-England administrative dataset that contains a statistical record of every hospital day case and inpatient admission in the NHS in England (population about 53 million). Records of successive admissions for each individual and data from the death registration record, if death occurred, were linked together into a cumulative record of care for each individual from 1999 (when linkable data first became available nationally) to December 2011. The linked file, which was intended to be multi-purpose and support a wide range of studies, was built by the Oxford record linkage team using anonymised unique personal identifiers. In brief, linkage was based on the encrypted values of the HES identification number (a national number that is unique to each person hospitalised for NHS care), the encrypted NHS number (unique to each individual registered for NHS care) and several other encrypted personal identifiers. The national HES records were supplied by the Health and Social Care Information Centre.

An 'exposure' cohort was constructed by identifying the first admission for each individual hospitalised for SCD. The latter was defined using the International Classification of Diseases (ICD) 10th revision codes D57.0-D57.2 and D57.8. Cases of sickle cell trait were not included. A control cohort was constructed to include people hospitalised with a variety of mainly minor medical and surgical conditions and injuries (see footnotes to Table 1). The cohorts were 'followed', as retrospective cohorts, to identify cancers as 'outcome' events. The start date for follow-up of each individual was the date of first hospital record for SCD or condition in the control cohort. Multiple cohort studies were undertaken, one for each individual cancer, such that the analysis for each individual cancer was undertaken independently of the analysis of each other cancer. For each specific cancer, people were included in the SCD or control cohort if they did not have a record of the specific cancer before or at the same time as the first admission for SCD or control condition. Anyone with both SCD and a control condition was counted in the SCD cohort and removed from the control cohort.

The rate ratio for each individual cancer was calculated as follows. The risk of each specific malignant neoplasm was calculated based on person-days at risk. 'Days at risk' was defined as the time period in days between the date of first admission with SCD or control condition and the date of subsequent admission with each cancer, death (regardless of cause) or end of data collection at 31 December 2011 (whichever occurred first). To compare the SCD cohort with the control cohort, rates of each type of cancer were calculated within age strata in five-year groups (by age at entry to the cohort) and, within each age stratum, by further stratification by sex, year of first admission (in single years), quintile of index of multiple deprivation (a standard UK measure of socio-economic status) and region of residence (9 in England). All the main analysis was confined to patients in the SCD and control cohorts with a HES record that specified that they were Black. We did this knowing that HES data on ethnicity is incomplete;[14] but we wanted, as best as we could, to distinguish associations attributable to SCD from those that may be attributable to ethnicity. Black patients (the term in HES) were selected as those with HES ethnicity codes for 'Caribbean (Black or Black British)', 'African (Black or Black British)' or 'Any other Black background'. Analyses were confined to people aged under 65 years. We age-standardised all rates using indirect standardisation, where the standard population was the combined SCD and

Table 1. Age and sex distribution of people admitted to hospital with sickle cell disease[a]: number of people in each age stratum, percentage male, stratum ratio[b] of controls[c] per case in each age group. Data confined to people with ethnic group recorded as Black.

Age (years)	Number (% male)	% of total	Stratum ratio[b]
0–4	597 (57)	7.9	12.5
5–9	710 (55)	9.5	15.5
10–14	642 (52)	8.5	13.8
15–19	804 (53)	10.7	10.1
20–24	892 (57)	11.9	10.6
25–29	830 (62)	11.0	12.2
30–34	745 (59)	9.9	14.0
35–39	688 (55)	9.2	17.7
40–44	569 (45)	7.6	21.4
45–49	422 (44)	5.1	23.8
50–54	259 (46)	3.4	28.0
55–59	188 (35)	2.5	30.4
60–64	166 (49)	2.2	36.9
Total/average	7512 (54)	100	15.8

[a]137 (96.4%) coded as D57.0 or D57.1 (Sickle cell anaemia), 5 (1.8%) as D57.2 (Double heterozygous sickling disorders) and 8 (5.6%) as D57.8 (Other sickle cell disorders).

[b]The number of people in the control cohort per person with SCD in each age stratum. Note that, in analysis, we included all people eligible to be in the control cohort and analysed the data within each age stratum comparing people with SCD and controls.

[c]Conditions used in control cohort, with Office of Population, Censuses and Surveys (OPCS) code edition 4 for operations and ICD10 code for diagnosis (with equivalent codes used for other coding editions): appendectomy (OPCS4 H01–H03), adenoidectomy (E20), dilation and curettage (Q10–Q11), hip replacement (W37–W39), knee replacement (W40-W42), squint (ICD10 H49–H51),cataract (H25), otitis (H60–H67), upper respiratory tract infections (J00–J06), varicose veins (I83), haemorrhoids (I84), deflected septum, nasal polyp (J33 + J34.2), impacted tooth and other disorders of teeth (K00–K03), inguinal hernia (K40), head injury (S06), ingrowing nail, toenail and other diseases of nail (L60), contraceptive management (Z30), internal derangement of knee (M23), bunion (727.1), dislocations, sprains and strains (S03, S13, S23, S33, S43, S53, S63, S73, S83, S93), selected limb fractures (S42, S52, S62, S82, S92), superficial injury and contusion (S00, S10, S20, S30, S40, S50, S60, S70, S80, S90). In analyses of colorectal cancers, we excluded appendectomy, haemorrhoids and inguinal hernia from the control cohort. From the analysis of IPD, we excluded upper respiratory tract infections.

Table 2. Risk of haematological and non-haematological malignancies in Black patients with sickle cell disease by time interval, number of observed (N) cases of cancer, rate ratio (RR) comparing the SCD cohort with the control cohort, and 95% confidence intervals (95% CI) for the rate ratio.

	All years			All years under observation excluding first year		
Cancers (ICD codes)	*N*	RR (95% CI)	*p*	*N*	RR (95% CI)	*p*
All cancers combined	142	2.07 (1.73–2.46)	<0.001	98	1.85 (1.49–2.27)	<0.001
Colon cancer (C18)	9	2.75(1.21–5.49)	0.006	<5[a]		
Lung/bronchus cancer (C34)	7	2.08 (0.81–4.46)	0.103	5	1.78 (0.56–4.35)	0.333
Non–melanoma skin cancer(C44)	5	4.42 (1.29–12.24)	0.005	5	5.37 (1.50–16.05)	0.002
Breast cancer (women) (C50)	7	0.58 (0.23–1.22)	0.001	5	0.51 (0.16–1.20)	<0.001
Prostate cancer (C61)	9	1.21 (0.55–2.33)	0.704			
Kidney cancer (C64)	9	5.41 (2.29–11.47)	<0.001	7	6.30 (2.33–14.69)	<0.001
Thyroid cancer (C73)	<5	5.08 (1.25–15.41)	0.006	<5[a]		
Hodgkin's lymphoma (C81)	8	3.71 (1.46–8.36)	0.001	6	4.32 (1.43–11.08)	0.002
Non–Hodgkin's lymphoma (C82–C85)	13	2.63 (1.34–4.76)	0.002	9	2.37 (1.03–4.86)	0.027
Multiple myeloma (C90)	14	5.46 (2.78–10.05)	<0.001	9	4.07 (1.74–8.49)	<0.001
Lymphoid leukaemia (C91)	8	3.33 (1.26–7.97)	0.006	6	2.72 (0.86–7.41)	0.067
Myeloid leukaemia (C92)	14	9.99 (4.56–21.53)	<0.001	11	8.6 (3.49–20.8)	<0.001
Chronic myeloid leukaemia (C92.0)	<5	11.35 (2.25–52.74)	<0.001	<5	7.54 (1.17–38.77)	0.010
Acute myeloid leukaemia (C92.1)	8	11.05 (3.86–30.17)	<0.001	6	10.69 (2.97–37.15)	<0.001

[a]Data not shown for cancers with fewer than five observed cases that were not statistically significant.

control cohorts. Stratum-specific cancer rates in the standard population were calculated. They were then applied to each corresponding stratum in the SCD cohort and, separately, to each stratum in the control cohort. A rate ratio of the standardised rates of cancer in the SCD group, compared with the standardised rates in the control group, was calculated using the formula $(O^{scd}/E^{scd})/(O^{contr}/E^{contr})$, where O and E are the observed and expected numbers in, respectively, the SCD and control cohorts. The rate ratios and their 95% confidence intervals and *p* values were calculated using standard methods.[15]

We present findings for all cases of individual cancers in the SCD or control cohort, and for all cases excluding those that occurred within a year of the first hospitalisation for SCD or control event. We did the latter to guard against surveillance bias, i.e. detection of cancer because the patient was under care for SCD or the control condition. Data on all cancers with five or more cases, and those with fewer than five but with statistically significant risks, are shown individually (Table 2).

Analysis was undertaken using Statistical Analysis Software (SAS, release 9.2, SAS Institute Inc., Cary, NC, USA).

Results

There were 7512 people with SCD and 118,821people in the control cohort. Table 1 shows the age distribution of people in the SCD cohort, the percentage in each age stratum who were men, and the number of people in the control cohort per person in the SCD cohort. As described in the Methods section, all analyses were initially done within age strata, e.g. for children aged under 5 years, 597 children in the SCD cohort were compared with 74,625 children of the same age (597 × 12.5) in the control cohort (Table 1).

Overall, the rate ratio (RR), comparing the SCD cohort with the control cohort, was 2.1 (95% confidence interval 1.7–2.5; Table 2). There were significantly high RRs for haematological malignancies, including Hodgkin's lymphoma (RR 3.7, 1.5–8.4), non-Hodgkin's lymphoma (2.6, 1.3–4.8), multiple myeloma (5.5, 2.8–10.1), lymphoid leukaemia (3.3, 1.3–8.0) and myeloid leukaemia (10.0, 4.6–21.5). There was only one case coded as Burkitt lymphoma.

There was also an increased risk of some solid malignant tumours (Table 2), namely cancer of the colon, kidney, thyroid, prostate and non-melanoma skin cancer. Cancers of other sites that were included in the analysis, but did not yield five or more cases, and without significant findings, comprised cancer of the oesophagus, stomach, rectum, bladder, liver, pancreas, bone and cartilage, malignant melanoma, chronic and acute lymphoid leukaemia analysed separately, ovary, uterus and testis.

Exclusion of first-year cases

In subgroup analysis, we estimated RRs excluding the first year after hospitalisation for SCD or conditions in the control group. The RR for cancers overall remained statistically significant at 1.9 (1.5–2.3). Among haematological malignancies, the risks remained elevated for all conditions studied, except for lymphoid leukaemia. Increased risks were also observed for kidney cancer and non-melanoma skin cancer (Table 2).

Face-validity test: Invasive pneumococcal disease in the SCD cohort

We used the same datasets and methods to study the risk of invasive pneumococcal disease (IPD) in the SCD cohort. We did so to be sure that the datasets and methods identify IPD as a known risk in people with SCD. The all-ages relative risk (RR) of IPD in the SCD cohort, compared with the reference cohort (in which the risk is 1), was 12.6 (95% CI 11.2–14.2). The rate ratios were 23.4 (14.0–41.5), based on 117 cases of IPD observed in children aged 0–4 years; 35.8 (24.3–54.5), based on 198 cases observed in children aged 5–14 years; 15.7 (13.0–19.0), based on 433 observed cases in people aged 15–44 years; and 5.1 (3.7–6.9), based on 55 observed cases in people aged 45–64 years.

Recording of ethnicity in HES

In the SCD cohort, there were 8662 people with a record of ethnicity of whom 7512 (87%) were Black. In the HES dataset, covering the 13-year period, the whole population of people aged under 65 years coded as SCD was 13,781, of whom 63% had a record of ethnicity and 37% did not. In the control cohort, there were 3.792 million people with a record of ethnicity of whom 118,821 (3.2%) were Black. The whole control cohort comprised 7.027 million people, 51% with a record of ethnicity and 49% without.

We repeated all the analyses of associations with individual cancers, above, on cohorts that included all SCD patients and equivalent controls under 65, regardless of ethnic group. The results (supplementary table) were broadly similar to those for Blacks only but, with larger numbers, confidence intervals were smaller. This comparison, as well as being between people with and without SCD, is also a comparison between a largely Black and a largely White cohort. The RR for all cancers combined was 1.4 (1.2–1.6). RRs for haematological cancers remained high: for example, 5.3 (3.0–8.8) for Hodgkin's disease and 8.9 (5.3–14.1) for multiple myeloma. There were significantly high RRs for cancers of the stomach, colon, pancreas, prostate and kidney. RRs became low for two cancers in which risks in Blacks are known to be lower than those in Whites: non-melanoma skin cancer (0.6, 0.3–1.0) and breast cancer in women (0.5, 0.3–0.8).

Discussion

Main findings

Black people with SCD were at higher risk than the Black control cohort of cancers overall and, notably, of haematological malignancies but also of some solid tumours.

Our rationale for analysing the rates of cancer in Black individuals was to take account of the effect of ethnicity on cancer risk in people with SCD. It can be assumed (whether or not ethnicity was recorded on individual records) that the great majority of the patients with SCD in England are from non-White ethnic groups, predominantly Black, but that the majority of people eligible to be in the control cohort, had we not allowed for ethnicity, would have been ethnic White. Unfortunately, the recording of ethnicity in HES is incomplete.[14] Nonetheless, the findings on cancers overall and on haematological malignancies, suggest that the increased risk of these is likely to be mediated through SCD rather than ethnicity.

Findings in this respect are less certain with some of the solid tumours. According to the National Cancer Intelligence Network, UK, the Black population has higher rates than the White of stomach, liver and prostate cancer.[16] Risks of these were not high in the comparisons between SCD and controls in the

Black cohorts; but they were high in the supplementary comparisons of all and unrecorded ethnicities. This suggests higher risks in Blacks than Whites, but no particular effect of SCD. By contrast, the risks of colon cancer and thyroid cancer were significantly elevated in cohorts that were restricted to Black patients. This suggests that the relationship between colon and thyroid cancer and SCD may be attributable more to SCD than ethnicity. The risk of female breast cancer and non-melanoma skin cancer, comparing SCD and control cohorts, was low in the 'all patients' cohorts, but this was not found in the analysis restricted to Black patients only (supplementary table). It is known that the rates of breast cancer and non-melanoma skin cancer are lower among Blacks than Whites.[16] The low risk of skin cancers in people with SCD no doubt reflects the well-known protective effects of ethnicity and skin colour. We note, however, the excess of non-melanoma skin cancer in people with SCD in the 'Black only' cohorts. We have no explanation for it.

Comparisons with literature and possible explanations for findings

This study is the first to quantify the overall risk of all cancers and the risk of a wide range of specific individual malignant neoplasms in a very large defined population of patients with SCD in England. However, there are publications reporting on increased risk of leukaemia and lymphoma, and on renal carcinoma, that accord with our findings.[3–9,17,18]

Biological mechanisms that could explain any relationship between SCD and cancer have not been sufficiently investigated experimentally. Among possible underlying mechanisms in an SCD-cancer association are endothelial damage, systemic inflammation, hypoxia, acidosis and compromised apoptosis.[19–24] Other potential causes and explanations for an excess risk of malignancies associated with SCD might include exposure to bone marrow transplantation, transfusion-related immunomodulation, and transmission of infections such as viral hepatitis and HIV.[25,26] The latter is probably less likely to occur in England, where standards of blood transfusion are better regulated, than in some other countries.

We have no information on prescribed drugs and cannot assess whether any of the excess risk of cancer is related to treatment of SCD patients. There have been reports suggesting that hydroxyurea, a drug widely used in management of patients with SCD, might trigger malignant transformation. However, findings from clinical trials do not support this.[27,28]

Strengths and weaknesses

This is a large population-based study, using a linked database containing all admissions to NHS hospitals in England. Its use of rates, i.e. population denominators, means that the scale of excess risk of cancer in people with SCD, compared with others, can be calculated. A strength is that the HES dataset includes information on patients' ethnicity. A weakness is that the data are incomplete and, even when recorded, are of unknown reliability.

Another limitation is that the dataset captures only hospital admissions and hospital day cases. This no doubt misses some people with SCD but is likely to identify the great majority of people with malignancies. The study design – a retrospective cohort study based on an administrative dataset – means that information about confounding factors is limited. It was not possible to take account of potentially confounding or effect-modifying factors including smoking, alcohol use and other environmental exposures which could predispose to cancer. We also have no data on treatment. The study is also reliant on diagnostic accuracy and the reliability of diagnostic coding. Current policy regulations preclude researchers from validating HES diagnoses by retrieving case-notes. However, in studies of other design, researchers have compared cancer diagnoses (the 'outcome' measure in this study) between individual records of HES and of cancer registries and have reported 98–99% agreement.[29]

We are not aware of studies of the reliability of the diagnosis of SCD in HES. However, given the fact that SCD predisposes to IPD, and that we found very high RRs for IPD in the SCD cohort, we suggest that this is strong evidence that most, if not all, people in the SCD cohort did in fact have SCD. Insofar as there may have been misclassification – for example, of sickle cell trait (even though we excluded the ICD code D57.3) as SCD – its effect would have been to attenuate the reported excess risk of cancer by 'dilution' of the SCD cohort. People with SCD in this study, namely those with a hospital day case or inpatient admission, come from a total SCD population of unknown size comprising all people with SCD under the age of 65 years who resided in England during the 13 years from 1999 to 2011. We have no information on people with SCD who do not get admitted as hospital day cases or inpatients. Therefore, our figures cannot be regarded as a point-prevalence census measure of SCD in England. As numerical context, the National Haemoglobinopathy Registry specifies that, nationally in England, 7338 patients with SCD were in the Registry in 2009–2014.[30]

Analysis of other linked datasets, to study cancer in people with SCD, may be possible: one example

would be the linkage of haemoglobinopathy registries to cancer registries.

The issue of studying multiple cancers, and therefore of multiple significance testing, needs consideration. We studied 26 cancers and, with a probability level of 1 in 20, would expect about one cancer to be significant by chance alone; 10 were significant. A Bonferroni correction would mean that significance could only be assumed at a significance level of $p = 0.002$ (0.05/26), and we give exact p values in Table 2. However, Bonferroni corrections can be unrealistically conservative in that, unless numbers of subjects are very large, they risk false negatives, i.e. rejection of associations that are real because the threshold for significance is set so high.

It is worth noting that numbers of individual cancers were very small, even when RRs were significant. For example, there were 14 cases of multiple myeloma in the 13 years of the study, i.e. about one a year in the whole of England. An individual clinician might not be able to have a clinical impression that the risk of multiple myeloma is elevated in people with SCD. The evidence is most likely to come from epidemiology.

Implications and conclusions

People with SCD may have an increased risk of cancer, particularly haematological malignancies and some individual solid tumours, and we have quantified these risks. Although the majority of our findings are consistent with the literature, it is important to treat our findings with caution and regard them as suggestive rather than conclusive. We have accounted for the effect of age and ethnicity on cancer risk but could not include other potential confounders. If our findings are confirmed elsewhere, physicians may wish to know that there might be an increased risk of malignancies in patients with SCD. If the cancer associations do indeed prove to be causally related to SCD itself, laboratory-based work on possible mechanisms to elucidate why people with SCD are at increased risk of quite a broad spectrum of malignancies might make contributions to fundamental understanding of some aspects of carcinogenesis.

Declarations

Competing interests: None declared

Funding: The Unit of Health-Care Epidemiology was funded by the English National Institute for Health Research [RNC/035/002] to build the linked datasets. MJG was part-funded by Public Health England.

Ethical approval: Ethical approval for analysis of the record linkage study data in a multi-purpose programme of research was obtained from the Central and South Bristol Multi-Centre Research Ethics Committee (04/Q2006/176).

Guarantor: MG

Contributorship: OO and MG conceived the study. MG designed the methodology. OS and DY did the analysis. OS and OO wrote the first draft. All authors contributed to the final draft.

Acknowledgements: Over many years, the linked datasets were built by Leicester Gill and Matt Davidson, Unit of Health–Care Epidemiology, University of Oxford. Mortality data were provided by the Office for National Statistics. The funders had no role in study design, data analysis, data interpretation, writing of the paper or the decision to submit for publication. The views expressed in this article do not necessarily reflect those of the funding bodies. All authors declare they are independent from all sources of funding.

Provenance: Not commissioned; peer-reviewed by Julie Morris and David Rees

References

1. Rees DC, Williams TN and Gladwin MT. Sickle-cell disease. *Lancet* 2010; 2376: 2018–2031.
2. National Health Service. *NHS choices*. Bristol, UK: IOP Publishing, See http://www.nhs.uk/conditions/Sickle-cell-anaemia/Pages/Introduction.aspx (last checked 4 September 2015).
3. Stricker RB, Linker CA, Crowley TJ and Embury SH. Haematological malignancy in sickle cell disease: report of four cases and review of the literature. *Am J Hematol* 1986; 21: 223–230.
4. Kim HS, Yospur L and Niihara Y. Chronic lymphocytic leukemia in a patient with sickle cell anaemia. *West J Med* 1998; 169: 114–116.
5. Paydas S. Sickle cell anaemia and haematological neoplasias. *Leuk Lymphoma* 2002; 43: 1431–1434.
6. Brown B and Kotila TR. Hodgkin lymphoma in a child with sickle cell anaemia. *Pediatr Hematol Oncol* 2007; 24: 531–532.
7. Chen L, Zhuang M, Shah HQ and Jen HL. Chronic myelogenous leukemia in sickle cell anemia. *Arch Pathol Lab Med* 2005; 129: 423–424.
8. Schultz WH and Ware RE. Malignancy in patients with sickle cell disease. *Am J Hematol* 2003; 74: 249–253.
9. Dawkins FW, Kim KS, Squires RS, Chisholm R, Kark JA, Perlin E, et al. Cancer incidence rate and mortality rate in sickle cell disease patients at Howard University Hospital: 1986–1995. *Am J Hematol* 1997; 55: 188–192.
10. Goldacre MJ, Kurina L, Yeates D, Seagroatt V and Gill L. Use of large medical database study associations between diseases. *QJM* 2000; 93: 669–675.
11. Goldacre MJ, Wotton CJ, Seagroatt V and Yeates D. Cancers and immune related diseases associated with Down's syndrome: a record linkage study. *Arch Dis Child* 2004; 89: 1014–1017.
12. Turner M, Goldacre R and Goldacre M. Reduced cancer incidence in Huntington's disease: record linkage study clue to an evolutionary trade-off? *Clin Genet* 2013; 83: 588–590.

13. Seminog OO and Goldacre MJ. Risk of benign tumours of nervous system, and of malignant neoplasms, in people with neurofibromatosis: population-based record-linkage study. *Br J Cancer* 2014; 108: 193–198.
14. Saunders CL, Abel GA, El Turabi A, Ahmed F and Lyratzopoulos G. Accuracy of routinely recorded ethnic group information compared with self-reported ethnicity: evidence from the English Cancer Patients Experience Survey. *BMJ Open* 2013; 28: 3. DOI: 10.1136/bmjopen-2013-002882.
15. Breslow NE and Day NE. Statistical methods in cancer research. In: Breslow NE and Day NE (eds) *The Design and Analysis of Cohort Studies*. Lyon: International Agency for Research in Cancer, 1987, pp.103–115.
16. National Cancer Intelligence Network and Cancer Research. *Cancer incidence and survival by major ethnic group, England, 2002–2006*. See http://www.google.co.uk/url?sa=t&rct=j&q=&esrc=s&source=web&cd=1&ved=0CDkQFjAA&url=http%3A%2F%2Fwww.ncin.org.uk%2Fview%3Frid%3D75&ei=ERy_UeKQF4jCPIGMgKAG&usg=AFQjCNGwcQ8VLlvNpS8WBwlm5_cftE271w&sig2=EC0J7BsH8-6jURiGziGs5g&bvm=bv.47883778,d.ZWU (2010, last checked 4 September 2015).
17. Simpson L, He X, Pins M, Huang X, Campbell SC, Yang XJ, et al. Renal medullary carcinoma and ABL gene amplification. *J Urol* 2005; 173: 1883–1888.
18. Noguera-Irizarry WG, Hibshoosh H and Papadopoulos KP. Renal medullary carcinoma: case report and review of the literature. *Am J Clin Oncol* 2003; 26: 489.
19. Muz B, De La Puente P, Azab F, Luderer M and Azab AK. The role of hypoxia and exploitation of the hypoxic environment in hematologic malignancies. *Mol Cancer Res* 2014; 12: 1347–1354.
20. Hebbel RP, Osarogiagbon R and Kaul D. The endothelial biology of sickle cell disease: inflammation and a chronic vasculopathy. *Microcirculation* 2004; 11: 129–151.
21. Makis A, Hatzimichael E and Bourantas K. The role of cytokines in sickle cell disease. *Ann Hematol* 2000; 79: 407–413.
22. Solovey A, Gui L, Ramakrishnan S, Steinberg MH and Hebbel RP. Sickle cell anemia as a possible state of enhanced anti-apoptotic tone: survival effect of vascular endothelial growth factor on circulating and unanchored endothelial cells. *Blood* 1999; 93: 3824–3830.
23. Majmundar AJ, Wong WJ and Simon MC. Hypoxia-inducible factors and the response to hypoxic stress. *Mol Cell* 2010; 40: 294–309.
24. Baron F, Dresse MF and Beguin Y. Transmission of chronic myeloid leukemia through peripheral-blood stem-cell transplantation. *N Engl J Med* 2003; 349: 913–914.
25. Vamvakas EC and Blajchman MA. Deleterious clinical effects of transfusion-associated immunomodulation: fact or fiction? *Blood* 2001; 97: 1180–1195.
26. Steinberg MH, McCarthy WF, Castro O, Ballas SK, Armstrong FD, Smith W, et al. The risks and benefits of long-term use of hydroxyurea in sickle cell anemia: a 17.5 year follow-up. *Am J Hematol* 2010; 85: 403–408.
27. McGann PT, Flanagan JM, Howard TA, Dertinger SD, He J, Kulharya AS, et al. Genotoxicity associated with hydroxyurea exposure in infants with sickle cell anemia: results from the BABY-HUG phase III clinical trial. *Pediatr Blood Cancer* 2012; 2: 254–257.
28. Møller H, Richards S, Hanchett N, Riaz SP, Lüchtenborg M, Holmberg L, et al. Completeness of case ascertainment and survival time error in English cancer registries: impact on 1-year survival estimates. *Br J Cancer* 2011; 105: 70–176.
29. Foster M. National Haemoglobinopathy Registry Report 2013/14. See http://r.search.yahoo.com/_ylt=A0LEViMCoQJX1ogA1jAnnIlQ;_ylu=X3oDMTByOHZyb21tBGNvbG8DYmYxBHBvcwMxBHZ0aWQDBHNlYwNzcg–/RV=2/RE=1459818883/RO=10/RU=http%3a%2f%2fwww.nhr.nhs.uk%2fwp-content%2fuploads%2f2015%2f10%2fNHR_AnnualReport_2014.pdf/RK=0/RS=cFZHE7d8egA7WHzjWEkE.Fl3f9g- (2014, last checked 17 May 2016).