

# Retrospective case-series analysis of haematological malignancies in goldmining areas of South Africa

T Inamasu,<sup>1</sup> PhD; M Patel,<sup>2</sup> PhD; C Espina,<sup>1</sup> PhD; A Pentz,<sup>3</sup> BSc; M Joffe,<sup>3</sup> MSc; F Winde,<sup>4</sup> PhD; J Schüz,<sup>1</sup> PhD

<sup>1</sup> Section of Environment and Radiation, International Agency for Research on Cancer, Lyon, France

<sup>2</sup> Division of Clinical Haematology, Department of Medicine, Chris Hani Baragwanath Academic Hospital and Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa

<sup>3</sup> Wits Health Consortium, MRC/Wits Development Pathways to Health Research Unit, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa

<sup>4</sup> Mine Water Research Group, School of Geo and Spatial Sciences, Faculty of Natural and Agricultural Sciences, North-West University, Vanderbijlpark, South Africa

**Corresponding author:** C Espina ([espina@iarc.fr](mailto:espina@iarc.fr))

**Background.** South Africa (SA)'s high levels of environmental contamination of mine tailings from uranium and its decay products, coupled with remarkably short distances between mine tailings and residential areas, raise concern about whether there is an association between environmental uranium exposure and risk of cancer, including haematological malignancies.

**Objectives.** We reviewed information on cases from the central hospital offering cancer diagnostics and treatment in a major mining area of SA to describe their basic clinical and demographic characteristics, as part of assessing whether a cancer epidemiological study in this area would be feasible.

**Methods.** Basic clinical, demographic and residential information on patients with haematological malignancy diagnosed between 2004 and 2013 was collected retrospectively from the patient files at Chris Hani Baragwanath Academic Hospital in Soweto, Johannesburg.

**Results.** In total, 1 880 patients aged 18 - 94 years were identified. Referral from distant provinces was not uncommon, but >80% lived within 50 km of the hospital. Non-Hodgkin's lymphoma accounted for 44% of the haematological malignancies, followed by leukaemia with 26%. HIV status was known for 93% of the patients, of whom 47% were HIV-positive.

**Conclusions.** Caution is required when interpreting spatial distributions of patients, given inaccuracies in residential addresses and referral patterns to the hospital, and with HIV and other infections probable important confounders. Our study therefore shows that active case recruitment is required for accurate assessment of residential information. However, some findings on spatial distributions in the study warrant the continuation of efforts to develop a study protocol to investigate the possible link between uranium exposure in mining areas and haematological malignancies in residents. Disproportionately high incidence rates of haematological malignancies observed in specific districts would be relevant for further investigation.

*S Afr Med J* 2018;108(10):858-864. DOI:10.7196/SAMJ.2018.v108i10.13175

Uranium has radioactive and chemotoxic properties. Evidence of cancer risk due to uranium as a radioactive element (as alpha-transmitter and from its decay products such as radon and radium of higher radioactivity) has been established.<sup>[1,2]</sup> According to the International Agency for Research on Cancer (IARC)'s Monograph programme on the evaluation of carcinogenic risks to humans, there is limited evidence for the carcinogenicity of mixtures of uranium (U) isotopes in humans.<sup>[1]</sup> There is sufficient evidence in experimental animals for the carcinogenicity of <sup>233</sup>U, <sup>234</sup>U, <sup>235</sup>U and <sup>238</sup>U (of which only the latter three isotopes occur naturally and are found in varying ratios in natural, enriched and depleted uranium).<sup>[1]</sup> An association between environmental uranium exposure and risk of haematological malignancies has been hypothesised, but the evidence is inconclusive.<sup>[2-4]</sup> Started in the mid-1880s, goldmining and milling over the past 130 years in the West Rand and Far West Rand basins of South Africa (SA) resulted in containment of an estimated >600 000 tons of uranium in deposited mining waste.<sup>[5]</sup> Locally high levels of environmental contamination from uranium and its decay products in water, soil, sediment and dust have been reported.<sup>[6]</sup> Owing to frequently short distances between mine tailings and residential areas, populations living around the goldmine tailings of the West Rand and Far

West Rand goldfields may be exposed to uranium and uranium decay products from the tailings through multiple pathways. These include ingestion of contaminated water, consumption of possibly contaminated local products (e.g. vegetables, milk, meat) and inhalation of dust. In addition, direct consumption of the contaminated soil and tailings material (geophagia) is a suspected pathway. Geophagia is intentional habitual soil ingestion, common in sub-Saharan Africa including SA and particularly prevalent among pregnant women, but small children also ingest soil via unintentional uptake by hand-to-mouth contact. Geophagia is practised for treatment of gastrointestinal symptoms and as a morning-sickness remedy, and is related to mineral (e.g. iron) deficiency as well as psychological and sociocultural factors.<sup>[7]</sup> Risks associated with ingestion of contaminated soil including exposure to toxic chemicals have been hypothesised.<sup>[8,9]</sup>

The scenario described above merits research for two reasons, outlined during an expert workshop held by the IARC in Johannesburg on 13 - 14 March 2013.<sup>[9]</sup> From a public health perspective, if environmental exposure to uranium does lead to increased human exposure, large populations may be affected, and resulting health risks need to be mitigated. From a scientific perspective, this may be a unique setting to address the still

inconclusive question of whether there is a detectable association between environmental uranium exposure and risk of cancer, including haematological malignancies.

In SA, ~2 000 - 3 000 cases of haematological malignancy were reported to the National Cancer Registry (NCR) each year between 2004 and 2011.<sup>[10]</sup> The NCR is a pathology-based cancer registry and its completeness of coverage of haematological malignancies, especially in less affluent areas, is not known. However, haematological malignancies accounted for 5.1% of total cancer cases during this period. More than half (50.9%) were non-Hodgkin's lymphomas (NHLs), followed by leukaemia (21.8%). In 2011, age-standardised rates per 100 000 population for all haematological malignancies combined were 9.25 for males and 6.19 for females. Only 37% of reported cases in 2011 were in black Africans, who account for 80% of the total population of SA.

## Objectives

Chris Hani Baragwanath Academic Hospital (CHBAH) is a tertiary hospital located in Soweto, Johannesburg. The largest public hospital on the entire African continent, it has a patient catchment area covering the whole of Soweto and surrounding areas up to 30 km away and receives patients from referral hospitals located up to 150 km away.

To date, there have been no epidemiological studies investigating the risk of cancer in the exposed populations living around the mine tailings. After the workshop mentioned above,<sup>[9]</sup> several pilot activities commenced in the framework of a feasibility study, with some aiming to quantify human exposure and how best to measure it, and one aiming to obtain an overview of patients with haematological malignancies in the area reported on here. The main objective of the case series analysis of haematological cancer patients from CHBAH was to describe the basic clinical and demographic characteristics of the patients and their spatial distribution. A second objective was to ascertain how many patients came from areas with reported elevated levels of uranium and uranium decay products. Acknowledging that a range of other factors such as referral patterns could play a major role in where the patients come from, this analysis was exploratory in nature. The outcomes would, however, assist in deciding whether a full-scale analytical study on uranium exposure and haematological cancer risk is feasible in this setting in terms of patient recruitment.

## Methods

All cases of haematological malignancy diagnosed at CHBAH between the years 2004 and 2013 inclusive were included in this retrospective study based on routine clinical documentation. We extracted general demographic (sex and age at diagnosis) and clinical (precise cancer diagnosis, date of diagnosis and death, HIV status at diagnosis) information from the medical records of the Division of Clinical Haematology in the Department of Medicine at CHBAH, where all patients with a haematological malignancy are seen. The procedure was that all medical records of the Division from the period 2004 - 2013 were individually checked for any mention of a haematological malignancy (by AP) and eligibility to be included in the study database was confirmed by the lead haematologist (MP). Diagnosis was given as either a suspected or initial diagnosis (mostly without a specified disease subtype) as well as a final diagnosis, the latter generally with a disease subtype confirmed by a pathology report where available. We coded these diagnoses according to the *International Classification of Diseases for Oncology*, 3rd edition (ICD-O-3) and grouped them into five broad disease categories: Hodgkin's lymphomas (HL), NHL, myeloma and related disorders (MM), leukaemia, and myelodysplastic/myeloproliferative neoplasms

(MDS). Cases that had only a suspected/initial diagnosis, and cases with a final diagnosis but without detailed classification, were treated and coded as subgroups 'not otherwise specified'. Cases with features of more than two subtypes within a broad disease group, such as 'NHL unclassifiable between diffuse large B-cell lymphoma and Burkitt's lymphoma', were kept as they appeared in the original data. Cases with features of more than two subtypes across two broad disease groups, such as 'Plasmablastic lymphoma and multiple myeloma', were treated as unclear diagnoses and excluded from detailed analysis (but shown in the total number).

We also obtained the addresses of the patients from the medical records, supplemented by outpatient files, inpatient files, electronic hospital records and National Health Laboratory Service laboratory data. The local study team collected the address information from the widely distributed electronic health records system in SA called the Medicom system, by linking to the hospital file number, and converted it into co-ordinate data using a geographic information system (GIS) and Google Earth, a virtual geographic environment (VGE). We recoded the obtained co-ordinate data in GIS and identified sub-place code, sub-place name, municipality code, municipality name, district code, district name, province, ward number and ward ID, which was possible most of the time even with uncertainties in exact GIS co-ordinates, as not all areas in and around Soweto use house numbers or even street names or those data were not recorded in detail. Sub-place refers to the smallest administrative boundary in SA, which is equivalent to a suburb, a village or a farm; ward is an electoral zone of a similar population size. We calculated the most direct distance from the patient residence to CHBAH (co-ordinates: -26.2597753, 27.9406257) using the haversine formula.

For the analyses, we mainly prepared frequency tables for clinical and for demographic variables. To calculate incidence rates, we used population denominators by administrative boundaries (provinces and wards) reported in the 2011 SA census. We combined haematological malignancy cases for the entire study period (10 years) and obtained 10-year average annual incidence rates to avoid reporting too-strong random fluctuations caused by the small number of cases in each year. We further calculated truncated age-standardised incidence rates (for ages >20 years) using the Segi World standard population<sup>[11]</sup> for wards with >20 cases during the entire study period, for the same reason. Note that the calculated rates represented the incident cases from CHBAH in relation to population sizes of provinces and wards, but are not to be confused with population-based incidence rates.

## Ethical considerations

The study was approved by the IARC Ethics Committee (project no. 14-19). Only the local study team had access to the raw and non-anonymised data.

The data that support the study findings are available from the Division of Clinical Haematology, Department of Medicine, CHBAH, but restrictions apply to the availability of these data, which were used under licence for the current study and so are not publicly available. Data are, however, available from the authors upon reasonable request and with permission from CHBAH.

## Results

### General clinical characteristics of patients

Initially, 1 925 files were entered into the study database as potentially relevant to the study. These included 5 non-cancerous and 2 non-haematological malignancy cases as well as 13 patients aged <18 years at diagnosis. Eleven patients were registered a second time during a repeat visit. All duplicate and irrelevant cases were excluded from the

study population, leaving 1 894 eligible cases. For 14 of these cases, the patient files were kept elsewhere within the hospital for academic purposes and detailed clinical information was not available to us by the end of the study period, leaving a total of 1 880 patients aged 18 - 94 years for the study. For one case, sex was not recorded, and for 27 cases the age was unknown. HIV status was available for 42% of the total cases in 2004, but for >95% from 2006 on. In total, HIV status at diagnosis was known for 93% of the patients, of whom 47% were HIV-positive.

Race was not reported in any of the files, but CHBAH as a public hospital traditionally serves the black African population living in its large catchment area (confirmed by the head of the Division of Clinical Haematology, Prof. M Patel, 6 September 2017).

Table 1 shows the basic characteristics of the cases of haematological malignancy identified in the study. NHLs comprised 44% of the total, followed by leukaemia (26%) and MM (17%). Overall the sex ratio (M:F) was 1.03, but it differed by diagnosis, with NHL and MM affecting slightly more women, whereas HL, leukaemia and MDS were seen more often in men. Overall, 15% of patients were aged ≤30 years at diagnosis and 38% were aged ≥51 years, but age distributions differed across diagnostic subtypes. In 2004 and 2005, the NHL subtype was known for only 30% of patients diagnosed with NHL (NHL subtype was not stated in the database for 2004 and

2005 in 70% of the patients, although it was known for most of the patients). This improved to 84 - 90% from 2008 onwards. Throughout the study period, the majority of NHL patients had diffuse large B-cell lymphoma (DLBCL). Burkitt's lymphoma/leukaemia (BL) and plasmablastic lymphoma (PBL) were also common in this study population. Patients with these three NHL subtypes or with primary effusion lymphoma (PEL) were predominantly HIV-positive (81 - 100%), confirming these subtypes to be HIV-associated, as previously reported.<sup>[12-15]</sup> More women with HIV-associated lymphomas had DLBCL or BL, while PEL and PBL were more often seen among men. HIV prevalence was generally low among patients with leukaemia, as well as among those with cancer types that affect older age groups such as MM and MDS.

### Geographical distribution of patients

For each year, up to 60% of the registered addresses were found to be either incomplete or inaccurate, because local street names are often incorrectly recorded in the database or are not recognisable in the GIS system (and sometimes do not even exist, e.g. in informal settlements). The study population also included a number of patients (n=27) whose current residence or duration of stay in the reported residence was unknown. For these patients, only the address of a temporary residence was available, such as a referral hospital,

**Table 1. Basic characteristics of 1 880 patients diagnosed with a haematological malignancy at Chris Hani Baragwanath Academic Hospital, Johannesburg, South Africa, between 2004 and 2013**

	Cumulative total, n (%)	Age group (years), %				Sex ratio (M/F)	HIV	
		18 - 30	31 - 40	41 - 50	≥51		With data, n (%)	Positive, n (%)
NHL	822 (44)	16	35	27	22	0.96	782 (95)	621 (76)
Myeloma and related disorders	325 (17)	0	6	23	71	0.92	291 (90)	26 (9)
HL	188 (10)	27	41	21	11	1.24	174 (93)	120 (69)
Leukaemias	487 (26)	20	13	18	50	1.11	453 (93)	54 (12)
MDS/MPN	45 (2)	2	11	7	80	1.25	45 (100)	3 (7)
Total (including 13 cases with unclear diagnosis)	1 880 (100)	15	24	23	38	1.03	1 757 (93)	830 (47)
NHL subtypes								
B-cell								
DLBCL	275 (34)	12	35	31	22	0.82	268 (97)	218 (81)
BL	127 (16)	28	42	21	10	0.81	124 (98)	120 (97)
PBL	115 (14)	14	43	34	10	1.59	113 (98)	110 (97)
DLBCL/BL	39 (5)	13	50	18	18	0.86	39 (100)	38 (97)
FL	15 (2)	0	13	27	60	1.14	15 (100)	0
PEL	8 (1)	25	50	13	13	1.67	7 (88)	7 (100)
MCL	7 (1)	0	0	0	100	2.50	7 (100)	0
MALT	7 (1)	14	0	43	43	2.50	7 (100)	3 (43)
DLBCL/PBL	5 (1)	0	40	40	20	1.50	5 (100)	5 (100)
T-cell								
ALCL	18 (2)	28	28	28	17	0.80	15 (83)	7 (47)
PTCL	8 (1)	13	13	13	63	0.14	7 (88)	2 (29)
Mycosis fungoides	6 (1)	0	17	17	67	1.00	6 (100)	0
NHL, other subtypes	16 (2)	19	13	13	56	0.78	16 (100)	1 (7)
NHL, not otherwise specified	176 (21)	16	31	28	24	1.02	153 (87)	110 (72)
Leukaemia subtypes								
Acute myeloid	204 (42)	25	16	20	40	1.00	196 (96)	28 (14)
Chronic lymphocytic	135 (28)	1	2	11	87	1.21	121 (90)	6 (5)
Chronic myeloid	112 (23)	18	20	25	37	1.00	102 (91)	16 (16)
Acute lymphocytic	36 (7)	64	17	14	6	2.00	34 (94)	4 (12)

NHL = Non-Hodgkin's lymphoma; HL = Hodgkin's lymphoma; MDS/MPN = myelodysplastic/myeloproliferative neoplasms; DLBCL = diffuse large B-cell lymphoma; BL = Burkitt's lymphoma/leukaemia; PBL = plasmablastic lymphoma; FL = follicular lymphoma; PEL = primary effusion lymphoma; MCL = mantle cell lymphoma; MALT = mucosa-associated lymphoid tissue lymphoma; ALCL = anaplastic large-cell lymphoma; PTCL = peripheral T-cell lymphoma.

hostel, nursing home for the elderly or prison. In some cases, GIS co-ordinates were estimated by a given post office box number, the name of a nearby train station or the centre point of an administrative region, instead of an actual complete address. The address was missing for 0.9 - 13.1% of the total cases reported each year. However, we were still able to locate the address for 1 762 cases for assignment to the smallest administrative unit. Four cases with a foreign address were excluded (three from neighbouring countries within Africa and one from Oceania). Among the remaining 122 cases with address unknown, 14 had an invalid patient identifier number and one had an incomplete address, but the majority had no address at all registered in any of the available databases.

Among 1 758 patients with an SA address available, 90.5% came from Gauteng Province and 8.7% from North West Province (Table 2). Referral from distant provinces does occur, but >85% of the patients came from within 50 km of the hospital (Table 3), most originating from Soweto, where the hospital is located.

Table 3 shows the crude incidence rate (CIR, cases per 100 000 residents per year) and the age-adjusted incidence rate (ASR) by province and their major geographical units (main place). The mean distance from the patients living in Soweto to the hospital was 6.9 km. In addition, 154 patients lived in the surrounding areas of Johannesburg, while 181 came from a few larger townships nearby: 87 from Orange Farm, 53 from Sebokeng and 41 from Katlehong. With a total of 46, the highest number of patients from North West came from a single referral hospital in Klerksdorp, 148.3 km from CHBAH.

Of the wards, 303 had at least one haematological malignancy patient during the study period, among them 24 wards with >20 patients over the 10 years. The incidence rates were highest in Ward

64003010 of Klerksdorp (CIR 30.2, ASR 52.7), and lowest in Ward 79800053 of Dobsonville Gardens (CIR 3.8, ASR 6.9). Table 4 shows the 21 wards that had a minimum of 20 patients during the study period and an ASR >10.0 per 100 000 per year, which is roughly double the ASR for the whole of SA according to SA NCR data.<sup>[10]</sup> HIV prevalence among patients from these wards was generally high. According to the census of 2011, the proportion of dwellings in informal settlements in these wards (as a percentage of the total number of dwellings) ranged from 2.1% to 34.5%. Black Africans comprised the overwhelming majority of the population (99 - 100%) in 19 of the 21 wards, except for Lenasia which is known to be an Indian community, and Eldorado Park, predominantly inhabited by coloured (mixed race) people.

Fig. 1 shows the spatial distribution of wards with elevated cancer rates (with yellow to red shading indicating increasing cancer rates) in relation to lease areas (white boundary lines) and tailings deposits (bright red shading) of surrounding goldmines. The majority of areas with moderately high incidence rates were found in Soweto, close to CHBAH and within its catchment area. High incidence rates outside Soweto (Eldorado Park and Lenasia) were within a similar distance from CHBAH as some western parts of Soweto. The few sporadic wards with high incidence rates not in close proximity to CHBAH were Klerksdorp (outside the map in Fig. 1, also located in a goldmining area), Orange Farm and Ennerdale.

## Discussion

Our retrospective case series found that a large number of patients can be recruited through CHBAH: each year during the 10-year study period, between 118 and 225 cases of haematological malignancy were seen at CHBAH. This corresponds to 4.7 - 7.1% of total cases of haematological malignancy and 9.7 - 12.3% of those in black Africans reported to the NCR between 2004 and 2011.<sup>[10]</sup>

The haematological malignancy patients at CHBAH came from a much larger catchment area than has previously been expected. This included neighbouring towns mostly within a 40 km radius of CHBAH, including the southern parts of Johannesburg, that are in fact closer to other tertiary or quaternary public hospitals such as Charlotte Maxeke Johannesburg Academic Hospital and Helen Joseph Hospital. Apart from Soweto, which straddles the Central Rand and West Rand goldfields, a relatively large number of patients also came from nearby townships with limited local medical

**Table 2. Geographical distribution of patients (registered address at diagnosis): distance to the hospital (N=1 758)**

Distance (km)	Patients, n (%)
<1	766 (43.5)
10 - 19	347 (19.7)
20 - 29	229 (13.0)
30 - 39	112 (6.4)
40 - 49	74 (4.2)
≥50	234 (13.3)

**Table 3. Geographical distribution of patients (registered address at diagnosis): incidence by registered address at diagnosis (N=1 758)\***

Province/main place	Mean distance (km)	n (% of total)	Population	CIR	ASR	Population (equivalent wards)
Gauteng	16.0	1 591 (90.5)	12 272 263	1.3	2.2	12 272 263
Soweto	6.9	814 (46.3)	1 271 628	6.4	10.7	1 321 333
Johannesburg	7.7	154 (8.8)	957 441	1.6	3.2	857 207
Orange Farm	24.5	87 (4.9)	76 767	11.3	13.6	118 682
Lenasia	12.4	58 (3.3)	89 714	6.5	6.7	156 140
Sebokeng	35.1	53 (3.0)	218 515	2.4	5.6	152 163
Katlehong	24.0	40 (2.3)	407 294	1.0	2.6	311 581
North West	182.7	153 (8.7)	3 509 953	0.4	0.8	3 509 953
Klerksdorp	148.3	46 (2.6)	186 515	2.5	6.7	107 480
Mpumalanga	204.2	7 (0.4)	-	-	-	-
Free State	65.7	3 (0.2)	-	-	-	-
Eastern Cape	713.1	2 (0.1)	-	-	-	-
KwaZulu-Natal	478.3	1 (0.1)	-	-	-	-
Western Cape	1 165.9	1 (0.1)	-	-	-	-

CIR = crude incidence rate (cases per 100 000 residents per year); ASR = age-adjusted incidence rate.

\*Population, CIR, ASR and Population (equivalent wards) were not calculated for Mpumalanga, Free State, Eastern Cape, KwaZulu-Natal and Western Cape owing to the very low numbers (<10 cases).

**Table 4. Wards with >20 patients with haematological malignancies diagnosed between 2004 and 2013 at Chris Hani Baragwanath Academic Hospital, Johannesburg, South Africa**

Ward no.	Sub-place(s)	Main place	Province	n	CIR	ASR	HIV-positive, %	Informal, %	Distance, km	Notes
64003010	Klerksdorp	Klerksdorp	North West	21	30.2	52.7	45.0	5.4	150.9	Referral hospital (20 from the same spot)
79800004	Orange Farm Ext. 8	Orange Farm	Gauteng	78	18.7	34.0	58.1	34.5	24.3	PO box registered (77 from the same spot)
79800031	Orlando East	Soweto	Gauteng	37	13.4	25.9	40.6	31.2	2.5	Hospice (21 from the same spot)
79800051	Zola	Soweto	Gauteng	37	13.0	22.8	42.9	7.3	10.5	13 from the same primary school; 10 from the same spot
79800012	Chiawelo Exts 2, 3, 4, 5	Soweto	Gauteng	31	11.4	22.7	44.8	11.0	8.8	GP's address registered (12 from the same spot); 7 from the same community/flat
79800035	Jabavu, Mofolo Central	Soweto	Gauteng	32	12.7	20.5	44.8	11.8	6.7	18 from the same primary school
79800026	Diepkloof Zones 2, 5, 6	Soweto	Gauteng	31	12.2	20.1	58.6	11.6	0.9	Exposed area. GP's address registered (6 from the same GP)
79800028	Diepkloof Zones 1, 2, Orlando East	Soweto	Gauteng	34	11.7	18.5	54.8	15.7	1.9	Exposed area. 9 from the same spot
79800016	Mapetla, Phiri	Soweto	Gauteng	24	8.9	18.4	31.8	6.5	9.3	8 from one spot; 7 from another spot
79800046	Jabulani, Zola, Zondi	Soweto	Gauteng	28	9.8	18.3	42.9	9.3	8.4	Hospital address registered (9 from the same district hospital)
79800036	Mofolo Central, Mofolo North	Soweto	Gauteng	24	9.8	17.9	45.5	7.9	6.2	5 from the same park
79800015	Phiri, Senaoane	Soweto	Gauteng	28	10.5	17.2	51.9	12.3	8.1	6 from the same spot
79800039	Orlando West, Meadowlands East Zone 5	Soweto	Gauteng	35	11.0	17.0	56.7	11.5	4.3	11 from the same tourist spot; 10 from another spot
79800017	Eldorado Park	Johannesburg	Gauteng	29	7.7	16.7	40.7	23.3	5.7	Black population is minority (ASR for black Africans 54.3). 20 from the same spot; 6 from another spot
79800048	Dobsonville, Dobsonville Exts 3, 5, 7	Soweto	Gauteng	28	8.7	15.6	45.8	12.6	9.1	GP's address registered (14 from the same GP)
79800020	Naledi Ext. 1	Soweto	Gauteng	29	9.1	14.7	56.0	4.4	11.1	15 from the same spot; 5 from the same police station
79800007	Ennerdale Exts 1, 3, 6, 12, 13, 14, Finetown, Mid-Ennerdale	Ennerdale	Gauteng	30	7.7	14.5	65.5	31.4	19.3	Black population is minority (ASR for black Africans 29.0). Clinic address registered (8 from the same clinic)
79800038	Dube, Orlando West	Soweto	Gauteng	20	9.2	14.5	47.4	3.4	5.3	9 from the same spot
79800009	Lenasia Exts 1, 2, 3, 5, 7, 8	Lenasia	Gauteng	28	9.3	13.1	38.5	2.1	12.7	Black population is minority (ASR for black Africans 137.6). 12 from the same religious institute
79800025	Klipspruit, Pimville Zones 1, 2, 4	Soweto	Gauteng	30	8.4	13.0	65.5	19.6	3.1	4 from the same church
79800033	Mofolo South, Molapo, Moroka	Soweto	Gauteng	22	7.9	12.2	38.1	9.4	6.7	5 from the same spot

CIR = crude incidence rate (cases per 100 000 residents per year); ASR = age-adjusted incidence rate; Informal = population living in informal dwellings; Distance = mean distance to hospital; Ext./Exts = Extension/s; GP = general practitioner.



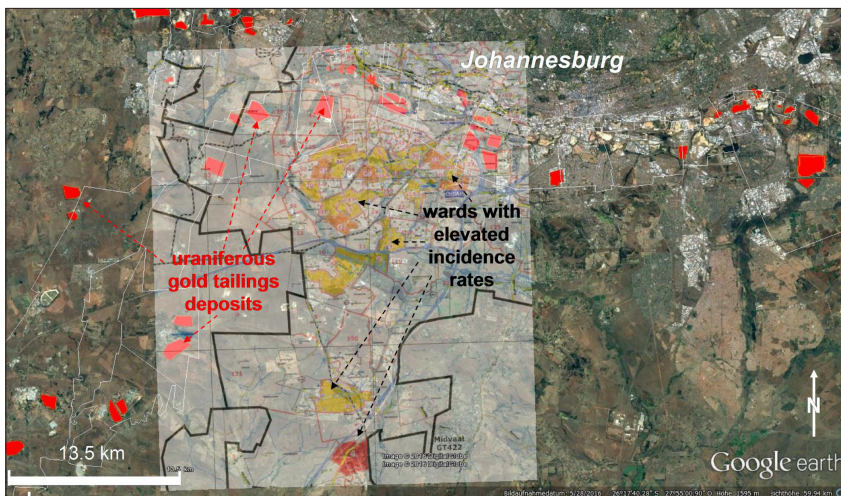


Fig. 1. Google Earth satellite image of the study area indicating the location of goldmines (lease areas, white boundary lines), uraniferous tailings deposits (red) and a superimposed map of wards with elevated rates of haematological malignancies (calculated based on total population, with yellow to red shading indicating increasing rates of haematological malignancies).

facilities, such as Orange Farm, Sebokeng and Katlehong, as well as from the Indian settlement of Lenasia. Overall, the catchment area in this study resembled that for breast cancer patients at CHBAH,<sup>[16]</sup> except that some haematological malignancy patients came from even more far-flung areas of up to 150 km away (i.e. the Klerksdorp goldfield).

The clinical information obtained in this study was relatively complete, and ~40% of the cases were confirmed by pathology reports. Given the largely unknown cancer patterns in sub-Saharan Africa, the data presented are of importance in understanding the haematological malignancy burden of the area. In accordance with the results of another study in a similar geographical setting (the broader metropolitan municipality of Johannesburg) between 2004 and 2009,<sup>[12]</sup> DLBCL was the leading NHL subtype over the entire study period.

The prevalence of HIV infection in our study population was ~2.5 - 4 times higher than the national level of 10.2% (all ages) or 16.8% (15 - 49-year-olds),<sup>[17]</sup> and more or less the same as in the abovementioned study (44.3 - 62.0%),<sup>[12]</sup> suggesting that it is the typical prevalence for adult haematological malignancy patients. Infections play an important role in the aetiology of haematological malignancies. For example, Epstein-Barr virus infection is related to risk of BL, HL, nasal natural killer T-cell lymphoma and other B-cell NHLs.<sup>[18]</sup> HIV infection increases the risk of developing NHL, HL and leukaemia. Previous studies have shown some NHL subtypes, such as DLBCL, BL and PBL, to be associated with HIV.<sup>[12-15]</sup> This has implications for any future studies, as infections need to be accounted

for when investigating environmental risk factors, and information on HIV status in particular needs to be collected. Our active case collection from clinical records revealed much more detailed clinical information than would have been obtainable through the NCR data. Information on detailed NHL subtypes is important in understanding the potential impact of infections on the observed burden.

The categorisation of haematological malignancies in broad disease groups showed NHL to be the most common in this study population, followed by leukaemia. This pattern corresponds to that of NCR data,<sup>[10,19]</sup> as well as figures for sub-Saharan Africa in general.<sup>[18]</sup> In black Americans in the USA and black African ethnic groups in the UK, on the other hand, exceptionally high incidence rates of myeloma and related disorders render this disease group the second most common following NHL.<sup>[20,21]</sup> However, if each of the four subtypes of leukaemia is regarded as a separate entity, multiple myeloma as a single disease entity is the second most common haematological malignancy after NHL. The different disease profile seen in the high-income countries would mainly be explained by demographic factors such as life expectancy and age structure of the countries, and also the HIV burden.

### Study limitations

The spatial distribution of incidence rates indicates the presence of geographical clusters. However, in our study this needs very careful interpretation, as with a hospital-based case collection clusters may have formed as a result of referral patterns rather

than the true disease burden in the area. For example, of the 21 patients from the ward in Klerksdorp with the highest ASR in this study, 20 originated from exactly the same geographical co-ordinates, reflecting the location of the referring hospital rather than the patients' residential addresses. Similarly, the third-largest number of patients came from Orlando East in Soweto, where 21 out of 37 patients were registered at the address of a specific hospice. Clusters at addresses of district hospitals or general practitioners were also observed for Chiawelo, Jabulani and Dobsonville. This illustrates the practical difficulties arising from the frequent lack of street addresses and house numbers in many townships.

In addition to such referral patterns, several other factors such as the poor population registration of the area, migration of residents and individual choice of hospital were found to preclude any future study from population-based recruitment via the CHBAH route alone. A specific sub-place (Extension 8) in Orange Farm had the largest number of patients, but the co-ordinates of their addresses showed that 77 out of 78 patients from Orange Farm used a post box number as their address. Some areas including Orange Farm are particularly impoverished and have large informal settlements, so that accurate residential co-ordinates cannot be established from hospital records alone. In addition, there are substantial differences between the census population and the actual number of people living in the areas, introducing further uncertainty in the estimation of any rates. In some other cases, exactly the same address (school, religious institution, train station, police station, park, etc.) was used for several patients. In any future epidemiological studies, patients would have to be interviewed about their true residential history, as neither the hospital records nor the NCR has sufficiently reliable and detailed information. This is important for investigating the possible relationship between uranium exposure and cancer, as patients may have moved away from the address they lived at during the relevant exposure time window, which may range from several years before diagnosis up to decades before diagnosis.

Another limitation of the retrospective collection of clinical data in this study was that data on patient ethnicity were not available, although 95% of patients were assumed to be black. Ethnic composition of the study population is directly related to different incidence patterns as well as patient choice of hospital. Previous studies have

revealed that white patients mostly go to private hospitals.<sup>[16,19]</sup> Any calculation of incidence rates needs to be specific according to ethnic group, if data are available.<sup>[18]</sup> For ascertainment of cases in children and adolescents, the respective paediatric oncology clinics would need to be included.

Another limitation is that mere distance to goldmines is a relatively poor predictor of uranium exposure, since factors such as uranium levels in tailings, the status of tailings deposits (vegetated or not, active or decommissioned), prevailing wind directions and resulting dust plumes, and existence and course of polluted streams are all crucial for determining the extent of possible exposure of adjacent residents. In future studies it will therefore be essential to analyse the areas of interest in terms of exposure patterns rather than simply distance to mining sites. Results extracted from >60 previous measurement surveys georeferenced and collated in a combination of GIS and VGE provide levels of uranium (and other radioactive isotopes such as radium 226) in water, soil, air and biota (including foodstuff) that can be used for detailed mapping of exposure pathways and patterns.<sup>[22]</sup> Furthermore, occupational exposure to uranium of patients who worked or are still working at the goldmines also needs to be captured.

## Conclusions

With regard to the feasibility of any future epidemiological studies on haematological malignancies, our data show that case ascertainment through CHBAC would provide good coverage of the true disease burden of the majority of the local black population and allow for recruitment of large numbers of cases. However, the data also show that active recruitment is necessary to obtain accurate residential information, including on residential mobility, which is not sufficiently well covered by the hospital records or in the cancer registry. Furthermore, information on possible occupational exposure and dietary habits as well as sources of drinking water, food, etc. are required. Data on HIV, on the other hand, also essential for any interpretation of disease clustering, are sufficiently complete in the hospital records.

Our study also shows that retrospective case collections cannot be used to link observed disease patterns to environmental uranium exposure, as the clinical records do not provide sufficiently good address data and do not reflect population-based incidence rates owing to patterns of referral to the hospital. However, some of the areas observed to have an elevated incidence may not be fully explained by factors such as referral patterns or poor population registration. Acknowledging all the above limitations, we found that high incidence rates were all observed in proximity to historical or currently operating goldmines. As they were also in close proximity to CHBAH, these two factors cannot be disentangled on the basis of our data alone. Nevertheless, this finding suggests that further research on environmental contamination and the risk of haematological malignancies in this area is warranted.

**Acknowledgements.** Ms Natasha Morris and Dr Nisha Naicker of the South African Medical Research Council contributed to the data management for the study. Dr Sarah Schonfeld started this work while working for IARC and therefore contributed to the design and conduct of the study. Mr Emile Hoffmann (Mine Water Research Group, North-West University) assisted with GIS/VGE-related work.

**Author contributions.** TI contributed to the acquisition of the data, carried out the analysis and interpretation of the data and drafted the

manuscript. MP contributed to the acquisition and interpretation of the data and helped draft the manuscript and revise it critically for important intellectual content. CE helped with interpreting the data, drafting the manuscript and revising it critically for important intellectual content, and finalising it. AP contributed to the data extraction from medical files and helped draft the manuscript. MJ contributed to the acquisition of the data and supervision of data extraction and helped draft the manuscript. FW built up the maps, contributed to the interpretation of the data and helped with drafting the manuscript and revising it critically for important intellectual content. JS conceived the study and contributed to interpreting the data, drafting the manuscript and revising it critically for important intellectual content, and finalising it. All authors read and approved the final manuscript.

**Funding.** This study was partially supported by a grant made to the IARC by the Ministry of Health, Labour and Welfare of Japan (grant agreement no. 2012-02-21-01). The funding source had no involvement in the content or preparation of the manuscript.

**Conflicts of interest.** None.

1. International Agency for Research on Cancer. A Review of Human Carcinogens. IARC Monographs Radiation Volume 100D. Lyon: IARC, 2012.
2. United Nations Scientific Committee on the Effects of Atomic Radiation. UNSCEAR 2016 Report to the General Assembly: Sources, Effects and Risks of Ionizing Radiation (Annex D: Biological effects of selected internal emitters – Uranium). Vienna: UNSCEAR, 2016.
3. Brugge D, Buchner V. Health effects of uranium: New research findings. *Rev Environ Health* 2011;26(4):231-249. <https://doi.org/10.1515/REVEH.2011.032>
4. Canu JG, Laurent O, Pires N, Laurier D, Dublineau I. Health effects of naturally radioactive water ingestion: The need for enhanced studies. *Environ Health Perspect* 2011;119(12):1676-1680. <https://doi.org/10.1289/ehp.1003224>
5. Winde F. Uranium pollution of the Wonderfonteinpruit, 1997 - 2008: Part 1. Uranium toxicity; regional background and mining-related sources of uranium pollution. *Water SA* 2010;36(3):239-256. [www.scielo.org.za/pdf/wsa/v36n3/a06v36n3.pdf](http://www.scielo.org.za/pdf/wsa/v36n3/a06v36n3.pdf) (accessed 7 September 2018).
6. Barthel R. Radiological impact assessment of mining activities in the Wonderfonteinpruit catchment area, South Africa. In: Merkel B, Schipek M, eds. *The New Uranium Mining Boom*. Berlin: Springer, 2012:517-527. [https://doi.org/10.1007/978-3-642-22122-4\\_60](https://doi.org/10.1007/978-3-642-22122-4_60)
7. Nyanza EC, Joseph M, Premji SS, Thomas DS, Mannion C. Geophagy practices and the content of chemical elements in the soil eaten by pregnant women in artisanal and small scale gold mining communities in Tanzania. *BMC Pregnancy Childbirth* 2014;14:144. <https://doi.org/10.1186/1471-2393-14-144>
8. Lar UA, Agene JI, Umar AI. Geophagic clay materials from Nigeria: A potential source of heavy metals and human health implications in mostly women and children who practice it. *Environ Geochem Health* 2015;37(2):363-375. <https://doi.org/10.1007/s10653-014-9653-0>
9. Schonfeld SJ, Winde F, Albrecht C, et al; workshop participants. Health effects in populations living around the uraniumiferous gold mine tailings in South Africa: Gaps and opportunities for research. *Cancer Epidemiol* 2014;38(5):628-632. <https://doi.org/10.1016/j.canep.2014.06.003>
10. National Cancer Registry, National Institute for Occupational Health, National Health Laboratory Service. *Cancer in South Africa 2000 - 2011 Full Report*. [http://www.nioh.ac.za/?page=national\\_cancer\\_registry&id=41](http://www.nioh.ac.za/?page=national_cancer_registry&id=41) (accessed 14 September 2016).
11. Segi M. *Cancer Mortality for Selected Sites in 24 countries (1950 - 1957)*. Sendai, Japan: Department of Public Health, Tohoku University of Medicine, 1960.
12. Wiggill TM, Mantina H, Willem B, Perner Y, Stevens WS. Changing pattern of lymphoma subgroups at a tertiary academic complex in a high-prevalence HIV setting: A South African perspective. *J Acquir Immune Defic Syndr* 2011;56(5):460-466. <https://doi.org/10.1097/QAI.0b013e31820bb06a>
13. Dal Maso L, Franceschi S. Epidemiology of non-Hodgkin lymphomas and other haemolymphopoietic neoplasms in people with AIDS. *Lancet Oncol* 2003;4(2):110-119. [https://doi.org/10.1016/S1473-2045\(03\)00983-5](https://doi.org/10.1016/S1473-2045(03)00983-5)
14. Jacobson CA, Abramson JS. HIV-associated Hodgkin's lymphoma: Prognosis and therapy in the era of cART. *Adv Hematol* 2012;2012:507257. <https://doi.org/10.1155/2012/507257>
15. Patel M, Philip V, Fazel F. Human immunodeficiency virus infection and Hodgkin's lymphoma in South Africa: An emerging problem. *Adv Hematol* 2011;2011:578163. <https://doi.org/10.1155/2011/578163>
16. Dickens C, Joffe M, Jacobson J, et al. Stage at breast cancer diagnosis and distance from diagnostic hospital in a periurban setting: A South African public hospital case series of over 1,000 women. *Int J Cancer* 2014;135(9):2173-2182. <https://doi.org/10.1002/ijc.28861>
17. Statistics South Africa. Mid-year population estimates 2014. Statistical release P0302. Pretoria: Stats SA, 2014. <https://www.statssa.gov.za/publications/P0302/P03022014.pdf> (accessed 4 February 2016).
18. Gopal S, Wood WA, Lee SJ, et al. Meeting the challenge of hematologic malignancies in sub-Saharan Africa. *Blood* 2012;119(22):5078-5087. <https://doi.org/10.1182/blood-2012-02-387092>
19. Schonfeld SJ, Erdmann F, Wiggill T, et al. Hematologic malignancies in South Africa 2000 - 2006: Analysis of data reported to the National Cancer Registry. *Cancer Med* 2016;5(4):728-738. <https://doi.org/10.1002/cam4.597>
20. Waxman AJ, Mink PJ, Devesa SS, et al. Racial disparities in incidence and outcome in multiple myeloma: A population-based study. *Blood* 2010;116(25):5501-5506. <https://doi.org/10.1182/blood-2010-07-298760>
21. Shirley MH, Sayeed S, Barnes I, Finlayson A, Ali R. Incidence of haematological malignancies by ethnic group in England, 2001 - 2007. *Br J Haematol* 2013;163(4):465-477. <https://doi.org/10.1111/bjh.12562>
22. Winde F, Hoffmann E. Virtual geographical environments (VGE) as a tool to map human exposure to mining-related radionuclides. In: Merkel BJ, Arab A, eds. *Uranium - Past and Future Challenges: Proceedings of the 7th International Conference on Uranium Mining and Hydrogeology*. Heidelberg, New York, Dordrecht, London: Springer International Publishing Switzerland, 2014:192-200. <https://doi.org/10.1007/978-3-319-11059-2>

Accepted 3 April 2018.